

Anwai C. Ku · Kingsley Davis · Wilbert E. Moore · Melvin M. Tumin · Claude Fischer · Michael Hout · Martín Sánchez Jankowski · Samuel R. Lucas · Ann Swidler · Kim Ross · Alan B. Krueger · Martina Morris · Bruce Western · Thomas Picketty · Emmanuel Saez · Karl Marx · Ralf Dahrendorf · Erik Olin Wright · Immanuel Wallerstein · Max Weber · Anthony Giddens · Frank Parkin · Emile Durkheim · Jesper B. Sørensen · Peter M. Blau · Otis Dudley Duncan · Andrea Tyree · Donald J. Treiman · Keith Hope · David L. Featherman · Robert M. Hauser · Robert W. Hodge · John Robert Warren · Aage Sørensen · Amartya Sen · Kim Weeden · Gaetano Mosca · C. Wright Mills · G. William Domhoff · Alvin W. Gouldner · David Brooks · Gilval · Iván Széleányi · Eleanor Townsley · Barbara Ehrenreich · Timothy Smeeding · William Julius Wilson · Douglas S. Massey · Nancy A. Denton · Anne R. Pebley · Narayan Sastry · Robert J. Sampson · Jeffrey D. Morenoff · James Rosenbaum · Stefanie DeLuca · Elizabeth Clark-Kauffman · Greg J. Duncan · Pamela Morris · Sharon Hays · Ralph H. Turner · Robert Erikson · John H. Aldthorpe · Richard Breen · Gary Solon · Christopher Jencks · Marshall Smith · Henry Acland · Mary Jo Bane · David Cohen · Herbert Gintis · Barbara Heyns · Stephan Michelson · David Harding · Leonard M. Lopoo · Susan E. Mayer · William H. Sewell · Archibald O. Haller ·

THIRD EDITION

SOCIAL STRATIFICATION

*Class, Race, and Gender
in Sociological
Perspective*

Robert Erikson · John H. Goldthorpe · Richard Breen · Gary Solon · Christopher Jencks · Marshall Smith · Henry Acland · Mary Jo Bane · David Cohen · Herbert Gintis · Barbara Heyns · Stephan Michelson · David J. Harding · Leonard M. Lopoo · Susan E. Mayer · William H. Sewell · Archibald O. Haller · Alejandro Portes · Jay MacLeod · Stephen Morgan · Michael J. Piore · Arne L. Kalleberg · Mark S. Granovetter · Nan Lin · Ronald S. Burt · Roberto M. Fernandez · Isabel Fernandez-Mateo · Dalton Conley · Reynolds Farley · Edna Bonacich · Robert D. Manning · Min Zhou · Mary C. Waters · Marianne Bertrand · Sendhil Mullainathan · Claude Steele · Leah Pager · J. Morgan Kousser · David P. Taub · Arlie Hochschild · Charles Rouse · Shelley Correll · Stephen Benard · In Paik · Barbara Reskin · Trond Petersen · William T. Bielby · Perry A. Jacobs · Asaf Levanon · Laurie A. Morgan · Paula England · Tony Tam · Francine Law · Lawrence M. Kahn · Thorstein Veblen · Pierre Bourdieu · Melvin Kohn · John Mullahy · Stephanie Robert · Barbara Wolfe · Tak Wing Chan · Annette Lareau · Eszter Hargittai · Benjamin Moodie · Clark Kerr · John T. Dunlop · Frederick H. Harbison · Charles A. Myers · Daniel Bell · Gøsta Esping-Andersen · Victor Nee · John W. Meyer · Jan Pakulski · Malcolm Waters · Ulrich Beck · Christoph Lau · Joseph E. Stiglitz · Glenn Firebaugh · Jitka Malečková

EDITED BY DAVID B. GRUSKY

Social Stratification

THIRD EDITION

Social Stratification

Class, Race, and Gender
in Sociological Perspective

EDITED BY

David B. Grusky

(in collaboration with
Manwai C. Ku and Szonja Szelényi)

Stanford University



Westview Press

A Member of the Perseus Books Group

Copyright © 2008 by Westview Press

Published by Westview Press,
A Member of the Perseus Books Group

All rights reserved. Printed in the United States of America. No part of this book may be reproduced in any manner whatsoever without written permission except in the case of brief quotations embodied in critical articles and reviews. For information, address Westview Press, 2465 Central Avenue, Boulder, Colorado 80301.

Find us on the World Wide Web at www.westviewpress.com.

Westview Press books are available at special discounts for bulk purchases in the United States by corporations, institutions, and other organizations. For more information, please contact the Special Markets Department at the Perseus Books Group, 2300 Chestnut Street, Suite 200, Philadelphia, PA 19103, or call (800) 255-1514, or e-mail special.markets@perseusbooks.com.

Designed by Brent Wilcox

Library of Congress Cataloging-in-Publication Data

Social stratification : class, race, and gender in sociological perspective /
edited by David B. Grusky, in collaboration with Manwai C. Ku and Szonja Szelényi.
— 3rd ed.

p. cm.

Includes bibliographical references and index.

ISBN-13: 978-0-8133-4373-0 (pbk. : alk. paper)

ISBN-10: 0-8133-4373-9 (pbk. : alk. paper)

1. Social stratification. 2. Equality. 3. Social status. 4. Social classes.

I. Grusky, David B. II. Ku, Manwai C. III. Szelényi, Szonja, 1960-

HM821.S65 2008

305.5'12—dc22

2007040457

10 9 8 7 6 5 4 3 2 1

To the Elfenworks Foundation

CONTENTS

Study Guide xv
Preface and Acknowledgments xvii

Part I Introduction

1 ■ David B. Grusky and Manwai C. Ku
Gloom, Doom, and Inequality 2

Part II Forms and Sources of Inequality

THE FUNCTIONS OF INEQUALITY

2 ■ Kingsley Davis and Wilbert E. Moore
Some Principles of Stratification 30

3 ■ Gerhard Lenski
New Light on Old Issues: The
Relevance of “Really Existing Socialist
Societies” for Stratification Theory ... 33

THE DYSFUNCTIONS OF INEQUALITY

4 ■ Melvin M. Tumin
Some Principles of Stratification:
A Critical Analysis 41

5 ■ Claude S. Fischer, Michael Hout,
Martín Sánchez Jankowski, Samuel
R. Lucas, Ann Swidler, and Kim Voss
Inequality by Design 49

6 ■ Alan B. Krueger
Inequality, Too Much of a Good
Thing 53

WHY IS INCOME INEQUALITY INCREASING?

7 ■ Martina Morris and Bruce
Western
Inequality in Earnings: Trends and
Implications 61

8 ■ Thomas Piketty and Emmanuel
Saez
The Evolution of Top Incomes:
A Historical and International
Perspective 67

Part III The Structure of Contemporary Inequality

THEORIES OF CLASS

Marx and Post-Marxists

9 ■ Karl Marx
Alienation and Social Classes 74
Classes in Capitalism and
Pre-Capitalism 79

Ideology and Class 89

10 ■ Ralf Dahrendorf

Class and Class Conflict in
Industrial Society 91

| | |
|--|-----|
| 11 ■ Erik Olin Wright | |
| A General Framework for the | |
| Analysis of Class Structure | 98 |
| 12 ■ Immanuel Wallerstein | |
| Class Conflict in the Capitalist | |
| World Economy | 111 |
| Weber and Post-Weberians | |
| 13 ■ Max Weber | |
| Class, Status, Party | 114 |
| Status Groups and Classes | 124 |
| Open and Closed Relationships | 128 |
| 14 ■ Anthony Giddens | |
| The Class Structure of the | |
| Advanced Societies | 132 |
| 15 ■ Frank Parkin | |
| Marxism and Class Theory: | |
| A Bourgeois Critique | 143 |
| Durkheim and | |
| Post-Durkheimians | |
| 16 ■ Emile Durkheim | |
| The Division of Labor in Society | 159 |
| 17 ■ David B. Grusky and Jesper B. | |
| Sørensen | |
| Are There Big Social Classes? | 165 |
| 18 ■ Kim A. Weeden | |
| Occupational Closure and | |
| Earnings Inequality | 176 |
| GRADATIONAL STATUS GROUPINGS | |
| Classic Statements | |
| 19 ■ Peter M. Blau and Otis Dudley | |
| Duncan, with the collaboration of | |
| Andrea Tyree | |
| Measuring the Status of | |
| Occupations | 187 |
| Part IV Inequality at the Extremes | |
| THE RULING CLASS, ELITES, AND | |
| THE UPPER CLASS | |
| Classic Statements | |
| 28 ■ Gaetano Mosca | |
| The Ruling Class | 268 |
| 20 ■ Donald J. Treiman | |
| Occupational Prestige in | |
| Comparative Perspective | 191 |
| 21 ■ John H. Goldthorpe and Keith | |
| Hope | |
| Occupational Grading and | |
| Occupational Prestige | 195 |
| 22 ■ David L. Featherman and | |
| Robert M. Hauser | |
| Prestige or Socioeconomic Scales | |
| in the Study of Occupational | |
| Achievement? | 202 |
| Are Socioeconomic Scales Obsolete? | |
| 23 ■ Robert W. Hodge | |
| The Measurement of | |
| Occupational Status | 205 |
| 24 ■ Robert M. Hauser and John | |
| Robert Warren | |
| Socioeconomic Indexes for | |
| Occupations: A Review, Update, | |
| and Critique | 213 |
| The New Gradationalism? | |
| 25 ■ Aage B. Sørensen | |
| Foundations of a Rent-Based | |
| Class Analysis | 219 |
| 26 ■ Amartya K. Sen | |
| From Income Inequality to | |
| Economic Inequality | 235 |
| 27 ■ Kim A. Weeden, Young-Mi Kim, | |
| Matthew Di Carlo, and David B. | |
| Grusky | |
| Is the Labor Market Becoming | |
| More or Less Gradational? | 249 |
| 29 ■ C. Wright Mills | |
| The Power Elite | 275 |
| 30 ■ Anthony Giddens | |
| Elites and Power | 285 |

Contemporary Statements

- 31 ■ G. William Domhoff**
Who Rules America? Power and
Politics290
- 32 ■ Alvin W. Gouldner**
The Future of Intellectuals and
the Rise of the New Class295
- 33 ■ David Brooks**
Bobos in Paradise: The New Upper
Class and How They Got There ...304
- 34 ■ Gil Eyal, Iván Széleányi, and
Eleanor Townsley**
Post-Communist Managerialism ...311

POVERTY AND THE UNDERCLASS

The Experience of Poverty

- 35 ■ Barbara Ehrenreich**
Nickel-and-Dimed: On (not)
Getting by in America317

How Much Poverty is There?

- 36 ■ Timothy M. Smeeding**
Poverty, Work, and Policy:
The United States in
Comparative Perspective327

**The Sources of Poverty and
the Underclass**

- 37 ■ William Julius Wilson**
Jobless Poverty: A New Form of
Social Dislocation in the
Inner-City Ghetto340

Part V Generating Inequality

SOCIAL MOBILITY

Class Mobility

- 45 ■ Ralph H. Turner**
Sponsored and Contest Mobility
and the School System420
- 46 ■ David L. Featherman and
Robert M. Hauser**
A Refined Model of Occupational
Mobility426

- 38 ■ Douglas S. Massey and Nancy
A. Denton**
American Apartheid: Segregation
and the Making of the Underclass ...349

- 39 ■ Anne R. Pebley and Narayan
Sastry**
Neighborhoods, Poverty, and
Children's Well-Being360

- 40 ■ Robert J. Sampson and Jeffrey
D. Morenoff**
Durable Inequality: Spatial
Dynamics, Social Processes, and
the Persistence of Poverty in
Chicago Neighborhoods371

Poverty and Social Policy

- 41 ■ Bruce Western**
Incarceration, Unemployment,
and Inequality388

- 42 ■ James E. Rosenbaum and
Stefanie DeLuca**
Does Changing Neighborhoods
Change Lives? The Chicago
Gautreaux Housing Program393

- 43 ■ Elizabeth Clark-Kauffman, Greg
J. Duncan, and Pamela Morris**
How Welfare Policies Affect Child
and Adolescent Achievement400

- 44 ■ Sharon Hays**
Flat Broke with Children: Women
in the Age of Welfare Reform407

- 47 ■ Robert Erikson and John H.
Goldthorpe**
Trends in Class Mobility: The
Post-War European Experience ...437

- 48 ■ Richard Breen**
Social Mobility in Europe465

Income Mobility

- 49 ■ Gary Solon**
Intergenerational Income
Mobility481

STATUS AND INCOME ATTAINMENT

Basic Models

50 ■ Peter M. Blau and Otis Dudley Duncan, with the collaboration of Andrea Tyree
The Process of Stratification486

51 ■ Christopher Jencks, Marshall Smith, Henry Acland, Mary Jo Bane, David Cohen, Herbert Gintis, Barbara Heyns, and Stephan Michelson
Inequality: A Reassessment of the Effect of Family and Schooling in America498

52 ■ David J. Harding, Christopher Jencks, Leonard M. Lopoo, and Susan E. Mayer
Family Background and Income in Adulthood, 1961–1999505

Social Psychological Models

53 ■ William H. Sewell, Archibald O. Haller, and Alejandro Portes
The Educational and Early Occupational Attainment Process516

54 ■ Jay MacLeod
Ain't No Makin' It: Leveled Aspirations in a Low-Income Neighborhood528

55 ■ Stephen L. Morgan
A New Social Psychological Model of Educational Attainment542

Labor Markets

56 ■ Michael J. Piore
The Dual Labor Market: Theory and Implications550

57 ■ Aage B. Sørensen and Arne L. Kalleberg
An Outline of a Theory of the Matching of Persons to Jobs553

58 ■ Arne L. Kalleberg
Nonstandard Employment Relations and Labour Market Inequality: Cross-national Patterns562

Social Capital, Networks, and Attainment

59 ■ Mark S. Granovetter
The Strength of Weak Ties576

60 ■ Nan Lin
Social Networks and Status Attainment580

61 ■ Ronald S. Burt
Structural Holes583

62 ■ Roberto M. Fernandez and Isabel Fernandez-Mateo
Networks, Race, and Hiring587

Sibling Models

63 ■ Dalton Conley
What Do Low (or High) Sibling Correlations Tell Us About Social Ascription?596

Rational Action Approaches

64 ■ Richard Breen and John H. Goldthorpe
Explaining Educational Differentials: Towards a Formal Rational Action Theory610

Part VI Race and Ethnicity

CONSTRUCTING RACIAL CATEGORIES

- 65 ■ **Reynolds Farley**
Racial Identities in 2000:
The Response to the
Multiple-Race Response Option . . .624

MODES OF INCORPORATION

Classic Modes

- 66 ■ **Edna Bonacich**
A Theory of Ethnic Antagonism:
The Split Labor Market632
- 67 ■ **Alejandro Portes and Robert D. Manning**
The Immigrant Enclave: Theory
and Empirical Examples646

New Modes

- 68 ■ **Alejandro Portes and Min Zhou**
The New Second Generation:
Segmented Assimilation and
Its Variants658
- 69 ■ **Mary C. Waters**
Black Identities: West Indian
Immigrant Dreams and
American Realities669

DISCRIMINATION

- 70 ■ **Marianne Bertrand and Sendhil Mullainathan**
Are Emily and Greg More
Employable Than Lakisha and
Jamal? A Field Experiment on
Labor Market Discrimination673

Part VII Gender Inequality

LABOR FORCE PARTICIPATION

- 77 ■ **Arlie Russell Hochschild**
The Time Bind: When Work
Becomes Home and Home
Becomes Work730
- 78 ■ **Lisa Belkin**
The Opt-Out Revolution735

- 71 ■ **Claude Steele**
Stereotype Threat and African-
American Student Achievement678
- 72 ■ **Devah Pager**
Marked: Race, Crime, and
Finding Work in an Era of Mass
Incarceration683

ARE RACIAL AND ETHNIC
DISTINCTIONS DECLINING IN
SIGNIFICANCE?

- 73 ■ **William Julius Wilson**
The Declining Significance of
Race: Blacks and Changing
American Institutions691
- 74 ■ **Joe R. Feagin**
The Continuing Significance of
Race: Antiblack Discrimination
in Public Places703
- 75 ■ **Melvin L. Oliver and Thomas M. Shapiro**
Black Wealth/White Wealth: A New
Perspective on Racial Inequality709
- 76 ■ **William J. Wilson and Richard P. Taub**
There Goes the Neighborhood: Race,
Ethnic, and Class Tensions in Four
Chicago Neighborhoods and Their
Meaning in America716

- 79 ■ **Heather Boushey**
Is the Opt-Out Hypothesis
Convincing?739

DISCRIMINATION

- 80 ■ **Claudia Goldin and Cecilia Rouse**
Orchestrating Impartiality:
The Impact of “Blind” Auditions
on Female Musicians747

81 ■ Shelley J. Correll, Stephen Benard, and In Paik
 Getting a Job: Is There a Motherhood Penalty?759

82 ■ Barbara F. Reskin
 Rethinking Employment
 Discrimination and Its Remedies . . .770

83 ■ Trond Petersen
 Discrimination: Conscious or Nonconscious?780

SEX SEGREGATION

84 ■ William T. Bielby
 The Structure and Process of Sex Segregation786

85 ■ Jerry A. Jacobs
 Revolving Doors: Sex Segregation and Women’s Careers797

86 ■ Barbara F. Reskin
 Labor Markets as Queues: A Structural Approach to Changing Occupational Sex Composition802

87 ■ David B. Grusky and Asaf Levanon
 Four Gloomy Futures for Sex Segregation812

GENDER GAP IN WAGES

88 ■ Trond Petersen and Laurie A. Morgan
 The Within-Job Gender Wage Gap826

89 ■ Paula England
 Devaluation and the Pay of Comparable Male and Female Occupations834

90 ■ Tony Tam
 Why Do Female Occupations Pay Less?838

91 ■ Francine D. Blau and Lawrence M. Kahn
 The Gender Pay Gap: Have Women Gone as Far as They Can? . . .843

Part VIII The Consequences of Inequality

THEORY

92 ■ Thorstein Veblen
 The Theory of the Leisure Class862

93 ■ Pierre Bourdieu
 Distinction: A Social Critique of the Judgement of Taste870

94 ■ Melvin L. Kohn
 Job Complexity and Adult Personality893

HEALTH

95 ■ John Mullahy, Stephanie Robert, and Barbara Wolfe
 Health, Income, and Inequality904

CULTURE

96 ■ Tak Wing Chan and John H. Goldthorpe
 The Social Stratification of Theatre, Dance, and Cinema Attendance916

97 ■ Annette Lareau
 Unequal Childhoods: Class, Race, and Family Life926

98 ■ Eszter Hargittai
 The Digital Reproduction of Inequality936

POLITICS

99 ■ Michael Hout and Benjamin Moodie
 The Realignment of U.S. Presidential Voting, 1948–2004945

Part IX The Future of Inequality

THEORIES OF INDUSTRIALISM AND MODERNITY

- 100 ■ **Clark Kerr, John T. Dunlop, Frederick H. Harbison, and Charles A. Myers**
Industrialism and Industrial Man . . .954

THEORIES OF POST-INDUSTRIALISM, POST-SOCIALISM, AND POST-MODERNITY

Post-Industrialism and the New Class

- 101 ■ **Daniel Bell**
The Coming of Post-Industrial Society966
- 102 ■ **Gøsta Esping-Andersen**
Social Foundations of Postindustrial Economies978

Post-Socialism

- 103 ■ **Victor Nee**
Postsocialist Stratification994
- 104 ■ **Gil Eyal, Iván Szelényi, and Eleanor Townsley**
Making Capitalism without Capitalists1000

Post-Modernity and High Modernity

- 105 ■ **John W. Meyer**
The Evolution of Modern Stratification Systems1006
- 106 ■ **Anthony Giddens**
Social Justice and Social Divisions1016
- 107 ■ **Jan Pakulski and Malcolm Waters**
The Death of Class1022
- 108 ■ **Ulrich Beck and Christoph Lau**
Theoretical and Empirical Explorations in the ‘Meta-Change’ of Modern Society1030

GLOBALIZATION AND INEQUALITY

- 109 ■ **Joseph E. Stiglitz**
Globalism’s Discontents1036
- 110 ■ **Glenn Firebaugh**
The New Geography of Global Income Inequality1044
- 111 ■ **Alan B. Krueger and Jitka Malečková**
Does Poverty Cause Terrorism? The Economics and the Education of Suicide Bombers1056

- About the Editors1063*
- About the Book and Our New Series of Inequality Readers1065*
- Index1067*

STUDY GUIDE

The chapters of *Social Stratification* were designed to be used in conjunction with additional study materials available at www.inequality.com. This site provides the latest trend data, the latest working papers by top scholars in the field, and additional reporting on new and emerging developments relevant to issues of poverty and inequality.

In assembling *Social Stratification*, every effort was made to select articles that were both path-breaking *and* readable, yet on occasion it proved necessary to compromise on either of these two objectives. The following chapters, in particular, rest on concepts or methods that might be challenging to some undergraduate students:

Karl Marx, *Alienation and Social Classes*

Aage B. Sørensen, *Foundations of a Rent-Based Class Analysis*

Robert J. Sampson and Jeffrey D. Morenoff, *Durable Inequality: Spatial Dynamics, Social Processes, and the Persistence of Poverty in Chicago Neighborhoods*

David L. Featherman and Robert M. Hauser, *A Refined Model of Occupational Mobility*

Richard Breen, *Social Mobility in Europe*

Gary Solon, *Intergenerational Income Mobility*

Richard Breen and John H. Goldthorpe, *Explaining Educational Differentials: Towards a Formal Rational Action Theory*

Although the foregoing chapters present materials that should be mastered by all advanced students, they can be safely excised for the purposes of a purely introductory course. The remaining readings were selected to ensure that introductory students will still be acquainted with the most important concepts, findings, and debates in the field.

PREFACE AND ACKNOWLEDGMENTS

The study of poverty and inequality is no longer a sleepy little enterprise in sociology. In the last decade, issues of poverty and inequality have not only burst into the academic mainstream, but are now routinely covered in mainstream newspapers, popular magazines, and trade books. The most obvious precipitant behind this development is the historic and unprecedented takeoff of income inequality in many countries. If the defining social problem of the postwar period was developing policy tools for increasing economic production, the emerging social problem of our time is developing an analogous machinery for bringing the unequal distribution of economic production under control.

By virtue of these developments, we had no choice but to undertake a radical revision of *Social Stratification*, even if by our slow-science standards the last edition was published quite recently. In the end, we have of course retained most of the classics appearing in the second edition, but we had to replace many of the contemporary pieces with yet newer selections that provided important extensions, revisions, and even rebuttals of prior research. We also incorporated selections from new and emerging subfields of study. The third edition is in the end nearly half-new; that is, of the 111 selections appearing in this edition, only 62 appeared also in the second edition.

The third edition is distinctive not only by virtue of such heavy updating. In addition, recent developments in the field demand a new breed of anthology that is quite unlike existing ones, even our own second edition. There are four developments, in particular, that have revolutionized the field and have motivated us to rethink the rationale behind the book. These developments, which we list below, are part of the ongoing transformation of sociology as it refashions a place for itself within an increasingly competitive social science landscape.

Rise of a multidisciplinary inequality field: For better or worse, there is no longer a consensual, interdisciplinary division of labor that has economists studying the total output of goods and sociologists studying how goods and resources are unequally distributed. Although such a simple division of labor was never obtained even in the past, clearly economists have turned to the study of poverty and inequality, once the special province of sociologists, with unprecedented vigor. As a result, sociology is no longer a dominant force in the field, and “head-in-the-sand anthologies” that continue to pretend otherwise by relying exclusively on contributions by sociologists clearly do not serve our students well. We have sought to incorporate some of the most important contributions by economists and other social scientists, contributions that have already entered into sociological discourse and have accordingly become, in effect, part of sociology. If the third edition is

now more deeply multidisciplinary, it is mainly because disciplinary boundaries are growing more porous and “hit papers” in economics and other disciplines are accordingly crossing over and becoming hits in sociology as well.

Public intellectuals and inequality: As poverty and inequality emerge as mainstream concerns, public intellectuals have become yet another new group of practicing sociologists, although in this case they have contributed mainly to the qualitative side of inequality research. The latter research has taken not only the usual small-scale form characteristic of qualitative research in sociology but also a more heavily-funded “big journalism” form in which cadres of investigators investigate a well-defined sociological question as part of a long-term research project. If one wants a provocative hypothesis, it often pays to turn to such journalists and public intellectuals. Although some of our readers might have wanted us to turn even more frequently to them (if only for their pleasing writing style), it bears emphasizing that sociologists still hold a distinct advantage in rigorous hypothesis-testing, an advantage that implies a somewhat circumscribed role for public-intellectual sociology.

The new post-survey world: Over the last decade, the standard survey instrument has come to be appreciated as but one useful approach to studying poverty and inequality, while other approaches, especially experimental and quasi-experimental ones, have clearly gained in prominence. We are referring, for example, to experimental studies of hiring discrimination, neighborhood effects, and returns to schooling. It is important, we think, to feature this experimental turn in our anthology, as it has been exceedingly influential in many subfields of poverty and inequality research. We have also drawn more heavily than in prior editions on ethnographic and mixed-method studies. In many cases, these qualitative studies have attracted much attention and become widely celebrated, and not just among a narrow academic audience.

New empirical developments: We live in an extraordinary period in which the structure of social inequality is changing in rapid, complicated, and sometimes mystifying ways. Whereas scholars in some disciplines, such as sociology, are seemingly entranced by the possibility that class-based inequalities are weakening, scholars elsewhere attend to unprecedented increases in income inequality, wealth inequality, and an often deepening spatial divide between rich and poor. The fields of race and gender inequality likewise reveal a tension between scholars who emphasize the rise of formally egalitarian attitudes and institutional practices and those who emphasize the persistence of discrimination and segregation or the emergence of anti-egalitarian backlashes that may undo gains of the past. Even as these backlashes play out, the relentless march of egalitarian sensibilities serves to expose and delegitimize “new” inequalities that, not so long ago, were taken for granted, rarely discussed, or even barely seen (e.g., inequalities based on sexual orientation, disability, citizenship). We have included many new contributions that address this constellation of trends and ask whether any unified understanding of the forces making for equality and inequality is feasible in light of them.

The revolution in poverty and inequality analysis has therefore played out in conceptual, methodological, and empirical terms. We have reviewed these developments in our introductory chapter and won’t attempt to review them again here. It should suffice to note that recent scholarship has turned to (1) examining the

dramatic increases in inequality (e.g., the spectacular growth in income inequality) as well as recent slowdowns in the rate of change (e.g., slowing rate of decline in the gender pay gap), (2) exploring new types of inequality that were either ignored in the past or have been spawned by new technologies or institutions (e.g., the “digital divide”), and (3) teasing out the implications of inequality for both individual-level (e.g., health) and macro-level (e.g., terrorism, GNP) outcomes. We have included many examples of each of these developments in this latest edition.

As the foregoing suggests, the research literature has become so large and complex that the task of reducing it to manageable form poses difficulties of all kinds, not the least of which is simply defining defensible boundaries for a field that at times seems indistinguishable from social science at large. What types of priors do we bring to that task? Perhaps most obviously, we have sought to represent the diversity of research traditions on offer, while at the same time giving precedence to those traditions that have so far borne the greatest fruit. As is often the case, the pool of disciplinary knowledge has developed in uneven and ramshackle fashion, so much so that any attempt to cover all subjects equally would grossly misrepresent the current strengths and weaknesses of contemporary stratification research. This sensitivity to disciplinary fashion reveals itself, for example, in the relatively large number of selections addressing issues of race, ethnicity, and gender. These subfields rose to a special prominence in the 1970s and continue to be popular even after 30 years of intensive and productive research. If the concepts of class, status, and power formed the holy trinity of postwar stratification theorizing, then the partly overlapping concepts of class, race, and gender continue to play analogous roles now.

We have sought, secondarily, to give precedence to contemporary research, while also appreciating that there is value in exposing students not just to the standard classics but also to near-classics that were written well after the foundational contributions of, say, Karl Marx or Max Weber. The latter body of intervening work is often ignored by editors of anthologies, thereby perpetuating (in some small way) the view that all sociological research can or should be stamped with an exclusively Marxian or Weberian imprimatur. We have of course selected those classics or near-classics that seem to best anticipate or motivate current disciplinary interests. The novice reader might be left, then, with the impression that the discipline is more straight-line cumulative than it truly is, an impression that should naturally be resisted.

The final, and most difficult, task faced by editors of anthologies is to chart an optimal course between the Scylla of overly aggressive excerpting and the Charybdis of excessive editorial timidity. By the usual standards of anthologies, the course charted here was very much an average one, as the objective was to eliminate all inessential material while still preserving the analytic integrity of the contributions. We have excised many clarifying and qualifying footnotes, almost all decorative theorizing and literature reviews, and much analysis that wasn't crucial in advancing the argument. To be sure, some of our readers and contributors would no doubt oppose *all* excerpting, yet the high cost of implementing such a radical stance would be a substantial reduction in the number of articles that could be reproduced. We apologize to our authors for being unable to present the selections in their entirety and encourage our readers to consult the original and full versions of our excerpted pieces.

The editing rules adopted throughout this anthology were in most cases conventional. For example, brackets were used to mark off a passage that was inserted for the purpose of clarifying meaning, whereas ellipses were used whenever a passage appearing in the original contribution was excised altogether. The latter convention was violated, however, if the excised text was a footnote, a minor reference to a table or passage (e.g., “see table 1”) that was itself excerpted out, or otherwise inconsequential. When necessary, tables and footnotes were renumbered without so indicating in the text, and all articles that were cited in excised passages were likewise omitted, without indication, from the list of references appearing at the end of each chapter.

The spelling, grammar, and stylistic conventions of the original contributions were otherwise preserved. In this respect, the reader should be forewarned that some of the terms appearing in the original contributions would now be regarded as inappropriate or offensive, whereas others have passed out of common usage and will possibly be unfamiliar. Although a strong argument could clearly be made for eliminating all language that is no longer acceptable, this type of sanitizing would not only exceed usual editorial license but would also generate a final text that contained inconsistent and possibly confusing temporal cues.

Among the various functions that an anthology fills, one of the more obvious ones is to define and celebrate what a field has achieved, and in so doing to pay tribute to those who made such achievement possible. We are grateful to the many dozens of scholars who allowed their work to be reproduced for this anthology or who agreed to write one of the many commissioned essays. This book provides a well-deserved occasion to recognize the many successes of a field that is perhaps better known for its contentiousness and controversy.

The selections reproduced here have for the most part been pre-tested in graduate and undergraduate inequality classes at the University of Chicago, Cornell University, and Stanford University. We are indebted to the many instructors and students who shared their reactions to the selections and thereby shaped the final product more than they may realize. The anonymous reviewers of an earlier proposal for this book also provided unusually constructive criticisms that were helpful in assembling the final book.

We are equally grateful for the formal market research carried out by Karl Yambert, for Michelle Trader’s careful work securing permissions to reprint, for Erica Lawrence’s excellent advertising and marketing support, and for Carol Smith’s and Annie Lenth’s patient overseeing of a complicated, fast, and high-pressure production process. Finally, we must thank our Westview Press Editor, Steve Catalano, for choreographing the entire process so expertly and somehow inducing us to produce far more quickly than usual, all the while maintaining a superhuman, calming presence that is typically found only in monks and yoga instructors.

David B. Grusky, Manwai C. Ku, and Szonja Szelényi
Palo Alto, California 2007

PART I

Introduction

1 ■ David B. Grusky and Manwai C. Ku

Gloom, Doom, and Inequality

There is a growing consensus among academics, policy makers, and even politicians that poverty and inequality should no longer be treated as soft “social issues” that can safely be subordinated to more fundamental interests in maximizing total economic output. The most important sources of this newfound concern with poverty and inequality are (1) the spectacular increase in economic inequality and other forms of disadvantage in many late-industrial countries (the *takeoff account*); (2) the striking persistence of other forms of inequality (e.g., racially segregated neighborhoods, occupational sex segregation) despite decades of quite aggressive egalitarian reform (the *persistence account*); (3) an emerging concern that poverty and inequality may have negative macro-level effects on terrorism, total economic production, and ethnic unrest (the *macro-level externalities account*); (4) a growing awareness of the negative individual-level effects of poverty on health, political participation, and a host of other life conditions (the *micro-level externalities account*); (5) the rise of a “global village” in which spatial disparities in the standard of living have become more widely visible and hence increasingly difficult to ignore (the *visibility account*); (6) the ongoing tendency to expose and delegitimize new types of inequalities (based on sexual orientation, disability, or cit-

izenship) that, not so long ago, were taken for granted, rarely discussed, and barely seen (the *new inequalities account*); and (7) a growing commitment to a broader conception of human entitlements that encompasses rights to basic social amenities (e.g., housing) as well as rights to basic forms of social participation, such as employment (the *social inclusion account*).

This list is remarkable in two ways. First, only two of the seven reasons for our newfound interest in poverty and inequality are about brute empirics (i.e., the growth or intransigence of inequality), while all others are about changes in how we have come to view, study, and evaluate those empirics. When scholars now argue, for example, that inequality has multifarious unintended effects (i.e., an *externalities account*), they presumably don't mean to suggest that such effects suddenly multiplied in the contemporary period (although perhaps they have). Rather, we are to understand that inequality was always rife with externalities, however inadequately we may have appreciated them in the past. While changes in empirics hardly exhaust, then, the sources of our growing concern with inequality, this is not to gainsay the equally important point that such changes, especially the recent takeoff in income inequality, are likely a core reason why inequal-

This is a commissioned chapter that draws heavily on material in previous publications (David B. Grusky and Szonja Szelényi, “The Rise and Fall of Benign Narratives About Inequality,” David B. Grusky and Szonja Szelényi, eds., *The Inequality Reader: Contemporary and Foundational Readings in Race, Class, and Gender*, 2007 by Westview Press; David B. Grusky, “The Past, Present, and Future of Social Inequality,” David B. Grusky, ed., *Social Stratification: Class, Race, and Gender in Sociological Perspective*, 2nd edition 2001 by Westview Press).

ity has come to be understood as a fundamental social problem of our time.

The above list is no less remarkable for the relatively minor role that normative concerns play. To be sure, there appears to be a growing sentiment that, at minimum, contemporary social systems should guarantee an opportunity to all citizens to participate in economic life and hence avoid the most extreme forms of social and economic exclusion (i.e., the *social inclusion account*). It would nonetheless be a mistake to understand the rising interest in poverty and inequality as principally fueled by some sudden realization that social inclusion is a fundamental social good. Indeed, far from treating inequality as exclusively a moral problem, there is a growing tendency to emphasize its profound consequences and threats for the world community as a whole (i.e., the *macro-level externalities account*). The rhetoric of “sustainability,” although more frequently featured in discussions of environmental problems, is increasingly taken as relevant to discussions of inequality as well (see, in this volume, Mullahy, Robert, and Wolfe, Ch. 95; Stiglitz, Ch. 109; Krueger and Malečková, Ch. 111). In adopting this rhetoric, the claim is that extreme inequality is counterproductive not just because it reduces total economic output, but also because other very legitimate objectives, such as reducing mortality rates or the threat of terrorism, might be compromised if inequality remains so extreme. By this logic, social policy must simultaneously be oriented to increasing economic output *and* restraining the rise of debilitating and counterproductive forms of inequality, a rather more complicated maximization problem than that conventionally taken on (Fischer et al., Ch. 5; Krueger, Ch. 6).

Core Fields of Inquiry

The foregoing is to suggest that the main task of stratification scholarship, that of describing the contours of inequality and explaining its

causes, has come to be viewed as an increasingly important and central endeavor, and not just within the narrow world of academia. The purpose of this introductory essay is to lay out how the field of stratification has developed, where it currently stands, and how this volume represents and conveys that field.

It may be useful to begin such a task by specifying the main questions around which the field has to this point developed. For each of the eight subfields listed below, we provide some examples of the types of questions that have been featured, our intent being to convey the prevailing style of inquiry rather than some more exhaustive listing.

Forms and Sources of Stratification

What have been the major forms of inequality in human history? Can the ubiquity of inequality be attributed to individual differences in talent or ability (e.g., Fischer et al., Ch. 5)? Is some amount of inequality an inevitable feature of human life (e.g., Davis and Moore, Ch. 2; Tumin, Ch. 4)? Why is income inequality increasing in so many late-industrial countries (e.g., Morris and Western, Ch. 7; Piketty and Saez, Ch. 8)?

Structure of Contemporary Inequality

Is inequality organized into a small number of social classes or occupational groups (e.g., Grusky and Sørensen, Ch. 17)? Or does it take on a largely gradational form featuring wholly incremental differences of income or status (e.g., Blau and Duncan, Ch. 19; Hauser and Warren, Ch. 24; Sen, Ch. 26)? Which of these two forms became more prominent with the transition into modernity or late modernity (e.g., Weeden et al., Ch. 27)?

Inequality at the Extremes

Do political, economic, and cultural elites come together to form a single “upper class” (e.g., Mosca, Ch. 28; Mills, Ch. 29; Giddens, Ch. 30)? Is there likewise a well-formed “underclass” of the poor and dispossessed (e.g., Wilson, Ch. 37)? Are elites principally drawn

from long-standing elite families and the poor from long-standing poverty-stricken families (e.g., Domhoff, Ch. 31)?

Generating Inequality

How frequently do individuals move into new classes, occupations, or income groups (e.g., Breen, Ch. 48)? Is the United States an especially mobile society (e.g., Erikson and Goldthorpe, Ch. 47)? To what extent are occupational outcomes determined by such forces as intelligence, effort, schooling, aspirations, social contacts, and individual luck (e.g., Jencks et al., Ch. 51; Harding et al., Ch. 52; Granovetter, Ch. 59; Fernandez and Fernandez-Mateo, Ch. 62)?

Race and Ethnicity

What accounts for racial and ethnic differences in grades, test scores, high school graduation, and college attendance (e.g., Steele, Ch. 71)? What accounts for racial and ethnic differences in hiring, promotion, and pay (e.g., Bertrand and Mullainathan, Ch. 70)? Why are neighborhoods so deeply segregated by race and ethnicity (e.g., Massey and Denton, Ch. 38; Wilson, Ch. 76)? Why are interracial marriages becoming more common (e.g., Farley, Ch. 65)?

Gender Inequality

Why has the gender gap in educational attainment disappeared altogether in some countries? Why has occupational sex segregation proven, by contrast, to be so durable (e.g., Reskin, Ch. 86; Grusky and Levanon, Ch. 87)? What accounts for the gender pay gap (e.g., Petersen and Morgan, Ch. 88; England, Ch. 89; Tam, Ch. 90)? Why has the historic decline in the pay gap stalled of late (e.g., Blau and Kahn, Ch. 91)?

Consequences of Inequality

Does our class position determine our politics, attitudes, and behaviors (Bourdieu, Ch. 93; Chan and Goldthorpe, Ch. 96; Lareau, Ch. 97; Hout and Moodie, Ch. 99)? Are such

effects of class weakening (e.g., Hout and Moodie, Ch. 99)?

Future of Inequality

Is a new “high-modern” form of stratification emerging (e.g., Giddens, Ch. 106; Beck and Lau, Ch. 108)? Will inequality continue to grow in many late-industrial countries (e.g., Firebaugh, Ch. 110)? Is the concept of social class still useful in describing contemporary forms of inequality (e.g., Pakulski and Waters, Ch. 107)? Are inequality regimes gradually shedding their distinctive features and converging towards some common regime (e.g., Esping-Andersen, Ch. 102)?

The foregoing questions all adopt a critical orientation to human stratification systems that is distinctively modern in its underpinnings. For the greater part of human history, the existing stratification order was regarded as an immutable feature of society, and the implicit objective of commentators was to explain or justify this order in terms of religious or quasi-religious doctrines (e.g., Bottomore 1965; Tawney 1931). It was only with the Enlightenment that a critical “rhetoric of equality” emerged in opposition to the civil and legal advantages of the aristocracy and other privileged status groupings. In its most radical form, this egalitarianism led to Marxist interpretations of human history, and it ultimately provided the intellectual underpinnings for socialist stratification systems. While much of stratification theory has been formulated in reaction and opposition to these early forms of Marxist scholarship,¹ the field nonetheless shares with Marxism the distinctively Enlightenment premise that individuals are to be equally valued and respected regardless of differences in their economic or social standing (see Meyer, Ch. 105; also Tawney 1931).

Basic Concepts

The eight lines of questioning described above are elaborated in the remainder of this

essay. Before turning to such elaboration, it will be helpful to introduce some of the basic concepts used to describe the amount and structure of inequality, the genesis and persistence of inequality, and the effects of inequality on social behavior of all kinds.

The key components of stratification systems are (1) the institutional processes that define certain types of goods as valuable and desirable, (2) the rules of allocation that distribute these goods across various jobs or occupations in the division of labor (e.g., doctor, farmer, “housewife”), and (3) the mobility mechanisms that link individuals to jobs and thereby generate unequal control over valued resources. It follows that inequality is produced by two types of matching processes: The social roles in society are first matched to “reward packages” of unequal value, and individual members of society are then allocated to the positions so defined and rewarded.² In all societies, there is a constant flux of occupational incumbents as newcomers enter the labor force and replace dying, retiring, or out-migrating workers, yet the positions themselves and the reward packages attached to them typically change only gradually. As Schumpeter (1953) puts it, the occupational structure can be seen as “a hotel . . . which is always occupied, but always by different persons” (p. 171).

What types of rewards are distributed via these two matching processes? It is increasingly fashionable to recognize that inequality is “multidimensional,” that income is accordingly only one of many important resources, and that income redistribution in and of itself would not eliminate inequality (e.g., Sen, Ch. 26). When a multidimensionalist approach is accordingly taken, one might usefully distinguish among the eight types of assets listed in Table 1, each understood as valuable in its own right rather than a mere investment item.³ The assets of Table 1 may of course serve investment as well as consumption functions. For example, most economists regard schooling as an invest-

ment that generates future streams of income, while some sociologists likewise regard social networks as forms of capital that can be parlayed into educational credentials, income, and other valued goods.⁴ While most of the assets listed in Table 1 are clearly convertible in this fashion, they are not necessarily regarded as investments by the individuals involved. In fact, many valuable assets can be secured at birth or through childhood socialization (e.g., the “good manners” of the aristocracy), and they are therefore acquired without the beneficiaries explicitly weighing the costs of acquisition against the benefits of future returns.⁵

The implicit claim underlying Table 1 is that the listed assets exhaust all possible consumption goods and, as such, constitute the raw materials of stratification systems. The stratification field has developed a vocabulary that describes how these raw materials are distributed across members of the population. The main parameters of interest pertain to (1) the overall amount of inequality, (2) the extent to which individuals are locked permanently into certain positions (i.e., rigidity), (3) the extent to which ascription (e.g., racial discrimination) is used for purposes of allocation, and (4) the degree to which the various dimensions of inequality cohere (or are “crystallized”). These four classes of parameters are reviewed below.

The overall *amount of inequality* in any given resource (e.g., income) may be defined as its dispersion or concentration among the individuals in the population. Although many scholars seek to characterize the overall level of societal inequality with a single parameter, such attempts will obviously be compromised insofar as some types of rewards are distributed more equally than others. This complexity clearly arises in the case of contemporary inequality regimes. The recent diffusion of “citizenship rights,” for example, implies that civil goods have come to be widely dispersed across all citizens, whereas economic and political goods continue to be

Table 1. Types of Assets and Examples of Advantaged and Disadvantaged Groups

| Assets | | Examples | |
|--------------------|---|---|---|
| <i>Asset group</i> | <i>Examples of types</i> | <i>Advantaged</i> | <i>Disadvantaged</i> |
| 1. Economic | Wealth Income Ownership | Billionaire Professional Capitalist | Bankrupt worker Laborer Worker (i.e., employed) |
| 2. Power | Political power Workplace authority Household authority | Prime minister Manager Head of household | Disenfranchised person Subordinate worker Child |
| 3. Cultural | Knowledge Digital culture “Good” manners | Intelligentsia Silicon Valley residents Aristocracy | Uneducated persons Residents of other places Commoners |
| 4. Social | Social clubs Workplace associations Informal networks | Country club member Union member Washington A-list | Nonmember Nonmember Social unknown |
| 5. Honorific | Occupational Religious Merit-based | Judge Saint Nobel Prize winner | Garbage collector Excommunicate Nonwinner |
| 6. Civil | Right to work Due process Franchise | Citizen Citizen Citizen | Illegal immigrant Suspected terrorist Felon |
| 7. Human | On-the-job training General schooling Vocational training | Experienced worker College graduate Law school graduate | Inexperienced worker High school dropout Unskilled worker |
| 8. Physical | Mortality Physical disease Mental health | Person with long life Healthy person Healthy person | A “premature” death Person with AIDS, asthma Depressed, alienated |

disproportionately controlled by a relatively small elite.

The *rigidity* of a stratification system is indexed by the continuity (over time) in the social standing of its members. The stratification system is highly rigid, for example, if the current wealth, power, or prestige of individuals can be accurately predicted on the basis of their prior statuses or those of their parents. It should again be emphasized that such rigidity (or “social closure”) will typically vary in degree across the different types of assets listed in Table 1.

The stratification system rests on *ascriptive processes* to the extent that conditions present at birth (e.g., parental wealth, sex, race) influence the subsequent social standing of individuals. If ascriptive processes of this sort are in operation, it is possible (but by no means guaranteed) that the underlying traits

themselves will become bases for group formation and collective action (e.g., race riots, feminist movements). In contemporary societies, ascription of all kinds is usually seen as undesirable or discriminatory, and much governmental policy (e.g., antidiscrimination legislation) is therefore directed toward fashioning a stratification system in which individuals acquire resources solely by virtue of their achievements.⁶

The degree of *crystallization* is indexed by the correlations among the various resources (e.g., income, wealth, education) that are socially valued. If these correlations are strong, then the same individuals (i.e., the “upper class”) will consistently appear at the top of all hierarchies, while other individuals (i.e., the “lower class”) will consistently appear at the bottom of the stratification system. By contrast, various types of status inconsisten-

cies (e.g., a poorly educated millionaire) will emerge in stratification systems with weakly correlated hierarchies, and it is correspondingly difficult in such systems to define a unitary set of classes that have predictive power with respect to all resources.

Measuring Inequality

Given the complexity of contemporary inequality, one might expect most scholars to adopt a multidimensional approach to characterizing it, an approach that might begin by specifying the multivariate “inequality space” of valued goods listed in Table 1. It nonetheless remains conventional among sociologists to describe stratification systems in terms of discrete classes or strata whose members are presumed to have similar levels or types of assets. In the most extreme versions of this approach, the resulting classes are assumed to be real entities that preexist the distribution of assets, and many scholars therefore refer to the “effects” of class location on the assets that their incumbents control. We present here a stylized history of such class models as well as competing approaches to characterizing the structure of inequality.

The claim that inequality takes on a “class form” is one of the few distinctively sociological contributions to inequality measurement and stands as the main alternative to approaches that either focus exclusively on income inequality or analyze the many dimensions of inequality independently and separately. The main advantage of class-based measurement, as argued by sociologists, is that conventional class categories (e.g., professional, manager, clerk, craft worker, laborer, farmer) are institutionalized within the labor market and are accordingly more than purely nominal or statistical constructions. The labor market, far from being a seamless and continuous distribution of incomes, is instead understood as a deeply lumpy entity, with such lumpiness mainly taking the form

of institutionalized groups (i.e., “classes”) that constitute prepackaged combinations of the valued goods listed in Table 1.

Within sociology, the implicit critique, then, of income-based approaches rests not so much on the argument that the income distribution is just one of many distributions of interest (i.e., multidimensionalism), but rather on the argument that measurement strategies based on the income distribution alone impose an excessively abstract, analytic, and statistical lens on a social world that has much institutionalized structure to it. This structure takes the tripartite form of a set of social classes that are privileged under capitalist labor markets (e.g., capitalists, professionals, managers), a set of social classes that are less privileged under advanced capitalism (e.g., routine nonmanuals, craft workers, operatives), and an “underclass” that stands largely outside the labor market and is accordingly deeply disadvantaged in market systems. The rise of class models should therefore be understood as a distinctively sociological reaction to the individualism of the income paradigm and other unidimensional approaches to measuring inequality.

The foregoing account, which is a largely consensual rendition of the rationale for social class measurement, nonetheless conceals much internal debate within the field on how best to identify and characterize the boundaries dividing the population into classes. We review below the three main phases in which these debates have played out.

Structuralist Phase (ca. 1945–1985)

The class models of the postwar period rested implicitly or explicitly on the assumption that classes are bundles of endowments (e.g., education levels), working conditions (e.g., amount of autonomy), and reward packages (e.g., income) that tend to cohere together (esp. Giddens, Ch. 14; Parkin, Ch. 15). The middle class of “craft workers,” for example, comprises individuals with moderate educational investments (i.e., secondary school

credentials), considerable occupation-specific investments in human capital (i.e., on-the-job training), average income coupled with substantial job security (at least until deindustrialization), middling social honor and prestige, quite limited authority and autonomy on the job, and comparatively good health outcomes (by virtue of union-sponsored health benefits and regulation of working conditions). By contrast, the underclass may be understood as comprising a rather different package of conditions, a package that combines minimal educational investments (i.e., secondary school dropouts), limited opportunities for on-the-job training, resulting in intermittent labor force participation and low income, virtually no opportunities for authority or autonomy on the job (during those brief bouts of employment), relatively poor health (by virtue of lifestyle choices and inadequate health care), and much social denigration and exclusion. The other classes appearing in conventional class schemes (e.g., professional, managerial, routine nonmanual) may likewise be understood as particular combinations of scores on the dimensions of Table 1.

For the purposes of illustration, consider a simplified case in which the multidimensional “inequality space” comprises only three individual-level variables (e.g., education, autonomy, income), thus allowing the class hypothesis to be readily graphed. Additionally, assume that the class structure can be represented by six classes (e.g., professional, managerial, sales and clerical, craft, laborer, farm), signified in Figure 1 by six different symbols (dark squares, light squares, dark circles, etc.). As shown in this figure, the two main claims underlying the class hypothesis are that (1) the structural conditions of interest tend to cluster together into characteristic packages, and (2) these packages of conditions correspond to occupational or employment groupings. For structuralists, the inequality space is presumed to have a relatively low dimensionality, indeed a di-

dimensionality neither more nor less than the number of postulated classes. The individuals falling within the classes constituting this scheme will accordingly have endowments, working conditions, and reward packages that are close to the averages prevailing for their classes. Moreover, even when individual scores deviate from class averages, the conventional class-analytic assumption (albeit wholly untested) is that the contextual effect of the class is dominant and overcomes any such individual-level deviations. For example, a full professor who lacks a Ph.D. is presumably just as marketable as a fully credentialed (but otherwise comparable) full professor, precisely because membership in the professorial class is a master status that tends to dominate all others.

The postwar period also was notable for a flourishing of gradational measurement approaches that again treated occupations as the fundamental units of analysis but assumed that such occupations may be ordered into a unidimensional socioeconomic or prestige scale (e.g., Treiman, Ch. 20; Goldthorpe and Hope, Ch. 21). In Figure 1, we assumed that the class structure cannot be understood in simple gradational terms, meaning that at least some classes were formed by combining high values on one dimension with low values on another. It is possible, however, that the structural conditions of interest tend to co-

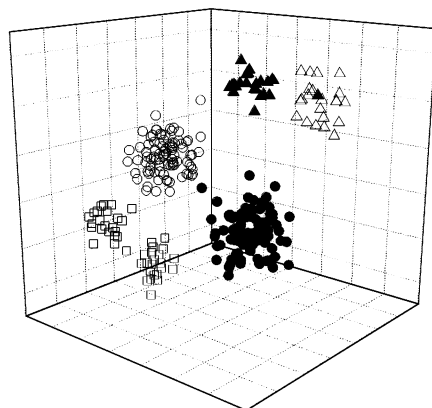


Figure 1. Class Regime

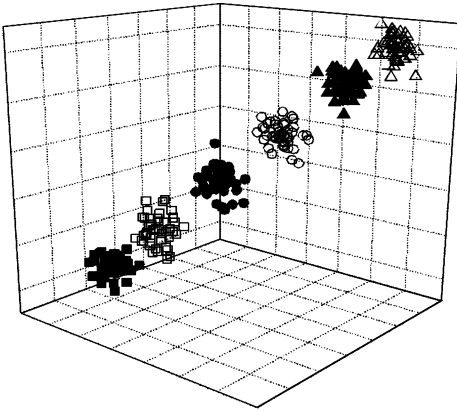


Figure 2. Gradational Regime

vary linearly, thus generating a class structure of the very simple type represented in Figure 2. In a regime of this sort, inequality becomes rather stark (i.e., crystallized), as privilege on one dimension implies very reliably privilege on another. There should accordingly be much interest in determining whether inequality indeed takes this form. Unfortunately, inequality scholars of the postwar period didn't typically test the crystallization assumption, but rather simply assumed that it held and proceeded to develop socioeconomic scales that treated education and income as the main dimensions of interest (and ranked occupations by averaging scores on these two dimensions).

It may be noted that many neo-Marxian scholars during this period also deviated from a strict multidimensional stance by nominating particular dimensions within Table 1 as being theoretically crucial and hence the appropriate basis upon which social classes might be defined. There are nearly as many claims of this sort as there are dimensions in Table 1. To be sure, Marx is most commonly criticized for placing "almost exclusive emphasis on economic factors as determinants of social class" (Lipset 1968, p. 300), but in fact much of what passed for stratification theorizing during this period amounted to reductionism of some kind, albeit often an expanded version of reduction-

ism in which two or three dimensions were nominated as especially crucial. When a reductionist position is adopted, the rationale for a class model is not typically that classes are coherent packages of conditions (as represented in Figure 1), but rather that the nominated dimension or dimensions are crucial in defining interests and will accordingly come to be the main sources of social action. The classic Marxian model, for example, has workers ultimately appreciating that their status as workers (i.e., nonowners) defines their interests.

Culturalist Phase (ca. 1985–1995)

In the mid-1980s, Bourdieu (Ch. 93) and other sociologists (esp. Wilson, Ch. 37) sought to develop a culturalist rationale for class models, a rationale that rested on the claim that classes are not merely constellations of structural conditions (e.g., working conditions, rewards) but are also socially closed groupings in which distinctive cultures emerge and come to influence attitudes, behaviors, or even preferences of class members. Throughout this period, many sociologists continued to work with more narrowly structuralist definitions of class (e.g., Wright, Ch. 11; Erikson and Goldthorpe, Ch. 47), but Bourdieu (Ch. 93) and Wilson (Ch. 37) were instrumental in pressing the claim that class-specific cultures are a defining feature of inequality systems (also, Lareau, Ch. 97).

The two main forms of closure that serve to generate class-specific cultures are residential segregation (e.g., urban ghettos) and workplace segregation (e.g., occupational associations). As Wilson notes, members of the underclass live in urban ghettos that are spatially isolated from mainstream culture, thus allowing a distinctively oppositional culture to emerge and reproduce itself. The effects of residential segregation operate, by contrast, in more attenuated form for other social classes; after all, residential communities map only imperfectly onto class categories

(i.e., the demise of the “company town”), and social interaction within contemporary residential communities is in any event quite superficial and cannot be counted upon to generate much in the way of meaningful culture. If distinctive cultures emerge outside the underclass, they do so principally through the tendency for members of the same occupation to interact disproportionately with one another in the workplace and in leisure activities. In accounting, for example, for the humanist, antimaterialist, and otherwise left-leaning culture and lifestyle of sociologists, class analysts would stress the forces of social closure within the workplace, especially the liberalizing effects of (1) lengthy professional training and socialization into the “sociological worldview,” and (2) subsequent interaction in the workplace with predominantly liberal colleagues.

When occupations or classes are allowed to have cultures in this fashion, one naturally wishes to better understand the content of those cultures and, in particular, the relationship between such content and the structural conditions (i.e., endowments, outcomes, institutional setting) that a class or occupational situation implies. At one extreme, class cultures may be understood as nothing more than “rules of thumb” that encode optimizing behavioral responses to prevailing institutional conditions, rules that allow class members to forgo optimizing calculations themselves and rely instead on cultural prescriptions that provide reliable and economical shortcuts to the right decision. For example, Breen and Goldthorpe (Ch. 64) argue that working-class culture is disparaging of educational investments not because of some maladaptive oppositional culture, but because such investments expose the working class (more so than other classes) to a real risk of downward mobility. In most cases, working-class children lack insurance in the form of substantial family income or wealth, meaning that they cannot easily recover from an educational investment gone awry (i.e.,

dropping out); and those who nonetheless undertake such an investment therefore face the real possibility of substantial downward mobility. The emergence, then, of a working-class culture that regards educational investments as frivolous may be understood as encoding that conclusion and thus allowing working-class children to undertake optimizing behaviors without explicitly engaging in decision tree calculations. The behaviors that a “rule of thumb” culture encourages are, then, deeply adaptive because they take into account the endowments and institutional realities that class situations encompass (also see Morgan, Ch. 55).

The foregoing example may be understood as one in which a class-specific culture instructs recipients about appropriate (i.e., optimizing) means for achieving ends that are widely pursued by *all* classes. Indeed, the prior “rule-of-thumb” account assumes that members of the working class share the conventional interest in maximizing labor market outcomes, with their class-specific culture merely instructing them about the approach that is best pursued in achieving that conventional objective. At the other extreme, one finds class-analytic formulations that represent class cultures as more overarching worldviews, ones that instruct not merely about the proper means to achieve ends but additionally about the proper valuation of the ends themselves. For example, some class cultures (e.g., aristocratic ones) place an especially high valuation on leisure, with market work disparaged as “common” or “polluting” (Veblen, Ch. 92). This orientation presumably translates into a high reservation wage within the aristocratic class. Similarly, “oppositional cultures” within the underclass may be understood as worldviews that place an especially high valuation on preserving respect and dignity for class members, with of course the further prescription that these ends are best achieved by (1) withdrawing from and opposing conventional mainstream pursuits, (2) representing conventional mo-

bility mechanisms (e.g., higher education) as tailor-made for the middle class and, by contrast, unworkable for the underclass, and (3) pursuing dignity and respect through other means, most notably total withdrawal from and disparagement of mainstream pursuits. This is a culture, then, that advocates that respect and dignity deserve an especially prominent place in the utility function and that further specifies how those ends might be achieved.

It should by now be clear that sociologists operating within the class-analytic tradition have adopted very strong assumptions about how inequality and poverty are structured. As noted, intrinsic to the class concept are such claims as (1) the space of outcomes and capabilities has a (low) dimensionality equaling the number of social classes, (2) the class locations of individuals become master statuses that dominate (or at least supplement) the effects of individual-level endowments, and (3) such class locations are socially closed and come to be associated with adaptive or maladaptive cultures. The prior claims have been unstated articles of faith among class analysts in particular and sociologists more generally. In this sense, class analysts have behaved rather like stereotypical economists, the latter frequently being parodied for their willingness to assume most anything provided that it leads to an elegant model.

Postmodernist Phase (ca. 1995–Present Day)

The third phase of conceptual work within sociology has been marked, however, by an increased willingness to challenge the assumptions underlying the class-analytic status quo. In recent years, such criticisms of the class-analytic enterprise have escalated, with many postmodernist scholars now feeling sufficiently emboldened to argue that the concept of class should be abandoned altogether. Although the postmodern literature is notoriously fragmented, the variant of postmodernism that is most relevant here pro-

ceeds from the assumption that the labor movement is rooted in the old and increasingly irrelevant conflicts of industrial capitalism, that political parties have abandoned class-based platforms in favor of those oriented toward values and lifestyles, and that class-based identities accordingly become ever weaker and more attenuated (see Hout and Moodie, Ch. 99, for a relevant discussion). The resulting “individualization of inequality” (e.g., Pakulski and Waters, Ch. 107; Beck and Lau, Ch. 108) implies that lifestyles and consumption practices are becoming decoupled from work identities as well as other status group memberships. The stratification system may be regarded, then, as a “status bizarre” (Pakulski and Waters, Ch. 107) in which identities are actively constructed as individuals select and are shaped by their multiple statuses.⁷

This hypothesis, which is represented in extreme form by Figure 3, has not yet been subjected to convincing empirical test and may well prove to be premature. Moreover, even if lifestyles and life chances are truly decoupling from economic class, this ought not be misunderstood as a more general decline in inequality per se. The brute facts of inequality will still be with us even if social classes of the conventional form are weakening. Most obviously, income inequality is clearly on the rise (e.g., Piketty and Saez, Ch.

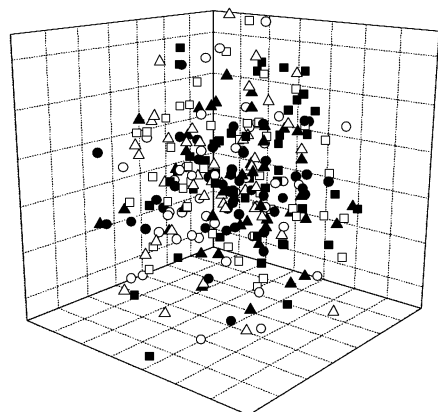


Figure 3. Disorganized Inequality

8; Morris and Western, Ch. 7), while other forms of inequality show no signs of withering away. The postmodernist hypothesis speaks, then, to the way in which inequality is organized, not to the overall amount of such inequality.

Analyzing Allocation

Although inequality scholars have long sought to understand how different “reward packages” are attached to different social positions, an equally important task within the field is that of understanding the rules by which individuals are allocated to the social positions so defined and rewarded. The language of stratification theory makes a sharp distinction between the distribution of social rewards (e.g., the income distribution) and the distribution of opportunities for securing these rewards. As sociologists have frequently noted, it is the latter distribution that governs popular judgments about the legitimacy of stratification: The typical American, for example, is quite willing to tolerate substantial inequalities in power, wealth, or prestige provided that the opportunities for securing these social goods are distributed equally. If the competition has been fairly run, we are quite willing to reward the winners and punish the losers.

The study of opportunities (or “capabilities”) is no less fashionable among economists. The main motivation, however, among economists for studying opportunities is not some intrinsic interest in mobility processes themselves, but rather a concern that standard outcome-based measures of inequality are tainted by the confounding effect of differential tastes. For example, an employee with a well-developed taste for leisure will presumably opt to work for relatively few hours, leading to low earnings but nonetheless optimal utility (by virtue of the high valuation placed on leisure). Whenever income inequality is generated through the operation of differential tastes, most econo-

mists would argue that it should be regarded as quite unproblematic, given that low-income workers are simply choosing, by virtue of their particular tastes, to trade off income for some other valuable good (e.g., leisure). This line of reasoning implies that inequality scholars should measure the distribution of opportunities that prevails before differential tastes can express themselves. The main task of an inequality scholar under this formulation is to determine whether “capabilities” (i.e., opportunities to secure rewards) are equally distributed, not whether rewards themselves, which reflect the operation of tastes, are equally distributed.

It follows that sociologists and economists have become quite interested, albeit for different reasons, in the study of opportunity and how it is unequally distributed. In most of the resulting research, the liberal ideal of an open and discrimination-free system is treated as an explicit benchmark, and the usual objective is to expose any inconsistencies between this ideal and the empirical distribution of life chances. This objective underlies analyses of the gross effects of class origins on class destinations (e.g., Featherman and Hauser, Ch. 46; Solon, Ch. 49; Conley, Ch. 63) as well as analyses of the net effects of gender, race, and class background after controlling education and related measures of achievement or merit (e.g., Blau and Duncan, Ch. 50; Sewell, Haller, and Portes, Ch. 53). We have presented a wide range of such analyses in this volume. Additionally, experimental approaches to measuring discrimination have recently become popular, most notably “audit studies” that proceed by (1) sending employers resumes that are identical save for the applicant’s gender, race, or class, and (2) then examining whether call back rates (for interviews) are nonetheless different across such groups. Although the available statistical and experimental studies all indicate that opportunities are far from equal (e.g., Bertrand and Mullainathan, Ch. 70; Pager, Ch. 72; also see Goldin and

Table 2. Parameters of Stratification for Eight Ideal-Typical Systems

| <i>Regime (1)</i> | <i>Principal Assets (2)</i> | <i>Major Strata or Classes (3)</i> | <i>Inequality (4)</i> | <i>Rigidity (5)</i> | <i>Crystallization (6)</i> | <i>Justifying Ideology (7)</i> |
|--|---|--|---------------------------|-------------------------|--------------------------------|--|
| <i>A. Hunting and gathering society</i> | | | | | | |
| 1. Tribalism | Human (hunting and magic skills) | Chiefs, shamans, and other tribe members | Low | Low | High | Meritocratic selection |
| <i>B. Horticultural and agrarian society</i> | | | | | | |
| 2. Asiatic mode | Political (i.e., incumbency of state office) | Office-holders and peasants | High | Medium | High | Tradition and religious doctrine |
| 3. Feudalism | Economic (land and labor power) | Nobility, clergy, and commoners | High | Medium-High | High | Tradition and Roman Catholic doctrine |
| 4. Slavery | Economic (human property) | Slave owners, slaves, “free men” | High | Medium-High | High | Doctrine of natural and social inferiority (of slaves) |
| 5. Caste society | Honorific and cultural (ethnic purity and “pure” life-styles) | Castes and subcastes | High | High | High | Tradition and Hindu religious doctrine |
| <i>C. Industrial society</i> | | | | | | |
| 6. Class system | Economic (means of production) | Capitalists and workers | Medium-High | Medium | High | Classical liberalism |
| 7. State socialism | Political (party and workplace authority) | Managers and managed | Low-Medium | Low-Medium | High | Marxism and Leninism |
| 8. Advanced industrial | Human (i.e., education, expertise) | Skill-based occupational groupings | Medium-High | Low-Medium | Medium | Classical liberalism |

Rouse, Ch. 80), there remains some debate about whether or to what extent such inequalities are declining or will continue to decline. We turn to such debates next.

A Compressed History of Inequality

The task of identifying the essential dynamics underlying social change has long been fundamental to sociology, but nowhere is this interest better developed or more fundamental than within the field of inequality analysis. The remainder of our essay is devoted, then, to laying out some of the most compelling narratives about how inequality has evolved. We begin this task by developing a quite conventional typology of inequality forms over the long span of human history and then turn to the question of how the narrative-building enterprise has changed in recent times (i.e., the last half-century).

Although stratification systems can of course be defined by a great many param-

eters, we proceed here by building a typology of inequality regimes in terms of the parameters of inequality that were introduced at the beginning of this essay (i.e., amount of inequality, rigidity, crystallization). For each of the eight regimes listed in Table 2, a particular class of assets (e.g., economic) has been assumed to be especially crucial in defining the structure of inequality, thus making it possible to specify the dominant social classes in terms of those assets. The rigidity of stratification systems can then be indexed by the amount of class persistence (see column 5), and crystallization can be indexed by the presumed correlation between class membership and each of the assets listed in Table 1 (see column 6).⁸ The final column in Table 2 rests on the further assumption that inequality regimes have (reasonably) coherent ideologies that legitimate the rules and criteria by which individuals are allocated to positions in the class structure (see column 7). In most cases, ideologies of this kind are

largely conservative in their effects, but they can sometimes serve as forces for change as well as stability. If, for example, the facts of labor market processes are inconsistent with the prevailing ideology (e.g., racial discrimination in advanced industrialism), then various sorts of ameliorative action might be anticipated (e.g., affirmative action programs).

The staple of modern classification efforts has been the tripartite distinction among class, caste, and estate, but there is also a long and illustrious tradition of Marxian typological work that introduces the additional categories of primitive communism, slave society, and socialism (e.g., Wright, Ch. 11). As shown in Table 2, these conventional approaches are largely complementary, and it is therefore possible to fashion a hybrid classification that incorporates most of the standard distinctions. This hybrid typology should be understood as a highly stylized and compressed history of inequality forms over the premodern, modern, and late modern periods.

The stratification forms represented in Table 2 should thus be taken as ideal types. In constructing these categories, our intention is not to make empirical claims about how existing systems operate in practice, but rather to capture and distill the accumulated wisdom about how these systems might operate in their purest form. These ideal-typical models can nonetheless assist us in understanding empirical systems. Indeed, insofar as societies evolve through the gradual overlaying of new stratification forms on older (and partly superseded) ones, it becomes possible to interpret contemporary systems as a complex mixture of several of the ideal types presented in Table 2.

The first panel in this table pertains to the tribal systems that dominated human society from the very beginning of human evolution until the Neolithic revolution of some 10,000 years ago. The characterizations of columns 2–7 necessarily conceal much variability in the structure of these systems: It is “merely in the night of our ignorance [that]

all alien shapes take on the same hue” (Anderson, 1974, p. 549). While such variable features of tribal societies are clearly of interest, for our purposes the important similarities are that (1) the total size of the distributable surplus was in all cases quite limited, and (2) this cap on the surplus placed corresponding limits on the overall level of economic inequality (but not necessarily on other forms of inequality). It is also noteworthy that customs such as gift exchange, food sharing, and the like were commonly practiced in tribal societies and had obvious redistributive effects. In fact, some observers (e.g., Marx [1939] 1971) treated these societies as examples of “primitive communism,” as the means of production (e.g., tools, land) were owned collectively and other types of property typically were distributed evenly among tribal members.

This is obviously not to suggest that a perfect equality prevailed. The more powerful medicine men (i.e., shamans) within tribal societies often secured a disproportionate share of resources, and the tribal chief could exert considerable influence on the political decisions of the day. In most cases, these residual forms of power and privilege were not directly inherited, nor were they typically allocated in accord with well-defined ascriptive traits (e.g., racial traits).⁹ It was only by demonstrating superior skills in hunting, magic, or leadership that tribal members could secure political office or acquire status and prestige. While meritocratic forms of allocation are often seen as prototypically modern, in fact they were present in incipient form at the very earliest stages of societal development.

With the emergence of agrarian forms of production, the economic surplus became large enough to support more complex systems of stratification. Among Marxist theorists, the “Asiatic mode” was often treated as an intermediate formation in the transition to advanced agrarian society (e.g., feudalism), and we have therefore led off our typol-

ogy with the Asiatic case (see Table 2, row B2).¹⁰ In doing so, we should emphasize that the explicit evolutionary theories of Godelier (1978) and others have not been well received, yet many scholars still take the fall-back position that Asiaticism is an important analytical category in the development of class society (e.g., Hobsbawm, 1965). The main features of this formation are (1) a large peasant class residing in agricultural villages that are “almost autarkic” (O’Leary 1989, p. 17), (2) the absence of strong legal institutions recognizing private property rights, (3) a state elite that extracts surplus agricultural production through rents or taxes and expends it on “defense, opulent living, and the construction of public works” (Shaw 1978, p. 127),¹¹ and (4) a constant flux in elite personnel due to “wars of dynastic succession and wars of conquest by nomadic warrior tribes” (O’Leary 1989, p. 18).

Beyond this skeletal outline, all else is open to dispute. There are long-standing debates, for example, about how widespread the Asiatic mode was and about the appropriateness of reducing all forms of Asian development to a “uniform residual category” (Anderson 1974, pp. 548–49). These issues are clearly worth pursuing, but for our purposes it suffices that the Asiatic mode provides a conventional example of how a “dictatorship of officialdom” can flourish in the absence of private property and a well-developed proprietary class (Gouldner 1980, pp. 327–28). Under this reading of Asiaticism, the parallel with modern socialism looms large (at least in some quarters), so much so that various scholars have suggested that Marx downplayed the Asian case for fear of exposing it as a “parable for socialism” (see Gouldner 1980, pp. 324–52).

Whereas the institution of private property was underdeveloped in the East, the ruling class under Western feudalism was, by contrast, very much a propertied one.¹² The distinctive feature of feudalism was that the nobility not only owned large estates or manors

but also held legal title to the labor power of its serfs (see Table 2, row B3).¹³ If a serf fled to the city, this was considered a form of theft: The serf was stealing that portion of his or her labor power owned by the lord (see Wright, Ch. 11). With this interpretation, the statuses of serf and slave differ only in degree, and slavery thereby constitutes a limiting case in which workers lose all control over their own labor power (see row B4). At the same time, it would obviously be a mistake to reify this distinction, given that the history of agrarian Europe reveals “almost infinite gradations of subordination” (Bloch 1961, p. 256) that confuse and blur the conventional dividing lines between slavery, serfdom, and freedom. The slavery of Roman society provides the best example of complete subordination, whereas some of the slaves of the early feudal period were bestowed with rights of real consequence (e.g., the right to sell surplus product), and some of the (nominally) free men were in fact obliged to provide rents or services to the manorial lord (Bloch 1961, pp. 255–74).¹⁴ The social classes that emerged under European agrarianism were thus structured in quite complicated ways. In all cases, we nonetheless find that property ownership was firmly established and that the life chances of individuals were defined, in large part, by their control over property in its differing forms. Unlike the ideal-typical Asiatic case, the nation-state was largely peripheral to the feudal stratification system, since the means of production (i.e., land, labor) were controlled by a proprietary class that emerged quite independently of the state.¹⁵

The historical record makes it clear that agrarian stratification systems were not always based on strictly hereditary forms of social closure (see Table 2, panel B, column 5). The case of European feudalism is especially instructive in this regard, since it suggests that stratification systems often become more rigid as the underlying institutional forms mature and take shape. Although it is

well known that the era of classical feudalism (i.e., post-twelfth century) was characterized by a “rigid stratification of social classes” (Bloch 1961, p. 325),¹⁶ there was greater permeability during the period prior to the institutionalization of the manorial system and the associated transformation of the nobility into a legal class. In this transitional period, access to the nobility was not yet legally restricted to the offspring of nobility, nor was marriage across classes or estates formally prohibited (see Bloch 1961, pp. 320–31). The case of ancient Greece provides a complementary example of a (relatively) open agrarian society. As Finley (1960) and others have noted, the condition of slavery was indeed heritable under Greek law, yet manumission was so common that the slave class had to be constantly replenished with new captives secured through war or piracy. The possibility of servitude was thus something that “no man, woman, or child, regardless of status or wealth, could be sure to escape” (Finley 1960, p. 161). At the same time, hereditary forms of closure were more fully developed in some slave systems, most notably the American one. As Sio (1965) notes, slavery in the antebellum South was “hereditary, endogamous, and permanent” (p. 303), with the annual manumission rate apparently as low as 0.04 percent by 1850 (see Patterson 1982, p. 273). The slave societies of Jamaica, South Africa, and rural Iraq were likewise based on largely permanent slave populations.

The most extreme examples of hereditary closure are of course found in caste societies (see Table 2, row B5). In some respects, American slavery might be seen as having “caste-like features” (see Berreman 1981), but Hindu India clearly provides the defining case of caste organization.¹⁷ The Indian caste system is based on (1) a hierarchy of status groupings (i.e., castes) that are ranked by ethnic purity, wealth, and access to goods or services, (2) a corresponding set of “closure

rules” that restrict all forms of intercaste marriage or mobility and thereby make caste membership both hereditary and permanent, (3) a high degree of physical and occupational segregation enforced by elaborate rules and rituals governing intercaste contact, and (4) a justifying ideology that induces the population to regard such extreme forms of inequality as legitimate and appropriate (Smaje 2000; Dumont 1970; Srinivas 1962). What makes this system so distinctive, then, is not merely its well-developed closure rules but also the fundamentally honorific (and noneconomic) character of the underlying social hierarchy. As indicated in Table 2, the castes of India are ranked on a continuum of ethnic and ritual purity, with the highest positions in the system reserved for castes that prohibit behaviors that are seen as dishonorable or polluting. Under some circumstances, castes that acquired political and economic power eventually advanced in the status hierarchy, yet they typically did so only after mimicking the behaviors and lifestyles of higher castes.

The defining feature of the industrial era (see Table 2, panel C) has been the emergence of egalitarian ideologies and the consequent delegitimation of the extreme forms of stratification found in caste, feudal, and slave systems. This can be seen, for example, in the European revolutions of the eighteenth and nineteenth centuries that pitted the egalitarian ideals of the Enlightenment against the privileges of rank and the political power of the nobility. In the end, these struggles eliminated the last residue of feudal privilege, but they also made new types of inequality and stratification possible. Under the class system that ultimately emerged (see row C6), the estates of the feudal era were replaced by purely economic groups (i.e., “classes”), and closure rules based on heredity were likewise supplanted by (formally) meritocratic processes. The resulting classes were neither legal entities nor

closed status groupings, and the associated class-based inequalities could therefore be represented and justified as the natural outcome of competition among individuals with differing abilities, motivation, or moral character (i.e., “classical liberalism”). As indicated in line C6 of Table 2, the class structure of early industrialism had a clear “economic base” (Kerbo 1991, p. 23), so much so that Marx (Ch. 9) defined classes in terms of their relationship to the means of economic production. The precise contours of the industrial class structure are nonetheless a matter of continuing debate (as discussed above). For example, a simple Marxian model focuses on the cleavage between capitalists and workers, while more elaborate Marxian and neo-Marxian models identify additional intervening or “contradictory” classes (e.g., Wright, Ch. 11), and yet other (non-Marxian) approaches represent the class structure as a continuous gradation of wealth or income.

Whatever the relative merits of these models might be, the ideology underlying the socialist revolutions of the nineteenth and twentieth centuries was of course explicitly Marxist. The intellectual heritage of these revolutions and their legitimating ideologies can again be traced to the Enlightenment, but the rhetoric of equality that emerged in this period was now directed against the economic power of the capitalist class rather than the status and honorific privileges of the nobility. The evidence from Eastern Europe and elsewhere suggests that these egalitarian ideals were only partially realized (e.g., Lenski, Ch. 3). In the immediate postrevolutionary period, factories and farms were indeed collectivized or socialized, and various fiscal and economic reforms were instituted for the express purpose of reducing income inequality and wage differentials among manual and nonmanual workers. Although these egalitarian policies were subsequently weakened through the reform efforts of

Stalin and others, inequality on the scale of prerevolutionary society was never reestablished among rank-and-file workers (cf. Lenski, Ch. 3). There nonetheless remained substantial inequalities in power and authority. Most notably, the socialization of productive forces did not have the intended effect of empowering workers, as the capitalist class was replaced by a “new class” of party officials and managers who continued to control the means of production and to allocate the resulting social surplus (see Eyal, Szélenyi, and Townsley, Chs. 34, 104). This class has been variously identified with intellectuals or intelligentsia, bureaucrats or managers, and party officials or appointees (Gouldner, Ch. 32). Regardless of the formulation adopted, the presumption is that the working class ultimately lost out in contemporary socialist revolutions, just as it did in the so-called bourgeois revolutions of the eighteenth and nineteenth centuries.

Whereas the means of production were socialized in the revolutions of Eastern Europe and the former Soviet Union, the capitalist class remained largely intact throughout the process of industrialization in the West. In the end, the propertied class may nonetheless be weakened by ongoing structural changes, with the most important of these being (1) the rise of a service economy and the growing power of the “service class” (Esping-Anderson, Ch. 102; Giddens, Ch. 106), (2) the increasing centrality of theoretical knowledge in the transition to a new “information age” (Bell, Ch. 101), and (3) the consequent emergence of technical expertise, educational degrees, and training certificates as “new forms of property” (Berg 1973, p. 183; Gouldner, Ch. 32; Brooks, Ch. 33). The foregoing developments all suggest that human and cultural capital are displacing economic capital as the principal stratifying forces in advanced industrial society (see Table 2, row C8). By this formulation, a new class of cultural elites may be emerging in the West, much as the transition to state

socialism (allegedly) generated a new class of intellectuals in the East.

The Role of Benign Narratives in Past Scholarship

The foregoing classification embodies a complicated amalgam of “benign narratives” about how inequality has developed within the industrial period, narratives that are benign in the sense that history is understood as operating in the main to reduce inequality, if only gradually and fitfully. Under such narratives, the problem of inequality becomes a tractable moral problem, an unfortunate side-circumstance of capitalism (and even socialism) that becomes yet more manageable with the transition into the increasingly affluent forms of advanced industrialism. This orientation to inequality is expressed in standard postwar narratives about three types of outcomes: (1) the distribution of income, power, and other valued resources; (2) the distribution of opportunities for securing income, power, and other valued resources; and (3) the formation of social classes and other institutionalized groups (e.g., racial groups, gender groups). We have hinted at these narratives in the preceding section and throughout this essay but turn in this section to laying them out more formally.

Trend in Inequality

The dominant inequality narrative of the postwar period featured the emergence of egalitarian ideologies and the consequent delegitimation of extreme forms of inequality (e.g., Kerr et al., Ch. 100; Bell, Ch. 101). The Enlightenment is understood in this context as unleashing one of the most profound revolutions in human history. The resulting decline in inequality can be seen, for example, in (1) the European revolutions of the eighteenth and nineteenth centuries against the privileges of rank (honorific equality); (2) the gradual elimination of inequalities in the right to vote, own

property, and speak and assemble (civil equality); (3) the abolition of slavery and the establishment of the radically egalitarian principle of self-ownership (equality of human assets); and (4) the equalization of economic assets via the rise of socialism, welfare capitalism, and their many institutions (economic equality).

As is well known, the latter commitment to equalizing economic assets was rather weaker than the commitment to other forms of equalization (at least in the capitalist context), with the result that economic inequalities remained extreme in all market economies. There was nonetheless a gradual decline in economic inequality throughout the postwar period in the United States and other industrial countries. According to the classic Kuznets curve (Kuznets 1955), the initial stages of capitalist development bring about an increase in income inequality as capital is increasingly concentrated among a small number of investors, whereas more advanced forms of capitalism entail a growth in the size of the middle class and a consequent reversal of the upward trend. The causal dynamics behind the resulting inverted-U pattern remain unclear (see Piketty and Saez, Ch. 8), but most sociologists attribute the late-industrial decline in inequality to the increasingly crucial role that the skilled working class played in production, the associated growth in working-class productivity, and the leverage that this growth in skills and productivity conferred on skilled workers.

It is worth inquiring as to the less proximate mechanisms by which egalitarianism of this sort may diffuse and take hold. The conventional view in this regard is that a series of crucial historical events after the Enlightenment (e.g., the defeat of Nazism, the civil rights movement) served to define equality as one of our core cultural commitments. Absent some revolutionary event that changes this cultural trajectory, the course of human history then becomes the “working out” of this commitment, a task that in-

volves shedding subsidiary values that sometimes come into conflict with our deeper commitment to egalitarianism. The core mechanism that drives this cultural diffusion may therefore be understood as the gradual reconciling of competing values to a new value, that of equality, that has been elevated by one or more historical events to a position of prominence.

This is obviously not to suggest that *all* postwar sociologists and social scientists emphasized forces making for a decline in inequality. However, even when a benign narrative was not explicitly adopted, there was usually some effort to engage with it and to explain or defend the decision not to take it up. This defensiveness was especially apparent in neo-Marxian analyses of the postwar era. Although such analyses were based on deeply pessimistic subnarratives about the trajectory of capitalism, these subnarratives were typically attached to larger and more benign narratives about the postcapitalist trajectory (e.g., Marx, Ch. 9; also see Wright, Ch. 11; Wallerstein, Ch. 12).

Trend in Inequality of Opportunity

The second benign narrative of interest rests on the conventional distinction between the distribution of social rewards (e.g., income) and the distribution of opportunities for securing these rewards. In liberal welfare regimes, extreme inequalities in rewards may be tolerated, but only insofar as opportunities for attaining these rewards are understood to be equally distributed. It is inequalities of opportunity that are regarded, then, as especially illegitimate in the context of liberal welfare regimes.

The dominant narratives of the postwar period have these inequalities of opportunity gradually weakening. The narratives of this period may be understood as benign because they describe the withering away of precisely those types of inequalities (i.e., inequalities of opportunity) that are regarded as problematic or illegitimate. The trademark of the

benign narrative is this simple correspondence between what we want and what we think will likely happen. We describe below four benign subnarratives that characterize some of the processes by which inequalities of opportunity may come to be weakened.

The most famous such subnarrative pertains to the discrimination-reducing effects of competitive market economies. In his original formulation of the “taste for discrimination” model, Becker (1957) argued that discrimination would be eroded by competitive market forces because it requires employers to pay a premium to hire members of the preferred class of labor, whether these be males, whites, or any other ascriptively defined classes. This taste is “discriminatory” because it rests on exogenous preferences for a certain category of labor that cannot be understood as arising from some larger concern for maximizing profitability or market share. When managers make hiring decisions in accord with such tastes, their firms will not be competitive with nondiscriminating firms because they must pay extra to secure labor from the preferred class (without any compensating increase in productivity). In standard renditions of this account, it is presumed that discriminating firms will gradually be selected out by the market, although it is also possible that some discriminating firms will change their hiring practices to remain competitive.

This economic subnarrative works in tandem with a second, “organizational” one that emphasizes the diffusion of modern personnel practices in the form of universalistic hiring practices (e.g., open hiring, credentialism) and bureaucratized pay scales and promotion procedures (Weber, Ch. 13; Grusky and Levanon, Ch. 87). The essence of such bureaucratic personnel practices is a formal commitment to universalism (i.e., treating all workers equally) and to meritocratic hiring and promotion (i.e., hiring and promoting on the basis of credentials). In its ideal-typical form, the spread of bureaucracy

becomes an organizational process with its own dynamic, a process of diffusion that rests not on actual efficiencies, as with the economic subnarrative, but simply on the *presumption* that bureaucratic practices are efficient and that “modern firms” must therefore adopt them. This subnarrative, like the economic one, implies that firms will gradually come to embrace organizational procedures that reduce inequalities of opportunity.

The third subnarrative of interest is the political one. Whereas the economic and organizational subnarratives treat change in inequality as an unintended by-product of macro-level forces (i.e., competition and bureaucratization), the political subnarrative is about instrumental action explicitly oriented towards effecting a decline in inequality. In theory, such political action could be oriented toward reducing either inequalities of opportunity or outcome, but historically a main emphasis within liberal welfare regimes has been legislation aimed at reducing inequality of opportunity (e.g., antidiscrimination legislation, early education programs, educational loans). The distinctive assumption of the political subnarrative is that straightforward “social engineering” is an important source of change and that the unintended or unanticipated consequences of such engineering are too often overemphasized.

The final subnarrative, a simple cultural one, rests on the argument that Western ideals of justice and equality continue to be endogenously worked out through a logic that diffuses independently of the economic efficiency of such ideals. The cultural subnarrative can be straightforwardly distinguished from the economic one because the growing “taste” for equality is presumed to be an exogenous shift rather than some accommodation to the economic cost of exercising discriminatory tastes. Likewise, the cultural subnarrative is distinct from the organizational subnarrative by virtue of focusing on the spread of tastes for equality and equality-

enhancing practices, not the spread of organizational forms (e.g., bureaucratization) that are deemed efficient, normatively desirable, or both. Similarly, the cultural and political subnarratives are closely related because political commitments to equal opportunity, antidiscrimination legislation, and school reform may be partly or even largely motivated by these newfound tastes for equal opportunity. At the same time, the cultural commitment to equal opportunity is not expressed exclusively in such political terms but is additionally expressed in the attitudes, behaviors, and personnel practices of employers. Most obviously, employers may gradually shed their preferences for certain categories of labor and instead develop positive tastes for equality in hiring, firing, and promotion, tastes that might at the limit be exercised in the labor market even with some loss in profits or efficiency.

The spread of such tastes for equal opportunity may again be viewed as part of our Enlightenment legacy, albeit a particular “liberal” variant of that legacy that emphasizes equalizing opportunities, not outcomes. This commitment is expressed not only at the individual level (e.g., changes in attitudes) but also at the collective level through various types of political reform (e.g., antidiscrimination legislation) as well as the diffusion of bureaucratic personnel policies (e.g., open hiring).

Trend in Class Formation

The final benign narrative of interest describes the gradual transition from “lumpy” class-based labor markets to more purely gradational ones (see Weeden et al., Ch. 27). Within this narrative, the early-industrial economy is represented as deeply balkanized into partly independent labor markets defined by detailed occupations (e.g., economist, carpenter), big social classes (e.g., manager, farmer), or yet more aggregated factors of production (e.g., worker, capitalist). For our purposes, what is principally of

interest is our collective fascination with arguments describing how these classes, however they may be defined, tend to gradually dissipate and leave us with gradational labor markets that increasingly approximate the seamless neoclassical ideal. The first step in this transition, as described most famously by Dahrendorf (Ch. 10), is the gradual “institutionalization” of class conflict, a regularization of labor-capital relations achieved through the establishment of unions, collective bargaining agreements, and other laws defining how labor and capital should negotiate. The second step in this transition involves the dismantling of unions and other institutionalized residues of classes as the liberal welfare ideals of “deregulation” and flexibility are increasingly pursued (Sørensen, Ch. 25). Although the mechanisms differ, there are of course analogous narratives on offer pertaining to the gradual decline of status groupings based on race, ethnicity, or gender (e.g., Wilson, Ch. 73; also see Petersen, Ch. 83).

This line of argumentation is additionally expressed in postmodernist narratives that assume that class identities, ideologies, and organization are attenuating and that “new theories, perhaps more cultural than structural, [are] in order” (Davis 1982, p. 585). The core claim here is that politics, lifestyles, and consumption practices are no longer class-determined and increasingly become a “function of individual taste, choice, and commitment” (Crook, Pakulski, and Waters 1992, p. 222; Pakulski and Waters, Ch. 107). In more ambitious variants of postmodernism, the focus shifts away from simply claiming that attitudes and practices are less class-determined, and the older class-analytic objective of understanding macro-level stratificational change is resuscitated. This ambition underlies, for example, all forms of postmodernism that seek to represent “new social movements” (e.g., environmentalism) as the vanguard force behind future stratificatory change. As

argued by Eyerman (1992) and others (e.g., Pakulski and Waters, Ch. 107), the labor movement can be seen as a fading enterprise rooted in the old conflicts of the workplace and industrial capitalism, whereas new social movements provide a more appealing call for collective action by virtue of their emphasis on issues of lifestyle, personal identity, and normative change. With this formulation, the proletariat is stripped of its privileged status as a universal class, and new social movements emerge as an alternative and far more benign force “shaping the future of modern societies” (Haferkamp and Smelser 1992, p. 17).

New Approaches to Studying Inequality

The foregoing narratives, all of which were fixtures of the postwar intellectual landscape, describe the emergence of a world in which inequalities are less profound, opportunities are more equally distributed, and class conflicts and interclass differences become attenuated. These narratives are benign in the sense that they push us toward equilibria that most commentators, even neo-Marxian ones, might well regard as appealing. The benign narrative is accordingly built on the happy correspondence between what should be and what will be.

If there is any theme to contemporary analyses of inequality, it is that the benign narrative has fallen largely out of fashion. We have nonetheless laid out these standard narratives in some detail because they provide an important backdrop to current theorizing and are often used as foils by contemporary scholars seeking to motivate their own analyses. The benign narrative is in this sense lurking in the background of contemporary discussions of inequality. We turn now to a closer discussion of how contemporary analyses of inequality have developed partly in reaction to the benign narratives of the postwar period.

Multidimensionalism and New Inequalities

As a natural starting point for this discussion, we note that contemporary inequality scholarship is increasingly concerned with new forms of inequality, forms that were either ignored in the past or have been spawned by new technologies or institutions. This growing emphasis on new inequalities is consistent with the now fashionable view that inequality is multidimensional and that conventional studies of economic, socioeconomic, or cultural inequality hardly exhaust its many forms.

This approach may be understood as simply reemphasizing the importance of studying all eight forms of inequality listed in Table 1. If the resulting multidimensional space is taken seriously, we can then ask how individuals are distributed among the less conventionally studied dimensions of the inequality space. Are “new” assets as unequally distributed as old ones? Is inequality becoming more of an “all or nothing” affair in which upper-class workers are advantaged on all dimensions of interest and lower-class workers are disadvantaged on all dimensions of interest? Are new assets sometimes distributed in ways that compensate for shortfalls in older ones? In the present volume, multidimensionalist questions of this kind are posed for such “new” outcomes as health (Mullahy, Robert, and Wolfe, Ch. 95), computer literacy (Hargittai, Ch. 98), imprisonment or capital punishment (Western, Ch. 41), and networks and social capital (Granovetter, Ch. 59; Lin, Ch. 60; Burt, Ch. 61). These new types of inequality may be understood in some cases as truly new divides generated by new technologies (e.g., the digital divide) or new social institutions (e.g., modern mass prisons). More typically, the “new” outcomes are just increasingly popular topics of study among academics, not truly new forms (e.g., health inequalities).

The Intransigence of Poverty and Inequality

The foregoing line of research typically takes the form of an exposé of the extent to which seemingly basic human entitlements, such as living outside of prison, freely participating in digital culture, or living a long and healthy life, are unequally distributed in ways that sometimes amplify well-known differentials of income or education. The continuing attraction of such exposés (at least among academics) may be attributed to our collective discomfort with an economic system that generates rather more inequality than is palatable under contemporary cultural standards. Although the equalizing reforms of social democracy have historically been a main solution to this tension, the declining legitimacy of such reform (especially in Europe and the United States) leaves the tension an increasingly unresolved one.

Whereas the old narratives focused, then, on the forces making for decline in inequality, a more pessimistic assessment of the trajectory of late industrialism has now taken hold, and much scholarship accordingly focuses on documenting that inequality has persisted at higher levels than had been anticipated. This sensibility underlies, for example, contemporary research showing that residential segregation in the United States is so extreme as to constitute a modern form of “apartheid” (Massey and Denton, Ch. 38), that racial discrimination in labor markets likewise remains extreme (Bertrand and Mullainathan, Ch. 70; Pager, Ch. 72), that the occupational structure is “hypersegregated” by gender (Reskin, Ch. 86), that income inequality has increased markedly in many countries over the last thirty-five years (Piketty and Saez, Ch. 8), that poverty rates in the United States remain strikingly high (Smeeding, Ch. 36), that African Americans are routinely harassed, slighted, and insulted in public places (Feagin, Ch. 74), that working-class and middle-class children tend to

be raised in profoundly different ways (Lareau, Ch. 97), that political behavior continues to be strongly shaped by class-based politics (Hout and Moodie, Ch. 99), and that massive class disparities in access to health services persist (Mullahy, Robert, and Wolfe, Ch. 95).

The cynic might ask whether this new muckraking tradition is really all that necessary. Is there truly a large public that doesn't already appreciate the persistence of many of these inequalities? We live, after all, in a market society in which virtually everything is commodified, meaning that almost all goods and services (e.g., health care, housing) are allocated on the basis of our ability to pay for them. Because we are so deeply and (seemingly) irrevocably marketized, the real intellectual challenge would be to find a good, service, or outcome that is somehow untouched by class, one that is perfectly and equally distributed to all. The role of class is in this sense so obvious as to preclude any need for academic muckraking.

We suspect that our contributors would react to such (hypothetical) criticism by emphasizing that only some inequalities may be understood as the inevitable outcome of our collective decision to allocate resources on the basis of a market. It is at least possible to imagine markets that do not entail racial discrimination, that do not entail gender discrimination, and that encompass institutions that have a substantial inequality-moderating effect. It is likewise possible to imagine that class differences in politics, culture, and child rearing practices would have by now abated. The benign narratives of the postwar period in fact laid out precisely such imaginings. If it is now clear that these imagined futures have not been realized, surely we need to document that conclusion with all the rigor that can possibly be mustered. To be sure, many of us well know that the world is a massively unequal one, but even so the force of the known can

be readily lost when we live with profound inequality on a day-to-day basis. This commitment to remind us of what comes to be taken for granted is the cornerstone, we suspect, of the renewed interest in the inequality-documenting function.

The Rise of Less Benign Narratives

The rise of this muckraking exposé of inequality has been coupled, moreover, with increasing interest in developing narratives that explain why inequality has persisted or grown more extreme. These narratives are typically less grand than the quite encompassing narratives of the postwar period; that is, rather specialized narratives have recently developed around many of the various unit trends of interest (e.g., the expansion of income inequality, the stalling decline in the gender wage gap), and rather little attention has been paid to developing some grand meta-narrative that links these specialized narratives together. The signature, then, of the contemporary narrative is this highly delimited focus, a commitment to developing a rigorously empirical foundation, and a special interest in identifying those more insidious social forces that undermine the benign narratives of the past.

By way of example, consider the historic rise in income inequality, a development that has spawned one of the most sustained efforts at narrative-building of our time (see Morris and Western, Ch. 7). As noted above, the classic Kuznets curve aligns nicely with the facts of inequality up to the early 1970s, but then a dramatic, unprecedented upswing in inequality in the post-1970 period made it clear that inequality history had not ended. We have since witnessed one of the most massive research efforts in the history of social science as scholars sought to identify the “smoking gun” that accounted for this dramatic increase in inequality.

Initially, the dominant hypothesis was that deindustrialization (i.e., the relocation of

manufacturing jobs to offshore labor markets) brought about a decline in demand for less-educated manufacturing workers, a decline that generated increases in inequality by hollowing out the middle class and sending manufacturing workers into unemployment or into the ranks of poorly paid service work. Although this line of argumentation still has its advocates, it cannot easily be reconciled with evidence suggesting that the computerization of the workplace and related technological change has been at least one force behind a heightened demand for highly educated workers. Because of this result (and other supporting evidence), the deindustrialization story has now been largely supplanted by the converse hypothesis that “skill-biased technological change” has increased the demand for high-skill workers beyond the increase in supply, thus inducing a short-term disequilibrium and a correspondingly increased payoff for high-skill labor. At the same time, most scholars acknowledge that this story is at best an incomplete one and that other accounts, especially more narrowly political ones, must additionally be entertained (e.g., Card and DiNardo 2002). Most notably, some of the rise in income inequality in the United States was clearly attributable to the declining minimum wage (in real dollars), a decline that in turn has to be understood as the outcome of political processes that increasingly favor pro-inequality forces (Morris and Western, Ch. 7). The same conclusion applies yet more obviously to the recent round of tax cuts in the United States.

The future of income inequality depends on which of these underlying mechanisms is principally at work. The silver lining of the deindustrialization story is that within-country increases in inequality should be offset by between-country declines (as poor countries profit from new manufacturing jobs), whereas the silver lining of skill-biased technological change is that the heightened demand for high-skill workers is presumably a

one-time, short-term disequilibrium that will, by virtue of the higher payoff to high-skill jobs, trigger a compensating growth in the supply of high-skill workers. There is, unfortunately, no shortage of competing stories that imply more disturbing futures, even futures consistent with a classical Marxian account in which low-skill workers are emiserated within some countries by virtue of a globalization-induced “race to the bottom.” Indeed, accounts that focus on the political sources of rising inequality often take on this more disturbing character, given that social democratic ideologies have fallen largely out of fashion and no longer provide capitalists with a viable high road of “enlightened self-interest” (e.g., support for labor unions, redistribution). As social democratic agendas come to be viewed with suspicion, political support for the minimum wage and other inequality-reducing institutions may increasingly falter, and market-generated inequality may no longer be much restrained by pre-market or after-market interventions.

We have focused on the rise of income inequality and the various narratives it has generated only because this literature is especially well known and central to the field. The larger point that we seek to make is that, no matter the subfield, there appears to be much interest in developing narratives that explain why long-standing declines in inequality have slowed down, stalled altogether, or even reversed themselves. We are referring, for example, to (1) narratives of “globalization” that describe how the liberalization of financial and capital markets has harmed poor countries (Stiglitz, Ch. 109); (2) narratives of “deindustrialization” that describe the loss of inner-city jobs and the associated rise of an urban underclass (Ehrenreich, Ch. 35; Wilson, Ch. 37); (3) narratives of “deunionization” that describe the loss of middle-class unionized jobs and the emergence of nonstandard forms of employment (e.g., Piore, Ch. 56; Kalleberg, Ch. 58); (4) narratives of “segmented assimilation” that

describe the relatively bleak prospects for at least some new immigrant groups (Portes and Zhou, Ch. 68; Waters, Ch. 69); (5) narratives of “opting out” that have highly trained women eschewing stressful careers in favor of recommitting to their children, spouses, and domestic responsibilities (Belkin, Ch. 78; cf. Boushey, Ch. 79); (6) narratives of “essentialist segregation” that describe how sex-typed occupational ghettos continue to be built around presumed differences in male and female aptitudes (Grusky and Levanon, Ch. 87), (7) narratives of “nonconscious discrimination” that refer to subtle internalized prejudices that are especially difficult to extirpate (e.g., Correll, Benard, and Paik, Ch. 81; Reskin, Ch. 82); and (8) narratives of “spatial segregation” that emphasize the deeply institutionalized forces generating racial enclaves (Massey and Denton, Ch. 38) and poverty-stricken neighborhoods (Pebly and Sastry, Ch. 39; Sampson and Morenoff, Ch. 40). The key question of our time, and one which remains largely unanswered by the evidence of the last half-century, is whether the forces for equality featured in the “liberal theory” are strong enough to overcome the above countervailing processes. Although counternarratives of the more optimistic sort are also being developed (e.g., Firebaugh, Ch. 110), these seem not to be as frequently generated or as readily embraced, and the proponents of such narratives tend now to find themselves beleaguered, outnumbered, and on the defensive.

Conclusions

We have fashioned the latter part of our review around the subtle changes over the last half-century in the types of narratives that sociologists and other social scientists have applied to make sense of trends in inequality. The narratives of the postwar period took on a strikingly benign form in which the dominant logics of history were understood as operating in the main to reduce inequality. These

benign narratives, which now mainly seem naive and quaint, have been supplanted by a host of new narratives that give far greater weight to the forces making for inequality.

Has the pendulum swung too far? It is of course child’s play to posit any number of nonempirical sources of our fascination, some might say obsession, with the pessimistic narrative. The following are perhaps the most obvious of such accounts:

The Newsworthiness Account

If the postwar era of the benign narrative was indeed one of across-the-board declines in inequality, then a special incentive presumably emerged to ferret out results that were inconsistent with the prevailing wisdom and hence would be regarded as newsworthy. It is surely difficult to market analyses and forge careers on the basis of business-as-usual evidence. We might ask, for example, whether the “digital divide” emerged as a newsworthy topic because of early evidence of a substantial divide. If instead access to computing was found to be equal, would we have quickly discarded the topic and set off to find some other more unequal outcome? This type of selection on the dependent variable (i.e., inequality) will create a research literature that exaggerates how unequal the world is.

The “Moral Credentials” Account

It is well to bear in mind that the contemporary academic, far from taking a vow of poverty, is now firmly ensconced in the middle class, often the upper middle class. When relatively privileged scholars study poverty and disadvantage, they often feel a special obligation to demonstrate a strong commitment to amelioration and to display political sentiments that are liberal, progressive, or even radical. In some cases, this pressure may motivate them to downplay any evidence of decline in inequality or disadvantage, presumably out of concern that undue emphasis on the progress achieved so far will make it

appear that the remaining disadvantage is taken as acceptable or unproblematic.

The “Obsession with Small Differences” Account

The continuing diffusion of egalitarian values renders any departures from equality, no matter how small, as problematic and newsworthy. By this logic, even increasingly small intergroup differences will attract much attention, especially because ever more powerful models and statistical methods now make it possible to tease them out.

If these nonempirical interpretations shouldn't be dismissed out of hand, nor should they be taken too seriously. It is relevant in this regard that many of the pessimistic narratives involve indisputably dramatic changes in outcomes (e.g., income) that were frequently studied well before any reversal in the trendline was detected. Although some of the pessimistic narratives featured in this volume are still in incipient form and are not yet well researched, it will of course be difficult to continue to maintain these narratives should strong disconfirming evidence turn up. The implication, then, is that our exaggerated taste for pessimism might conceivably lead us to cycle through a great many pessimistic stories (and fail to develop enough benign ones), but at least the usual rules of evidence will have us excise egregiously flawed narratives, no matter how benign or pessimistic they may be.

NOTES

1. In fact, the term *stratification* has itself been seen as anti-Marxist by some commentators (e.g., Duncan 1968), since it places emphasis on the vertical ranking of classes rather than the exploitative relations between them. The geological metaphor implied by this term does indeed call attention to issues of hierarchy. Nonetheless, whenever it is used in the present essay, the intention is to refer generically to inequality of all forms.

2. In some stratification systems, the distribution of rewards can be described with a single matching algorithm, since individuals receive rewards directly

rather than by virtue of the social positions that they occupy. The limiting case here would be the tribal economies of Melanesia in which “Big Men” (Oliver 1955) secured prestige and power through personal influence rather than through incumbency of any well-defined roles (also see Granovetter 1981, pp. 12–14).

3. It goes without saying that the assets listed in Table 1 are institutionalized in quite diverse ways. For example, some assets are legally recognized by the state or by professional associations (e.g., civil rights, property ownership, educational credentials), others are reserved for incumbents of specified work roles (e.g., workplace authority), and yet others have no formal legal or institutional standing and are revealed probabilistically through patterns of behavior and action (e.g., high-status consumption practices, deference, derogation).

4. It is sometimes claimed that educational credentials are *entirely* investment goods and should therefore be excluded from any listing of the primitive dimensions underlying stratification systems (e.g., Runciman 1968, p. 33). In evaluating this claim, it is worth noting that an investment rhetoric for schooling became fashionable only quite recently (e.g., Becker 1975), whereas intellectuals and humanists have long viewed education as a simple consumption good.

5. This is not to gainsay the equally important point that parents often encourage their children to acquire such goods because of their putative benefits.

6. Although “native ability” is by definition established at birth, it is often seen as a legitimate basis for allocating rewards (because it is presumed to be relevant to judgments of merit).

7. Although Pakulski and Waters (Ch. 107) use the label *postmodern* in their analyses, other scholars have invented such alternative terms as *late modernity*, *high modernity*, or *reflexive modernization* (see Giddens, Ch. 106; Beck and Lau, Ch. 108), and yet others continue to use *modernity* on the argument that the changes at issue are mere extensions of those long underway. We frequently use the term *postmodern* in this essay without intending to disadvantage the analyses of those who prefer other labels.

8. The assumptions embedded in columns 4–6 of Table 2 are clearly far-reaching. Unless a stratification system is perfectly crystallized, its parameters for inequality and rigidity cannot be represented as scalar quantities, nor can the intercorrelations between the multiple stratification dimensions be easily summarized in a single parameter. Moreover, even in stratification systems that are perfectly crystallized, there is no reason to believe that persistence

over the life course (i.e., intragenerational persistence) will always vary in tandem with persistence between generations (i.e., intergenerational inheritance). We have nonetheless assumed that each of our ideal-typical stratification systems can be characterized in terms of a single “rigidity parameter” (see column 5).

9. This claim does not hold with respect to gender. That is, women and men were typically assigned to different roles, which led to consequent differences in the distribution of rewards.

10. It should again be stressed that our typology by no means exhausts the variability of agrarian stratification forms (see Kerbo 2000 for an extended review).

11. The state elite was charged with constructing and maintaining the massive irrigation systems that made agriculture possible in regions such as China, India, and the Middle East (cf. Anderson 1974, pp. 490–92).

12. This is not to suggest that feudalism can only be found in the West or that the so-called Asiatic mode is limited to the East. Indeed, the social structure of Japan was essentially feudalistic until the mid-nineteenth century (with the rise of the Meiji State), and the Asiatic mode has been discovered in areas as diverse as Africa, pre-Columbian America, and even Mediterranean Europe (see Godelier 1978). The latter “discoveries” were of course predicated on a broad and ahistorical definition of the underlying ideal type. As always, there is a tension between scholars who seek to construct ideal types that are closely tied to historical social systems and those who seek to construct ones that are broader and more encompassing in their coverage.

13. This economic interpretation of feudalism is clearly not favored by all scholars. For example, Bloch (1961, pp. 288–89) argues that the defining feature of feudalism is the monopolization of *authority* by a small group of nobles, with the economic concomitants of this authority (e.g., land ownership) thus being reduced to a position of secondary importance. The “authority classes” that emerge under his specification might be seen as feudal analogues to the social classes that Dahrendorf (Ch. 10) posits for the capitalist case.

14. In the so-called secondary stage of feudalism (Bloch 1961), the obligations of serfs and free men became somewhat more formalized and standardized, yet regional variations of various sorts still persisted.

15. It was not until the early fourteenth century that states of the modern sort appeared in Europe.

16. In describing this period of classical feudalism, Bloch (1961) noted that “access to the circle of knights . . . was not absolutely closed, [yet] the door was nevertheless only very slightly ajar” (p. 325).

17. The Indian caste system flourished during the agrarian period, yet it persists in attenuated form within modern industrialized India.

REFERENCES

- Anderson, Perry. 1974. *Lineages of the Absolutist State*. London: Verso.
- Becker, Gary S. 1957. *The Economics of Discrimination*. Chicago: University of Chicago Press.
- _____. 1975. *Human Capital*. Chicago: University of Chicago Press.
- Berg, Ivar. 1973. *Education and Jobs: The Great Training Robbery*. Harmondsworth, UK: Penguin Books.
- Berremán, Gerald. 1981. *Caste and Other Inequities*. Delhi: Manohar.
- Bloch, Marc. 1961. *Feudal Society*. London: Routledge and Kegan Paul.
- Bottomore, Thomas B. 1965. *Classes in Modern Society*. London: Allen and Unwin.
- Card, D., and J. E. DiNardo. 2002. “Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles.” *Journal of Labor Economics* 20: 733–83.
- Crook, Stephen, Jan Pakulski, and Malcolm Waters. 1992. *Postmodernization*. London: Sage Publications.
- Davis, James. 1982. “Achievement Variables and Class Cultures: Family, Schooling, Job, and Forty-Nine Dependent Variables in the Cumulative GSS.” *American Sociological Review* 47: 569–86.
- Dumont, Louis. 1970. *Homo Hierarchicus*. Chicago: University of Chicago Press.
- Duncan, Otis Dudley. 1968. “Social Stratification and Mobility: Problems in the Measurement of Trend.” Pp. 675–719 in *Indicators of Social Change*, edited by Eleanor B. Sheldon and Wilbert E. Moore. New York: Russell Sage Foundation.
- Eyerman, Ron. 1992. “Modernity and Social Movements.” Pp. 37–54 in *Social Change and Modernity*, edited by Hans Haferkamp and Neil J. Smelser. Berkeley: University of California Press.
- Finley, Moses J. 1960. *Slavery in Classical Antiquity*. Cambridge, UK: W. Heffer and Sons.
- Godelier, Maurice. 1978. “The Concept of the ‘Asiatic Mode of Production’ and Marxist Models of Social Evolution.” Pp. 209–57 in *Relations of*

- Production, Marxist Approaches to Economic Anthropology*, edited by D. Seddon. London: Cass.
- Gouldner, Alvin. 1980. *The Two Marxisms: Contradictions and Anomalies in the Development of Theory*. New York: Seabury Press.
- Granovetter, Mark. 1981. "Toward a Sociological Theory of Income Differences." Pp. 11–47 in *Sociological Perspectives on Labor Markets*, edited by Ivar Berg. New York: Academic Press.
- Haferkamp, Hans, and Neil J. Smelser. 1992. "Introduction." Pp. 1–33 in *Social Change and Modernity*, edited by Hans Haferkamp and Neil J. Smelser. Berkeley: University of California Press.
- Hobsbawm, Eric J. 1965. "Introduction." Pp. 9–65 in *Pre-Capitalist Economic Formations*, translated by Jack Cohen and edited by Eric J. Hobsbawm. New York: International Publishers.
- Kerbo, Harold R. 1991. *Social Stratification and Inequality: Class Conflict in Historical and Comparative Perspective*. New York: McGraw-Hill.
- _____. 2000. *Social Stratification and Inequality: Class Conflict in Historical, Comparative, and Global Perspective*. New York: McGraw-Hill.
- Kuznets, Simon. 1955. "Economic Growth and Income Inequality (Presidential Address)." *American Economic Review* 45: 1–28.
- Lipset, Seymour M. 1968. "Social Class." Pp. 296–316 in *International Encyclopedia of the Social Sciences*, edited by D. L. Sills. New York: Macmillan.
- Marx, Karl. [1939] 1971. *The Grundrisse*. New York: Harper and Row.
- O'Leary, Brendan. 1989. *The Asiatic Mode of Production*. Oxford: Basil Blackwell.
- Oliver, Douglas L. 1955. *A Solomon Island Society*. Cambridge: Harvard University Press.
- Patterson, Orlando. 1982. *Slavery and Social Death*. Cambridge: Harvard University Press.
- Runciman, Walter G. 1968. "Class, Status, and Power?" Pp. 25–61 in *Social Stratification*, edited by J. A. Jackson. New York: Cambridge University Press.
- Schumpeter, Joseph. 1953. *Aufsätze zur Soziologie*. Tübingen: Mohr/Siebeck.
- Shaw, William H. 1978. *Marx's Theory of History*. Stanford, CA: Stanford University Press.
- Sio, Arnold A. 1965. "Interpretations of Slavery: The Slave Status in the Americas." *Comparative Studies in Society and History* 7: 289–308.
- Smaje, Chris. 2000. *Natural Hierarchies: The Historical Sociology of Race and Caste*. Oxford: Basil Blackwell.
- Srinivas, Mysore N. 1962. *Caste in Modern India*. London: Asia Publishing House.
- Tawney, R. H. 1931. *Equality*. London: George Allen and Unwin.

PART II

Forms and Sources of Inequality

The Functions of Inequality

2 ■ Kingsley Davis and Wilbert E. Moore

Some Principles of Stratification

In a previous paper some concepts for handling the phenomena of social inequality were presented.¹ In the present paper a further step in stratification theory is undertaken—an attempt to show the relationship between stratification and the rest of the social order.² Starting from the proposition that no society is “classless,” or unstratified, an effort is made to explain, in functional terms, the universal necessity which calls forth stratification in any social system.

Throughout, it will be necessary to keep in mind one thing—namely, that the discussion relates to the system of positions, not to the individuals occupying those positions. It is one thing to ask why different positions carry different degrees of prestige, and quite another to ask how certain individuals get into those positions. Although, as the argument will try to show, both questions are related, it is essential to keep them separate in our thinking.

Most of the literature on stratification has tried to answer the second question (particularly with regard to the ease or difficulty of mobility between strata) without tackling the first. The first question, however, is logically prior and, in the case of any particular individual or group, factually prior.

The Functional Necessity of Stratification

Curiously the main functional necessity explaining the universal presence of stratification is precisely the requirement faced by any society of placing and motivating individuals in the social structure. As a functioning mechanism a society must somehow distribute its members in social positions and induce them to perform the duties of these positions. It must thus concern itself with motivation at two different levels: to instill in the proper individuals the desire to fill certain positions, and, once in these positions, the desire to perform the duties attached to them. Even though the social order may be relatively static in form, there is a continuous process of metabolism as new individuals are born into it, shift with age, and die off. Their absorption into the positional system must somehow be arranged and motivated. This is true whether the system is competitive or non-competitive. A competitive system gives greater importance to the motivation to achieve positions, whereas a non-competitive system gives perhaps greater importance to the motivation to perform the duties of the positions; but in any system both types of motivation are required.

Kingsley Davis and Wilbert E. Moore. “Some Principles of Stratification,” *American Sociological Review* 10 (April 1945), pp. 242–249.

If the duties associated with the various positions were all equally pleasant to the human organism, all equally important to societal survival, and all equally in need of the same ability or talent, it would make no difference who got into which positions, and the problem of social placement would be greatly reduced. But actually it does make a great deal of difference who gets into which positions, not only because some positions are inherently more agreeable than others, but also because some require special talents or training and some are functionally more important than others. Also, it is essential that the duties of the positions be performed with the diligence that their importance requires. Inevitably, then, a society must have, first, some kind of rewards that it can use as inducements, and, second, some way of distributing these rewards differentially according to positions. The rewards and their distribution become a part of the social order, and thus give rise to stratification.

One may ask what kind of rewards a society has at its disposal in distributing its personnel and securing essential services. It has, first of all, the things that contribute to sustenance and comfort. It has, second, the things that contribute to humor and diversion. And it has, finally, the things that contribute to self-respect and ego expansion. The last, because of the peculiarly social character of the self, is largely a function of the opinion of others, but it nonetheless ranks in importance with the first two. In any social system all three kinds of rewards must be dispensed differentially according to positions.

In a sense the rewards are “built into” the position. They consist in the “rights” associated with the position, plus what may be called its accompaniments or perquisites. Often the rights, and sometimes the accompaniments, are functionally related to the duties of the position. (Rights as viewed by the incumbent are usually duties

as viewed by other members of the community.) However, there may be a host of subsidiary rights and perquisites that are not essential to the function of the position and have only an indirect and symbolic connection with its duties, but which still may be of considerable importance in inducing people to seek the positions and fulfill the essential duties.

If the rights and perquisites of different positions in a society must be unequal, then the society must be stratified, because that is precisely what stratification means. Social inequality is thus an unconsciously evolved device by which societies insure that the most important positions are conscientiously filled by the most qualified persons. Hence every society, no matter how simple or complex, must differentiate persons in terms of both prestige and esteem, and must therefore possess a certain amount of institutionalized inequality.

It does not follow that the amount or type of inequality need be the same in all societies. This is largely a function of factors that will be discussed presently.

The Two Determinants of Positional Rank

Granting the general function that inequality subserves, one can specify the two factors that determine the relative rank of different positions. In general those positions convey the best reward, and hence have the highest rank, which (a) have the greatest importance for the society and (b) require the greatest training or talent. The first factor concerns function and is a matter of relative significance; the second concerns means and is a matter of scarcity.

Differential Functional Importance. Actually a society does not need to reward positions in proportion to their functional importance. It merely needs to give sufficient reward to them to insure that they will be

filled competently. In other words, it must see that less essential positions do not compete successfully with more essential ones. If a position is easily filled, it need not be heavily rewarded, even though important. On the other hand, if it is important but hard to fill, the reward must be high enough to get it filled anyway. Functional importance is therefore a necessary but not a sufficient cause of high rank being assigned to a position.³

Differential Scarcity of Personnel. Practically all positions, no matter how acquired, require some form of skill or capacity for performance. This is implicit in the very notion of position, which implies that the incumbent must, by virtue of his incumbency, accomplish certain things.

There are, ultimately, only two ways in which a person's qualifications come about: through inherent capacity or through training. Obviously, in concrete activities both are always necessary, but from a practical standpoint the scarcity may lie primarily in one or the other, as well as in both. Some positions require innate talents of such high degree that the persons who fill them are bound to be rare. In many cases, however, talent is fairly abundant in the population but the training process is so long, costly, and elaborate that relatively few can qualify. Modern medicine, for example, is within the mental capacity of most individuals, but a medical education is so burdensome and expensive that virtually none would undertake it if the position of the M.D. did not carry a reward commensurate with the sacrifice.

If the talents required for a position are abundant and the training easy, the method of acquiring the position may have little to do with its duties. There may be, in fact, a virtually accidental relationship. But if the skills required are scarce by reason of the rarity of talent or the costliness of training, the

position, if functionally important, must have an attractive power that will draw the necessary skills in competition with other positions. This means, in effect, that the position must be high in the social scale—must command great prestige, high salary, ample leisure, and the like.

How Variations Are to Be Understood. In so far as there is a difference between one system of stratification and another, it is attributable to whatever factors affect the two determinants of differential reward—namely, functional importance and scarcity of personnel. Positions important in one society may not be important in another, because the conditions faced by the societies, or their degree of internal development, may be different. The same conditions, in turn, may affect the question of scarcity; for in some societies the stage of development, or the external situation, may wholly obviate the necessity of certain kinds of skill or talent. Any particular system of stratification, then, can be understood as a product of the special conditions affecting the two aforementioned grounds of differential reward.

NOTES

1. Kingsley Davis, "A Conceptual Analysis of Stratification," *American Sociological Review*. 7:309–321, June, 1942.

2. The writers regret (and beg indulgence) that the present essay, a condensation of a longer study, covers so much in such short space that adequate evidence and qualification cannot be given and that as a result what is actually very tentative is presented in an unfortunately dogmatic manner.

3. Unfortunately, functional importance is difficult to establish. To use the position's prestige to establish it, as is often unconsciously done, constitutes circular reasoning from our point of view. There are, however, two independent clues: (a) the degree to which a position is functionally unique, there being no other positions that can perform the same function satisfactorily; (b) the degree to which other positions are dependent on the one in

question. Both clues are best exemplified in organized systems of positions built around one major function. Thus, in most complex societies the religious, political, economic, and educational functions are handled by distinct structures not easily interchangeable. In addition, each structure possesses many different positions, some clearly dependent on, if not subordinate to, others. In sum, when an institutional nucleus becomes differenti-

ated around one main function, and at the same time organizes a large portion of the population into its relationships, the *key* positions in it are of the highest functional importance. The absence of such specialization does not prove functional unimportance, for the whole society may be relatively unspecialized; but it is safe to assume that the more important functions receive the first and clearest structural differentiation.

3 ■ Gerhard Lenski

New Light on Old Issues

The Relevance of “Really Existing Socialist Societies” for Stratification Theory

Scholars have long debated the causes, consequences, and legitimacy of systems of social inequality, with some defending them as natural, inevitable, or even divinely ordained, and others challenging them as unnatural, unnecessary, and immoral (Lenski 1966, ch. 1). In the twentieth century, the most important challenges have come from groups and individuals inspired, directly or indirectly, by the work of Marx and his followers.

One does not need to look far in sociology to see the impact of Marx’s vision and the controversies it has created. As many have observed, the long-running debate between functionalists and their critics is, in many ways, a debate over the merits of Marxism: Functionalists maintain that economic inequality is both necessary for societies and beneficial for the vast majority of their members, whereas their critics argue that it is neither.

Unfortunately, from the standpoint of our understanding of the causes and conse-

quences of systems of stratification and the merits of Marx’s ideas, the debate among sociologists has focused almost entirely on the experience of Western “capitalist” societies.¹ Surprisingly little attention has been devoted to the experience of the former Soviet republics, Poland, East Germany before unification, the once-united Czechoslovakia, Hungary, the former Yugoslavia, Romania, Bulgaria, Albania, China, Cuba, North Korea, Vietnam, and other societies that were or have been governed for extended periods by dedicated Marxists. Yet, as East European sociologists have often pointed out in recent years, these societies have provided a unique set of laboratories for observing the effects of “really existing socialism.”² They allow us to observe socialist societies functioning in the real world under real-life conditions. In these societies, we can see what actually happens when private ownership is abolished and the emphasis in a society’s

Gerhard Lenski. “New Light on Old Issues: The Relevance of ‘Really Existing Socialist Societies’ for Stratification Theory,” in *Social Stratification*, 2nd Edition, edited by David B. Grusky, pp. 77–83. Copyright © 2001 by Westview Press, Inc. Reprinted by permission of Westview Press, Inc., a member of the Perseus Books Group.

system of rewards is shifted from material incentives to moral incentives. Imperfect though these tests have been, they shed valuable new light on the causes and consequences of inequalities in power and privilege.³ The results have been much too consistent to be ignored or written off as simply a matter of chance, and the consistency is especially impressive when one considers the great cultural diversity of the societies involved.

For many years, Western sociologists could justify their inattention to “really existing socialist societies” because of the difficulties of obtaining reliable data. By the early 1970s, however, a sufficient body of evidence had accumulated, and political conditions in a number of Marxist societies had improved to the point that one could, with some confidence, begin to form a fairly accurate view of a number of important aspects of the new Marxist systems of stratification. On the basis of materials available at the time, I concluded in an earlier article (Lenski 1978) that these “experiments in destratification” had enjoyed their greatest successes in reducing *economic* inequality: Differentials in wealth and income appeared to be substantially less in societies governed by Marxist elites than in other societies. These successes were offset, however, by two major failures: (1) *Political* inequalities in these societies were enormous, far greater than in any of the Western industrial democracies, and (2) none of these societies had achieved anything remotely resembling the critical transformation in human nature that Marx had predicted would follow the abolition of private property and would lay the foundation for the subsequent evolution of societies from socialism to communism. These failures, I concluded, were due in large measure to a critical flaw in Marxian theory—its unrealistic assumptions about human nature.

Looking back, I believe these conclusions have stood the test of time fairly well. Of course, information that has since emerged

and the wisdom of hindsight would lead me to modify and extend them. For example, recent revelations following the overthrow of the Marxist regimes in Eastern Europe indicate that the level of economic inequality in those societies was greater than I was then aware. To cite but three examples: (1) After the overthrow of Todor Zhivkov, the Bulgarian public and the rest of the world learned that during his years in power he had acquired no fewer than thirty separate homes for his personal use and that he and other top Communist Party leaders had accumulated millions of dollars in secret foreign bank accounts (Laber 1990); (2) the longtime Communist leader of Romania, Nicolae Ceaușescu, amassed forty villas and twenty palaces for himself and his family and accumulated millions in Swiss bank accounts at a time when the bulk of the population was often living without heat or light (*Washington Post* 1990); and (3) in East Germany, Erich Honecker accumulated millions of dollars in Swiss bank accounts by skimming profits from arms sales to Third World nations, while sharing with other top Communist Party leaders exclusive private hunting preserves and other luxuries that were denied to, and hidden from, the rest of the population. Although it has long been clear that Communist Party elites enjoyed many privileges that were denied to others (Matthews 1978), the extent of these privileges has proved to be much greater than most had supposed. That these were not merely aberrations of East European Marxism is indicated by non-European examples: In Nicaragua, the villas and much of the other property once owned by Anastasio Somoza and his associates became the personal property of top Sandinista leaders and their families, while in China and Vietnam, Communist Party elites continue to live in closed compounds (similar to those in the former East Germany) where living conditions are carefully hidden from public scrutiny (Salisbury 1992).

At the other extreme, poverty in these societies was more widespread and more serious than Western observers generally realized.

Reports by Soviet authorities in the late 1980s indicated that at least 20 percent of the population was living at or below the official poverty level (Fein 1989). Homelessness was also reported to be a problem in Moscow and other Soviet cities, while studies in Hungary at the end of the Communist era found that a quarter of the population was living in poverty (Kamm 1989).

Despite these revelations, it still appears that the level of economic inequality in Marxist societies never equaled the level found in Japan and most of the Western democracies. Wealthy and privileged though the Zhivkovs, Ceaușescus, and Honeckers were by comparison with their fellow citizens, the magnitude of their wealth never compared with the great fortunes amassed by leading Western and Japanese businessmen and by oil-rich Middle Eastern leaders. Furthermore, passing wealth on to the next generation has always been much more difficult in Marxist societies than elsewhere, as the unhappy experiences of the Leonid Brezhnev family and others indicate.⁴

A more serious flaw in my earlier assessment was its failure to anticipate the speed and magnitude of the changes that lay ahead. Although I anticipated that the gradual process of political liberalization that began after Stalin's death would continue, and that other changes would occur in response to problems encountered and to the changing needs and growing demands of a better educated population, I cannot pretend to have foreseen the sudden collapse of Communist Party hegemony, the rapid emergence of multiparty systems, or the radical economic changes that have occurred in most of Eastern Europe.

The benefit of hindsight makes clear that the internal, systemic problems of the command economies and one-party polities of Marxist societies were far more serious than most Western observers suspected. In fact, it now appears that the greatest success of Marxist regimes was their ability to dissimu-

late—a success that was too often achieved because of the readiness of large numbers of Western journalists, scholars, and others to accept glowing reports of socialist successes uncritically (Hollander 1981; Fang 1990). With the revelations that have followed in the wake of the democratic revolutions in Eastern Europe, we now know that the economies of these societies had been stagnating for years and that much of the population had become disaffected and hostile. Worse yet, Marxism and Marxist elites had lost whatever legitimacy they once enjoyed in the minds of many people, especially intellectuals and other opinion leaders and even Party members. (Ironically, this was at a time when Marxism was becoming increasingly fashionable among Western intellectuals.)

These developments have great relevance for our understanding of the causes and consequences of inequality, since it seems that many of the internal, systemic problems of Marxist societies were the result of inadequate motivational arrangements of the sort debated by stratification theorists such as Davis and Moore (1945), Davis (1953), and Tumin (1953). These problems were of two basic types: (1) undermotivation of ordinary workers and (2) misdirected motivation of managers, bureaucrats, and other decision-makers.

The first of these problems was summarized succinctly years ago by East European workers themselves who said, “They pretend to pay us, and we pretend to work” (Dobbs 1981). The rewards for most kinds of work simply did not justify anything more than minimal, perfunctory effort (Shlapentokh 1989, ch. 2). Shoddy workmanship, sullen workers, absenteeism, corruption, and bureaucratic pathologies of various kinds came to typify worker performance in Marxist societies (*The Economist* 1988). These problems are present in every society to some degree, but they became far more prevalent and far more serious in the socialist economies of Marxist societies than in most others. They

became so serious, in fact, that they had demoralizing consequences for the vast majority of citizens: endless hours spent in lines queuing for merchandise that was either of poor quality or in short supply, frequent confrontations with surly state employees, unsatisfactory housing, an inadequate health-care system, and more. To add insult to injury, most citizens became aware that a small minority of their fellows was exempted from most of these problems: For them, there were well-stocked stores with better quality merchandise in ample supply and more responsive employees, better housing, better health-care facilities, better schools for their children, second homes, and countless other perks. Worse, this elite preached socialism and the need for sacrifice while enjoying all these special privileges.

To describe the conditions that developed in these societies is to raise the question of why the system failed so badly. What went wrong, and why was the promise of freedom and affluence for the masses never achieved?

For many years, Marxist elites in Eastern Europe and their Western sympathizers explained away these problems on the grounds of *external factors*: the historic backwardness of Eastern Europe, the damage to the Soviet economy caused by the civil war that followed the 1917 revolution, and the hostility of the Western democracies. Although there was much truth to these claims, it has become increasingly clear that *internal, systemic factors* were also a major source of problems for many years. By the late 1980s, this had become obvious even to the leaders of these societies, with many of them becoming advocates of change, and some abandoning Marxism altogether.

Over the years, Marxist societies experimented with a variety of incentive systems, but the egalitarian nature of Marxist ideology always led to substantial limitations on wage differentials for the masses of workers.⁵ Over time, however, the severity of these limita-

tions varied as Party elites attempted either to improve the economic performance of their societies or, alternatively, to conform more closely to socialist principles. In a few instances, in an excess of socialist zeal, wage differentials were virtually eliminated: In Czechoslovakia in the early 1960s, for example, wage differences were reduced to the point that engineers and highly skilled workers earned only 5 percent more than unskilled workers. Because of this, large numbers of talented young people dropped out of school, feeling that it was not worth the effort required and the income that would be sacrificed to continue their education. Morale problems also developed among skilled workers, engineers, and other professionals. Within several years, problems had become so acute that authorities were forced to reverse themselves and increase rewards for better educated and more highly skilled workers. A similar crisis developed in the Soviet Union in the early 1930s, forcing Stalin to increase material incentives and wage differentials substantially (Inkeles 1950), and there is growing evidence that the economic crisis in the Soviet Union of the 1980s developed initially in response to a process of wage leveling begun under Brezhnev.

The chief reason for these problems appears to be a basic flaw in Marxist theory. Writing in the nineteenth century, Marx was heir to the eighteenth-century Enlightenment view of human nature—an optimistic view that saw the unattractive aspects of human life as products of corrupting social institutions that could be eliminated by rational social engineering. Whereas the French philosophes blamed the defects in human nature on the influence of church and state, Marx saw private property as the ultimate source of society's ills: If it were abolished, human nature would be transformed. Once socialism was established and the means of production were owned by all, moral incentives could replace material incentives and workers would find work intrin-

sically rewarding (see also Tumin [1953] on this point). They would work for the sheer joy of working and for the satisfaction of contributing to society's needs, not simply to earn a livelihood.

Unfortunately, the abolition of private property failed to produce the happy transformation in human nature that Marx anticipated. On the contrary, freed from the fear of unemployment and lacking adequate material incentives, worker performance deteriorated and production stagnated or declined in Marxist societies everywhere (Shlapentokh 1989; *The Economist* 1988; Silk 1990; Kamm 1989; Jones 1981; Scammel 1990; Huberman and Sweezy 1967; Zeitlin 1970). The most compelling evidence of this has come from the two Germans, which shared a common cultural heritage that involved a long tradition of worker pride. Yet by the closing days of the German Democratic Republic, reports of slack work patterns were widespread, and many East German workers were quoted as expressing concern that they would be unable to adapt to the more demanding standards of West German industry. In 1990, at the twenty-eighth Communist Party congress in the Soviet Union, President Mikhail Gorbachev's close associate, Aleksandr Yakovlev, asserted that labor productivity in capitalist South Korea was substantially greater than in socialist North Korea (New York Times News Service 1990). Tatiana Zaslavskaja, a leading Soviet sociologist, found that as many as a third of Soviet workers hated work and were unresponsive to incentives of any kind (Shlapentokh 1987).

But the motivational problems of Marxist societies stemmed from more than faulty assumptions about human nature. They were also due to defective organizational arrangements spawned by the command economies of those societies. Lacking the system of automatic controls inherent in a market economy, economic planners were forced to devise elaborate plans and assign production

quotas for the managers of every enterprise. To ensure fulfillment of these quotas, managers were awarded bonuses for meeting or exceeding them and were penalized severely for any shortfall. One unanticipated consequence of this seemingly rational procedure was that managers acquired a strong incentive to stockpile essential resources of every kind—including labor (Kostakov 1989; Smith 1976; Greenhouse 1989). Thus, labor resources in these societies came to be used very inefficiently; the result was that workers became cynical about the value of what they were called on to do.

Managers also developed a variety of other unfortunate adaptations to central planning. They learned, for example, that quantity, not quality, was what their bosses, the central planners, cared about (Parkhomovsky 1982).⁶ They also learned that production figures could be inflated without much risk because their bosses were also rewarded for good statistics and no one had any interest in seeing if actual performance matched reported performance (G. Medvedev 1989; Z. Medvedev 1990).

Finally, managers learned that there were only minimal rewards for reinvestment and for technological innovation. Lacking pressures from direct economic competition, Party leaders and planners failed to appreciate the importance of continuous modernization of their industrial plant. According to one account, Soviet managers received bonuses of 33 percent for fulfilling production quotas but only 8 percent for fulfilling the plan for new technology (*The Economist* 1988, 11). Thus, because capital investment and technological advance were badly neglected, the command economies of Marxist societies became less and less competitive in world markets.⁷

All of this evidence seems to confirm Davis's (1953) assertion that successful incentive systems involve (1) motivating the best qualified people to seek the most important positions and (2) motivating them to perform

to the best of their ability once they are in them. Marxist societies seem to have failed on both counts, using political criteria primarily both to allocate positions and to reward incumbents (Voslensky 1984; Kennedy and Bialecki 1989; Voinovich 1989).

The many malfunctions in the command economies of Marxist societies raise the question of whether they were more or less inevitable consequences of the system itself. This is a question of considerable importance, since command economies are not confined to Marxist societies. The public sector in every society functions as a command economy, and the public sector has been expanding in most societies in recent decades.

Although it is not possible to explore this question in depth here, several observations are in order. First, a substantial majority of the citizens in most of the once socialist societies of Eastern Europe rejected the system when given the chance. Even many Party leaders came to have little faith in central planning and the command economy. As one member of the Soviet Congress of People's Deputies said on the floor of that body, his nation taught the world a valuable lesson by testing, at great cost to itself, what proved to be "an impossible system of economic development" (Zakharov 1990).

Second, there have been remarkable similarities in the performance of command economies in otherwise widely divergent Marxist societies. Most of the pathologies found in Eastern Europe—absenteeism, poor work discipline, low levels of productivity, failure to reinvest in plants and to encourage innovation—have also been reported in China, Cuba, and elsewhere.

Finally, many of these same problems are also evident in the public sector of non-Marxist societies. Government workers and workers in state-owned Western enterprises are widely perceived as less diligent, innovative, enterprising, and responsive than work-

ers in private industry: Negative associations with the term "bureaucrat" are almost as strong in non-Marxist societies as in Marxist ones. In addition, government agencies in these societies are often noted for their inefficient use of human and other resources. Managers in these bureaucracies often find that they are more likely to maximize their own rewards by expanding the size of the work force and other resources under their supervision (regardless of need) than by using these resources efficiently.

Some observers have argued that the massive failures of the socialist economies of Marxist societies in Eastern Europe and elsewhere demonstrate the obvious superiority of capitalism and indicate that the future lies with capitalism. That conclusion, however, seems unwarranted. As noted earlier, even those societies that are usually referred to as "capitalist" have, in reality, very *mixed economies*. To paraphrase Marx, they are societies in which rewards are allocated partly on the basis of *need*, partly on the basis of *work*, and partly on the basis of *property*. In short, they combine elements of communism, socialism, and capitalism and are the product of trial-and-error experimentation guided, in large measure, by a spirit of pragmatism. Mixed economies are systems that recognize the need for material incentives and acknowledge the benefits of economic inequality. But they are also systems that recognize the necessity of allocating a part of the economic product on the basis of need and most on the basis of work.⁸ In short, the old view of societies as being either capitalist or socialist seems increasingly irrelevant.

Over time, an ever-increasing number of societies and their leaders have accomplished what scholarly theorists have so notably failed to achieve: They have created a workable synthesis out of seemingly contradictory principles of allocation. One of the urgent tasks for students of inequality in the years ahead will be to catch up with this new social

reality and create the kind of theoretical synthesis that does justice to the economic synthesis that has been created in most Western democracies in recent decades. Too much of stratification theory still resembles the work of the proverbial blind men struggling to describe an elephant.

No real synthesis is likely to emerge, however, so long as students of stratification ignore the crucial body of evidence that has accumulated concerning the effects on motivation and productivity of the massive experiments in destratification conducted in the twentieth century by Marxist elites. In effect, these experiments have provided us with far better evidence than any we have had before of the limits of what is possible in terms of the reduction of differentials in wealth and income. And although these tests cannot be considered definitive, neither can they be written off and ignored as most analysts have done so far.

NOTES

I wish to thank Peter Bearman, David Grusky, Michael Kennedy, and Anthony Oberschall for valuable suggestions concerning a prior draft of this paper. They are, of course, in no way responsible for flaws and errors in this final version.

1. I have qualified the label *capitalist* because all Western industrial societies now have mixed economies with substantial state controls over and limitations on the rights of ownership.

2. The terms *really existing socialism* and *really existing socialist societies* were coined by East European sociologists. Although the Marxist-Leninist societies, to which the terms have been applied, represent but one version of socialism, they are especially important for stratification theory because the former leaders of these societies were able to implement the basic socialist principle of abolishing private property far more successfully than socialists in Western Europe ever were.

3. Unfortunately, imperfect tests are a fact of life in the social sciences. If the tests of Marxist theory that are possible in Marxist societies fall short of the scientific ideal, the same is true of almost every test in the social sciences. To deny the relevance of evidence from imperfect tests would be to deny the relevance of most of what has been learned over the years in the social sciences.

4. Shortly after Brezhnev's death, his son-in-law was arrested and sentenced to prison on charges of corruption.

5. The salaries of Party leaders were also kept quite low, but they were compensated generously in a variety of other ways.

6. Quality controls are far more likely when consumers can choose among competing products. When people must use their own money to purchase goods and services, they are not nearly so willing to accept inferior products as when they are using public funds.

7. For example, only 23 percent of Soviet inventions were put to use within two years of their date of patenting, compared to 66 percent of American inventions and 64 percent of West German (*The Economist* 1988).

8. Internal Revenue Service data indicate that approximately 10 percent of U.S. GNP is allocated on the basis of need (public health, welfare, and education expenditures), 70 percent on work (wages and salaries), and 20 percent on property (interest, rents, dividends, capital gains) (Lenski 1984, 202).

REFERENCES

- Davis, Kingsley. 1953. "Reply [to Tumin, 1945]." *American Sociological Review* 18:394–397.
- Davis, Kingsley, and Wilbert Moore. 1945. "Some Principles of Stratification." *American Sociological Review* 10:242–249.
- Dobbs, Michael. 1981. "They Pretend to Pay Us, We Pretend to Work,' East Europeans Say." *Washington Post*, April 22.
- The Economist*. 1988. "The Soviet Economy." April 9, pp. 3–18.
- Fang, Lizhi. 1990. "The Chinese Amnesia." *New York Review of Books*, September 27, pp. 30–31.
- Fein, Esther. 1989. "Glasnost Is Opening the Door on Poverty." *New York Times*, January 29.
- Greenhouse, Steven. 1989. "Can Poland's Dinosaur Evolve?" *New York Times*, November 27.
- Hollander, Paul. 1981. *Political Pilgrims: Travels of Western Intellectuals to the Soviet Union, China, and Cuba, 1928–1978*. New York: Oxford University Press.
- Huberman, Leo, and Paul Sweezy. 1967. *Socialism in Cuba*. New York: Modern Reader Paperbacks.
- Inkeles, Alex. 1950. "Social Stratification and Mobility in the Soviet Union: 1940–1950." *American Sociological Review* 15:465–479.
- Jones, T. Anthony. 1981. "Work, Workers, and Modernization in the USSR." *Sociology of Work* 1:249–283.

- Kamm, Henry. 1989. "Hungarians Shocked by News of Vast Poverty in Their Midst." *New York Times*, February 6.
- Kennedy, Michael, and Ireneusz Bialecki. 1989. "Power and the Logic of Distribution in Poland." *Eastern European Politics and Societies* 3:300–328.
- Kostakov, Vladimir. 1989. "Employment: Scarcity or Surplus?" Pp. 159–175 in Anthony Jones and William Moskoff, eds., *Perestroika and the Economy*. Armonk, N.Y.: Sharpe.
- Laber, Jeri. 1990. "The Bulgarian Difference." *New York Review of Books*, May 17, pp. 34–36.
- Lenski, Gerhard. 1966. *Power and Privilege: A Theory of Social Stratification*. New York: McGraw-Hill.
- . 1978. "Marxist Experiments in Destratification: An Appraisal." *Social Forces* 57: 364–383.
- . 1984. "Income Stratification in the United States: Toward a Revised Model of the System." *Research in Social Stratification and Mobility* 3:173–205.
- Matthews, Mervyn. 1978. *Privilege in the Soviet Union*. London: Allen & Unwin.
- Medvedev, Grigorii. 1989. *Chernobyl'skaia Khronika*. Moscow: Sovremennik. Cited by David Holloway.
1990. "The Catastrophe and After." *New York Review of Books*, July 19, p. 5.
- Medvedev, Zhores. 1990. *The Legacy of Chernobyl*. New York: Norton.
- New York Times News Service. 1990. "Gorbachev Checks Headstrong Congress." July 8.
- Parkhovmosky, Elrad. 1982. "Can't Anybody Here Make Shoes?" *Izvestia*. Reprinted in *World Press Review*, July, p. 36.
- Salisbury, Harrison. 1992. *The New Emperors: China in the Era of Mao and Deng*. Boston: Little, Brown.
- Scammel, Michael. 1990. "Yugoslavia: The Awakening." *New York Review of Books*, June 28, pp. 42–47.
- Shlapentokh, Vladimir. 1987. "Soviet People: Too Rich for Reform." *New York Times*, November 23.
- . 1989. *Public and Private Life of the Soviet People*. New York: Oxford University Press.
- Silk, Leonard. 1990. "Soviet Crisis Worse, Economists Declare." *New York Times*, March 15.
- Smith, Hedrick. 1976. *The Russians*. New York: Quadrangle.
- Tumin, Melvin M. 1953. "Some Principles of Stratification: A Critical Analysis." *American Sociological Review* 18:387–394.
- Voinovich, Vladimir. 1989. *The Fur Hat*. New York: Harcourt Brace Jovanovich.
- Voslensky, Michael. 1984. *Nomenklatura: The Soviet Ruling Class*. Garden City, N.Y.: Doubleday.
- Washington Post*. 1990. May 6.
- Zakharov, Mark. 1990. "A Glimpse in 1990: Politics and Democracy." *Literaturnaya Gazeta* (international ed.), March, p. 5.
- Zeitlin, Maurice. 1970. *Revolutionary Politics and the Cuban Working Class*. New York: Harper & Row.

The Dysfunctions of Inequality

4 ■ Melvin M. Tumin

Some Principles of Stratification

A Critical Analysis

The fact of social inequality in human society is marked by its ubiquity and its antiquity. Every known society, past and present, distributes its scarce and demanded goods and services unequally. And there are attached to the positions which command unequal amounts of such goods and services certain highly morally toned evaluations of their importance for the society.

The ubiquity and the antiquity of such inequality has given rise to the assumption that there must be something both inevitable and positively functional about such social arrangements.

Clearly, the truth or falsity of such an assumption is a strategic question for any general theory of social organization. It is therefore most curious that the basic premises and implications of the assumption have only been most casually explored by American sociologists.

The most systematic treatment is to be found in the well-known article by Kingsley Davis and Wilbert Moore, entitled "Some Principles of Stratification."¹ More than twelve years have passed since its publication, and though it is one of the very few treatments of stratification on a high level of gen-

eralization, it is difficult to locate a single systematic analysis of its reasoning. It will be the principal concern of this paper to present the beginnings of such an analysis.

The central argument advanced by Davis and Moore can be stated in a number of sequential propositions, as follows:

1. Certain positions in any society are functionally more important than others, and require special skills for their performance.
2. Only a limited number of individuals in any society have the talents which can be trained into the skills appropriate to these positions.
3. The conversion of talents into skills involves a training period during which sacrifices of one kind or another are made by those undergoing the training.
4. In order to induce the talented persons to undergo these sacrifices and acquire the training, their future positions must carry an inducement value in the form of differential, i.e., privileged and disproportionate access to the scarce and desired rewards which the society has to offer.²

Melvin M. Tumin. "Some Principles of Stratification: A Critical Analysis," *American Sociological Review* 18 (August 1953), pp. 387–394.

5. These scarce and desired goods consist of the rights and perquisites attached to, or built into, the positions, and can be classified into those things which contribute to (a) sustenance and comfort, (b) humor and diversion, (c) self-respect and ego expansion.
6. This differential access to the basic rewards of the society has as a consequence the differentiation of the prestige and esteem which various strata acquire. This may be said, along with the rights and perquisites, to constitute institutionalized social inequality, i.e., stratification.
7. Therefore, social inequality among different strata in the amounts of scarce and desired goods, and the amounts of prestige and esteem which they receive, is both positively functional and inevitable in any society.

Let us take these propositions and examine them *seriatim*.³

(1) *Certain positions in any society are more functionally important than others, and require special skills for their performance.*

The key term here is “functionally important.” The functionalist theory of social organization is by no means clear and explicit about this term. The minimum common referent is to something known as the “survival value” of a social structure.⁴ This concept immediately involves a number of perplexing questions. Among these are: (a) the issue of minimum vs. maximum survival, and the possible empirical referents which can be given to those terms; (b) whether such a proposition is a useless tautology since any *status quo* at any given moment is nothing more and nothing less than everything present in the *status quo*. In these terms, all acts and structures must be judged positively functional in that they constitute essential portions of the *status quo*; (c) what kind of calculus of functionality exists which will enable us, at this point in our

development, to add and subtract long- and short-range consequences, with their mixed qualities, and arrive at some summative judgment regarding the rating an act or structure should receive on a scale of greater or lesser functionality? At best, we tend to make primarily intuitive judgments. Often enough, these judgments involve the use of value-laden criteria, or, at least, criteria which are chosen in preference to others not for any sociologically systematic reasons but by reason of certain implicit value preferences.

Thus, to judge that the engineers in a factory are functionally more important to the factory than the unskilled workmen involves a notion regarding the dispensability of the unskilled workmen, or their replaceability, relative to that of the engineers. But this is not a process of choice with infinite time dimensions. For at some point along the line one must face the problem of adequate motivation for *all* workers at all levels of skill in the factory. In the long run, *some* labor force of unskilled workmen is as important and as indispensable to the factory as *some* labor force of engineers. Often enough, the labor force situation is such that this fact is brought home sharply to the entrepreneur in the short run rather than in the long run.

Moreover, the judgment as to the relative indispensability and replaceability of a particular segment of skills in the population involves a prior judgment about the bargaining-power of that segment. But this power is itself a culturally shaped *consequence* of the existing system of rating, rather than something inevitable in the nature of social organization. At least the contrary of this has never been demonstrated, but only assumed.

A generalized theory of social stratification must recognize that the prevailing system of inducements and rewards is only one of many variants in the whole range of possible systems of motivation which, at least theoretically, are capable of working in human society. It is quite conceivable, of course, that a system of norms could be institutionalized in which the

idea of threatened withdrawal of services, except under the most extreme circumstances, would be considered as absolute moral anathema. In such a case, the whole notion of relative functionality, as advanced by Davis and Moore, would have to be radically revised.

(2) Only a limited number of individuals in any society have the talents which can be trained into the skills appropriate to these positions (i.e., the more functionally important positions).

The truth of this proposition depends at least in part on the truth of proposition 1 above. It is, therefore, subject to all the limitations indicated above. But for the moment, let us assume the validity of the first proposition and concentrate on the question of the rarity of appropriate talent.

If all that is meant is that in every society there is a *range* of talent, and that some members of any society are by nature more talented than others, no sensible contradiction can be offered, but a question must be raised here regarding the amount of sound knowledge present in any society concerning the presence of talent in the population.

For in every society there is some demonstrable ignorance regarding the amount of talent present in the population. *And the more rigidly stratified a society is, the less chance does that society have of discovering any new facts about the talents of its members.* Smoothly working and stable systems of stratification, wherever found, tend to build-in obstacles to the further exploration of the range of available talent. This is especially true in those societies where the opportunity to discover talent in any one generation varies with the differential resources of the parent generation. Where, for instance, access to education depends upon the wealth of one's parents, and where wealth is differentially distributed, large segments of the population are likely to be deprived of the chance even to *discover* what are their talents.

Whether or not differential rewards and opportunities are functional in any one

generation, it is clear that if those differentials are allowed to be socially inherited by the next generation, then the stratification system is specifically dysfunctional for the discovery of talents in the next generation. In this fashion, systems of social stratification tend to limit the chances available to maximize the efficiency of discovery, recruitment and training of "functionally important talent."⁵

Additionally, the unequal distribution of rewards in one generation tends to result in the unequal distribution of motivation in the succeeding generation. Since motivation to succeed is clearly an important element in the entire process of education, the unequal distribution of motivation tends to set limits on the possible extensions of the educational system, and hence, upon the efficient recruitment and training of the widest body of skills available in the population.⁶

Lastly, in this context, it may be asserted that there is some noticeable tendency for elites to restrict further access to their privileged positions, once they have sufficient power to enforce such restrictions. This is especially true in a culture where it is possible for an elite to contrive a high demand and a proportionately higher reward for its work by restricting the numbers of the elite available to do the work. The recruitment and training of doctors in the modern United States is at least partly a case in point.

Here, then, are three ways, among others which could be cited, in which stratification systems, once operative, tend to reduce the survival value of a society by limiting the search, recruitment and training of functionally important personnel far more sharply than the facts of available talent would appear to justify. It is only when there is genuinely equal access to recruitment and training for all potentially talented persons that differential rewards can conceivably be justified as functional. And stratification systems are apparently *inherently antagonistic* to the development of such full equality of opportunity.

(3) *The conversion of talents into skills involves a training period during which sacrifices of one kind or another are made by those undergoing the training.*

Davis and Moore introduce here a concept, "sacrifice," which comes closer than any of the rest of their vocabulary of analysis to being a direct reflection of the rationalizations, offered by the more fortunate members of a society, of the rightness of their occupancy of privileged positions. It is the least critically thought-out concept in the repertoire, and can also be shown to be least supported by the actual facts.

In our present society, for example, what are the sacrifices which talented persons undergo in the training period? The possibly serious losses involve the surrender of earning power and the cost of the training. The latter is generally borne by the parents of the talented youth undergoing training, and not by the trainees themselves. But this cost tends to be paid out of income which the parents were able to earn generally by virtue of *their* privileged positions in the hierarchy of stratification. That is to say, the parents' ability to pay for the training of their children is part of the differential *reward* they, the parents, received for their privileged positions in the society. And to charge this sum up against sacrifices made by the youth is falsely to perpetrate a bill or a debt already paid by the society to the parents.

So far as the sacrifice of earning power by the trainees themselves is concerned, the loss may be measured relative to what they might have earned had they gone into the labor market instead of into advanced training for the "important" skills. There are several ways to judge this. One way is to take all the average earnings of age peers who did go into the labor market for a period equal to the average length of the training period. The total income, so calculated, roughly equals an amount which the elite can, on the average, earn back in the first decade of professional work, over and above the earnings of his age

peers who are not trained. Ten years is probably the maximum amount needed to equalize the differential.⁷ There remains, on the average, twenty years of work during each of which the skilled person then goes on to earn far more than his unskilled age peers. And, what is often forgotten, there is then still another ten- or fifteen-year period during which the skilled person continues to work and earn when his unskilled age peer is either totally or partially out of the labor market by virtue of the attrition of his strength and capabilities.

One might say that the first ten years of differential pay is perhaps justified, in order to regain for the trained person what he lost during his training period. But it is difficult to imagine what would justify continuing such differential rewards beyond that period.

Another and probably sounder way to measure how much is lost during the training period is to compare the per capita income available to the trainee with the per capita income of the age peer on the untrained labor market during the so-called sacrificial period. If one takes into account the earlier marriage of untrained persons, and the earlier acquisition of family dependents, it is highly dubious that the per capita income of the wage worker is significantly larger than that of the trainee. Even assuming, for the moment, that there is a difference, the amount is by no means sufficient to justify a lifetime of continuing differentials.

What tends to be completely overlooked, in addition, are the psychic and spiritual rewards which are available to the elite trainees by comparison with their age peers in the labor force. There is, first, the much higher prestige enjoyed by the college student and the professional-school student as compared with persons in shops and offices. There is, second, the extremely highly valued privilege of having greater opportunity for self-development. There is, third, all the psychic gain involved in being allowed to delay the assumption of adult responsibilities such as earning a living and supporting a family.

There is, fourth, the access to leisure and freedom of a kind not likely to be experienced by the persons already at work.

If these are never taken into account as rewards of the training period it is not because they are not concretely present, but because the emphasis in American concepts of reward is almost exclusively placed on the material returns of positions. The emphases on enjoyment, entertainment, ego enhancement, prestige and esteem are introduced only when the differentials in these which accrue to the skilled positions need to be justified. If these other rewards were taken into account, it would be much more difficult to demonstrate that the training period, as presently operative, is really sacrificial. Indeed, it might turn out to be the case that even at this point in their careers, the elite trainees were being differentially rewarded relative to their age peers in the labor force.

All of the foregoing concerns the quality of the training period under our present system of motivation and rewards. Whatever may turn out to be the factual case about the present system—and the factual case is moot—the more important theoretical question concerns the assumption that the training period under *any* system must be sacrificial.

There seem to be no good theoretical grounds for insisting on this assumption. For while under any system certain costs will be involved in training persons for skilled positions, these costs could easily be assumed by the society-at-large. Under these circumstances, there would be no need to compensate anyone in terms of differential rewards once the skilled positions were staffed. In short, there would be no need or justification for stratifying social positions on *these* grounds.

(4) In order to induce the talented persons to undergo these sacrifices and acquire the training, their future positions must carry an inducement value in the form of differential, i.e., privileged and disproportionate access to the

scarce and desired rewards which the society has to offer.

Let us assume, for the purposes of the discussion, that the training period is sacrificial and the talent is rare in every conceivable human society. There is still the basic problem as to whether the allocation of differential rewards in scarce and desired goods and services is the only or the most efficient way of recruiting the appropriate talent to these positions.

For there are a number of alternative motivational schemes whose efficiency and adequacy ought at least to be considered in this context. What can be said, for instance, on behalf of the motivation which De Man called “joy in work,” Veblen termed “instinct for workmanship” and which we latterly have come to identify as “intrinsic work satisfaction”? Or, to what extent could the motivation of “social duty” be institutionalized in such a fashion that self interest and social interest come closely to coincide? Or, how much prospective confidence can be placed in the possibilities of institutionalizing “social service” as a widespread motivation for seeking one’s appropriate position and fulfilling it conscientiously?

Are not these types of motivations, we may ask, likely to prove most appropriate for precisely the “most functionally important positions”? Especially in a mass industrial society, where the vast majority of positions become standardized and routinized, it is the skilled jobs which are likely to retain most of the quality of “intrinsic job satisfaction” and be most readily identifiable as socially serviceable. Is it indeed impossible then to build these motivations into the socialization pattern to which we expose our talented youth?

To deny that such motivations could be institutionalized would be to overclaim our present knowledge. In part, also, such a claim would seem to derive from an assumption that what has not been institutionalized yet in human affairs is incapable of institutionalization. Admittedly, historical experience affords us evidence we cannot afford to

ignore. But such evidence cannot legitimately be used to deny absolutely the possibility of heretofore untried alternatives. Social innovation is as important a feature of human societies as social stability.

On the basis of these observations, it seems that Davis and Moore have stated the case much too strongly when they insist that a “functionally important position” which requires skills that are scarce, “must command great prestige, high salary, ample leisure, and the like,” if the appropriate talents are to be attracted to the position. Here, clearly, the authors are postulating the unavoidability of very specific types of rewards and, by implication, denying the possibility of others.

(5) These scarce and desired goods consist of the rights and perquisites attached to, or built into, the positions, and can be classified into those things which contribute to (a) sustenance and comfort, (b) humor and diversion, (c) self-respect and ego expansion.

(6) This differential access to the basic rewards of the society has as a consequence the differentiation of the prestige and esteem which various strata acquire. This may be said, along with the rights and perquisites, to constitute institutionalized social inequality, i.e., stratification.

With the classification of the rewards offered by Davis and Moore there need be little argument. Some question must be raised, however, as to whether any reward system, built into a general stratification system, must allocate equal amounts of all three types of reward in order to function effectively, or whether one type of reward may be emphasized to the virtual neglect of others. This raises the further question regarding which type of emphasis is likely to prove most effective as a differential inducer. Nothing in the known facts about human motivation impels us to favor one type of reward over the other, or to insist that all three types of reward must be built into the positions in comparable amounts if the position is to have an inducement value.

It is well known, of course, that societies differ considerably in the kinds of rewards they emphasize in their efforts to maintain a reasonable balance between responsibility and reward. There are, for instance, numerous societies in which the conspicuous display of differential economic advantage is considered extremely bad taste. In short, our present knowledge commends to us the possibility of considerable plasticity in the way in which different types of rewards can be structured into a functioning society. This is to say, it cannot yet be demonstrated that it is *unavoidable* that differential prestige and esteem shall accrue to positions which command differential rewards in power and property.

What does seem to be unavoidable is that differential prestige shall be given to those in any society who conform to the normative order as against those who deviate from that order in a way judged immoral and detrimental. On the assumption that the continuity of a society depends on the continuity and stability of its normative order, some such distinction between conformists and deviants seems inescapable.

It also seems to be unavoidable that in any society, no matter how literate its tradition, the older, wiser and more experienced individuals who are charged with the enculturation and socialization of the young must have more power than the young, on the assumption that the task of effective socialization demands such differential power.

But this differentiation in prestige between the conformist and the deviant is by no means the same distinction as that between strata of individuals each of which operates *within* the normative order, and is composed of adults. The *latter* distinction, in the form of differentiated rewards and prestige between social strata, is what Davis and Moore, and most sociologists, consider the structure of a stratification system. The *former* distinctions have nothing necessarily to do with the workings of such a system nor with the effi-

ciency of motivation and recruitment of functionally important personnel.

Nor does the differentiation of power between young and old necessarily create differentially valued strata. For no society rates its young as less morally worthy than its older persons, no matter how much differential power the older ones may temporarily enjoy.

(7) Therefore, social inequality among different strata in the amounts of scarce and desired goods, and the amounts of prestige and esteem which they receive, is both positively functional and inevitable in any society.

If the objections which have heretofore been raised are taken as reasonable, then it may be stated that the only items which any society *must* distribute unequally are the power and property necessary for the performance of different tasks. If such differential power and property are viewed by all as commensurate with the differential responsibilities, and if they are culturally defined as *resources* and not as rewards, then, no differentials in prestige and esteem need follow.

Historically, the evidence seems to be that every time power and property are distributed unequally, no matter what the cultural definition, prestige and esteem differentiations have tended to result as well. Historically, however, no systematic effort has ever been made, under propitious circumstances, to develop the tradition that each man is as socially worthy as all other men so long as he performs his appropriate tasks conscientiously. While such a tradition seems utterly utopian, no known facts in psychological or social science have yet demonstrated its impossibility or its dysfunctionality for the continuity of a society. The achievement of a full institutionalization of such a tradition seems far too remote to contemplate. Some successive approximations at such a tradition, however, are not out of the range of prospective social innovation.

What, then, of the “positive functionality” of social stratification? Are there other, nega-

tive, functions of institutionalized social inequality which can be identified, if only tentatively? Some such dysfunctions of stratification have already been suggested in the body of this paper. Along with others they may now be stated, in the form of provisional assertions, as follows:

1. Social stratification systems function to limit the possibility of discovery of the full range of talent available in a society. This results from the fact of unequal access to appropriate motivation, channels of recruitment and centers of training.
2. In foreshortening the range of available talent, social stratification systems function to set limits upon the possibility of expanding the productive resources of the society, at least relative to what might be the case under conditions of greater equality of opportunity.
3. Social stratification systems function to provide the elite with the political power necessary to procure acceptance and dominance of an ideology which rationalizes the *status quo*, whatever it may be, as “logical,” “natural” and “morally right.” In this manner, social stratification systems function as essentially conservative influences in the societies in which they are found.
4. Social stratification systems function to distribute favorable self-images unequally throughout a population. To the extent that such favorable self-images are requisite to the development of the creative potential inherent in men, to that extent stratification systems function to limit the development of this creative potential.
5. To the extent that inequalities in social rewards cannot be made fully acceptable to the less privileged in a society, social stratification systems function to encourage hostility, suspicion and distrust among the various

segments of a society and thus to limit the possibilities of extensive social integration.

6. To the extent that the sense of significant membership in a society depends on one's place on the prestige ladder of the society, social stratification systems function to distribute unequally the sense of significant membership in the population.
7. To the extent that loyalty to a society depends on a sense of significant membership in the society, social stratification systems function to distribute loyalty unequally in the population.
8. To the extent that participation and apathy depend upon the sense of significant membership in the society, social stratification systems function to distribute the motivation to participate unequally in a population.

Each of the eight foregoing propositions contains implicit hypotheses regarding the consequences of unequal distribution of rewards in a society in accordance with some notion of the functional importance of various positions. These are empirical hypotheses, subject to test. They are offered here only as exemplary of the kinds of consequences of social stratification which are not often taken into account in dealing with the problem. They should also serve to reinforce the doubt that social inequality is a device which is uniformly functional for the role of guaranteeing that the most important tasks in a society will be performed conscientiously by the most competent persons.

The obviously mixed character of the functions of social inequality should come as no surprise to anyone. If sociology is sophisticated in any sense, it is certainly with regard to its awareness of the mixed nature of any social arrangement, when the observer takes into account long- as well as short-range consequences and latent as well as manifest dimensions.

Summary

In this paper, an effort has been made to raise questions regarding the inevitability and positive functionality of stratification, or institutionalized social inequality in rewards, allocated in accordance with some notion of the greater and lesser functional importance of various positions. The possible alternative meanings of the concept "functional importance" has been shown to be one difficulty. The question of the scarcity or abundance of available talent has been indicated as a principal source of possible variation. The extent to which the period of training for skilled positions may reasonably be viewed as sacrificial has been called into question. The possibility has been suggested that very different types of motivational schemes might conceivably be made to function. The separability of differentials in power and property considered as resources appropriate to a task from such differentials considered as rewards for the performance of a task has also been suggested. It has also been maintained that differentials in prestige and esteem do not necessarily follow upon differentials in power and property when the latter are considered as appropriate resources rather than rewards. Finally, some negative functions, or dysfunctions, of institutionalized social inequality have been tentatively identified, revealing the mixed character of the outcome of social stratification, and casting doubt on the contention that

Social inequality is thus an unconsciously evolved device by which societies insure that the most important positions are conscientiously filled by the most qualified persons.⁸

NOTES

The writer has had the benefit of a most helpful criticism of the main portions of this paper by Professor W. J. Goode of Columbia University. In addition, he has had the opportunity to expose this paper to criticism by the Staff Seminar of the Sociology Section

at Princeton. In deference to a possible rejoinder by Professors Moore and Davis, the writer has not revised the paper to meet the criticisms which Moore has already offered personally.

1. *American Sociological Review*, 10 (April, 1945), pp. 242–249. An earlier article by Kingsley Davis, entitled, “A Conceptual Analysis of Stratification,” *American Sociological Review*, VII (June, 1942), pp. 309–321, is devoted primarily to setting forth a vocabulary for stratification analysis. A still earlier article by Talcott Parsons, “An Analytical Approach to the Theory of Social Stratification,” *American Journal of Sociology*, XLV (November, 1940), pp. 849–862, approaches the problem in terms of why “differential ranking is considered a really fundamental phenomenon of social systems and what are the respects in which such ranking is important.” The principal line of integration asserted by Parsons is with the fact of the normative orientation of any society. Certain crucial lines of connection are left unexplained, however, in this article, and in the Davis and Moore article of 1945 only some of these lines are made explicit.

2. The “scarcity and demand” qualities of goods and services are never explicitly mentioned by Davis and Moore. But it seems to the writer that the argument makes no sense unless the goods and services are so characterized. For if rewards are to function as differential inducements they must not only be differentially distributed but they must be both scarce

and demanded as well. Neither the scarcity of an item by itself nor the fact of its being in demand is sufficient to allow it to function as a differential inducement in a system of unequal rewards. Leprosy is scarce and oxygen is highly demanded.

3. The arguments to be advanced here are condensed versions of a much longer analysis entitled, *An Essay on Social Stratification*. Perforce, all the reasoning necessary to support some of the contentions cannot be offered within the space limits of this article.

4. Davis and Moore are explicitly aware of the difficulties involved here and suggest two “independent clues” other than survival value. See footnote 3 on p. 244 of their article.

5. Davis and Moore state this point briefly on p. 248 but do not elaborate it.

6. In the United States, for instance, we are only now becoming aware of the amount of productivity we, as a society, lose by allocating inferior opportunities and rewards, and hence, inferior motivation, to our Negro population. The actual amount of loss is difficult to specify precisely. Some rough estimate can be made, however, on the assumption that there is present in the Negro population about the same range of talent that is found in the White population.

7. These are only very rough estimates, of course, and it is certain that there is considerable income variation within the so-called elite group, so that the proposition holds only relatively more or less.

8. Davis and Moore, *op. cit.*, p. 243.

5 ■ Claude S. Fischer, Michael Hout, Martín Sánchez Jankowski, Samuel R. Lucas, Ann Swidler, and Kim Voss

Inequality by Design

Why do some Americans have a lot more than others? Perhaps, inequality follows inevitably from human nature. Some people are born with more talent than others; the first succeed while the others fail in life’s competition. Many people accept this explanation,

but it will not suffice. Inequality is not fated by nature, nor even by the “invisible hand” of the market; it is a social construction, a result of our historical acts. *Americans have created the extent and type of inequality we have, and Americans maintain it.*

To answer the question of what explains inequality in America, we must divide it in two. First, who gets ahead and who falls behind in the competition for success? Second, what determines how much people get for being ahead or behind? To see more clearly that the two questions are different, think of a ladder that represents the ranking of affluence in a society. Question one asks why this person rather than that person ended up on a higher or lower rung. Question two asks why some societies have tall and narrowing ladders—ladders that have huge distances between top and bottom rungs and that taper off at the top so that there is room for only a few people—while other societies have short and broad ladders—ladders with little distance between top and bottom and with lots of room for many people all the way to the top.

The answer to the question of who ends up where is that people's social environments largely influence what rung of the ladder they end up on.¹ The advantages and disadvantages that people inherit from their parents, the resources that their friends can share with them, the quantity and quality of their schooling, and even the historical era into which they are born boost some up and hold others down. The children of professors, our own children, have substantial head starts over children of, say, factory workers. Young men who graduated from high school in the booming 1950s had greater opportunities than the ones who graduated during the Depression. Context matters tremendously.

The answer to the question of why societies vary in their structure of rewards is more political. In significant measure, societies choose the height and breadth of their "ladders." By loosening markets or regulating them, by providing services to all citizens or rationing them according to income, by subsidizing some groups more than others, societies, through their politics, build their ladders. To be sure, historical and external constraints deny full freedom of action, but a

substantial freedom of action remains. In a democracy, this means that the inequality Americans have is, in significant measure, the historical result of policy choices Americans—or, at least, Americans' representatives—have made. In the United States, the result is a society that is distinctively *unequal*. Our ladder is, by the standards of affluent democracies and even by the standards of recent American history, unusually extended and narrow—and becoming more so.

To see how policies shape the structure of rewards (i.e., the equality of outcomes), consider these examples: Laws provide the ground rules for the marketplace—rules covering incorporation, patents, wages, working conditions, unionization, security transactions, taxes, and so on. Some laws widen differences in income and earnings among people in the market; others narrow differences. Also, many government programs affect inequality more directly through, for example, tax deductions, food stamps, social security, Medicare, and corporate subsidies.

To see how policies also affect which particular individuals get to the top and which fall to the bottom of our ladder (i.e., the equality of opportunity), consider these examples: The amount of schooling young Americans receive heavily determines the jobs they get and the income they make. In turn, educational policies—what sorts of schools are provided, the way school resources are distributed (usually according to the community in which children live), teaching methods such as tracking, and so on—strongly affect how much schooling children receive. Similarly, local employment opportunities constrain how well people can do economically. Whether and where governments promote jobs or fail to do so will, in turn, influence who is poised for well-paid employment and who is not.

Claiming that intentional policies have significantly constructed the inequalities we have and that other policies could change those inequalities may seem a novel idea in

the current ideological climate. So many voices tell us that inequality is the result of individuals' "natural" talents in a "natural" market. Nature defeats any sentimental efforts by society to reduce inequality, they say; such efforts should therefore be dropped as futile and wasteful. Appeals to nature are common and comforting. As Kenneth Bock wrote in his study of social philosophy, "We have been quick to seek explanations of our problems and failures in what we *are* instead of what we *do*. We seem wedded to the belief that our situation is a consequence of our nature rather than of our historical acts."² In this case, appeals to nature are shortsighted.

Arguments from nature are useless for answering the question of what determines the structure of rewards because that question concerns differences in equality *among societies*. Theories of natural inequality cannot tell us why countries with such similar genetic stocks (and economic markets) as the United States, Canada, England, and Sweden can vary so much in the degree of economic inequality their citizens experience. The answer lies in deliberate policies.

Appeals to nature also cannot satisfactorily answer even the first question: Why do some *individuals* get ahead and some fall behind? Certainly, genetic endowment helps. Being tall, slender, good-looking, healthy, male, and white helps in the race for success, and these traits are totally or partly determined genetically. But these traits matter to the degree that society makes them matter—determining how much, for example, good looks or white skin are rewarded. More important yet than these traits are the social milieux in which people grow up and live.

Realizing that intentional policies account for much of our expanding inequality is not only more accurate than theories of natural inequality; it is also more optimistic. We are today more unequal than we have been in seventy years. We are more unequal than any other affluent Western nation. Intentional policies could change those conditions, could

reduce and reverse our rush to a polarized society, could bring us closer to the average inequality in the West, could expand both equality of opportunity and equality of result.

Still, the "natural inequality" viewpoint is a popular one. Unequal outcomes, the best-selling *Bell Curve* argues, are the returns from a fair process that sorts people out according to how intelligent they are.³ But *The Bell Curve's* explanation of inequality is inadequate. The authors err in assuming that human talents can be reduced to a single, fixed, and essentially innate skill they label intelligence. They err in asserting that this trait largely determines how people end up in life. And they err in imagining that individual competition explains the structure of inequality in society. . . .

Disparities in income and wealth, [other] analysts argue, encourage hard work and saving. The rich, in particular, can invest their capital in production and thus create jobs for all.⁴ This was the argument of "supply-side" economics in the 1980s, that rewarding the wealthy—for example, by reducing income taxes on returns from their investments—would stimulate growth to the benefit of all. The 1980s did not work out that way, but the theory is still influential. We *could* force more equal outcomes, these analysts say, but doing so would reduce living standards for all Americans.

Must we have so much inequality for overall growth? The latest economic research concludes *not*; it even suggests that inequality may *retard* economic growth. In a detailed statistical analysis, economists Torsten Persson and Guido Tabellini reported finding that, historically, societies that had more inequality of earnings tended to have lower, not higher, subsequent economic growth. Replications by other scholars substantiated the finding: More unequal nations grew less quickly than did more equal societies.⁵ . . .

This recent research has not demonstrated precisely how greater equality helps economic growth,⁶ but we can consider a few possibilities. Increasing resources for those of lower

income might, by raising health, educational attainment, and hope, increase people's abilities to be productive and entrepreneurial. Reducing the income of those at the top might reduce unproductive and speculative spending. Take, as a concrete example, the way American corporations are run compared with German and Japanese ones. The American companies are run by largely autonomous managers whose main responsibility is to return short-term profits and high stock prices to shareholders and—because they are often paid in stock options—to themselves as well. Japanese and German managers are more like top employees whose goals largely focus on keeping the company a thriving enterprise. The latter is more conducive to reinvesting profits and thus to long-term growth.⁷ Whatever the mechanisms may be, inequality appears to undermine growth. Americans certainly need not feel that they must accept the high levels of inequality we currently endure in order to have a robust economy.

A related concern for Americans is whether “leveling” stifles the drive to get ahead. Americans prefer to encourage Horatio Alger striving and to provide opportunities for everyone. Lincoln once said “that some would be rich shows that others may become rich.”⁸ Many, if not most, Americans believe that inequality is needed to encourage people to work hard.⁹ But, if so, *how much* inequality is needed?

For decades, sociologists have been comparing the patterns of social mobility across societies, asking: In which countries are people most likely to overcome the disadvantages of birth and move up the ladder? In particular, does more or less equality encourage such an “open” society? The answer is that Western societies vary little in the degree to which children's economic successes are constrained by their parents' class positions. America, the most unequal Western society, has somewhat more fluid intergenerational mobility than do other nations, but so does Sweden, the most equal Western society.¹⁰ There is no case for encouraging inequality in this evidence, either.

In sum, the assumption that considerable inequality is needed for, or even encourages, economic growth appears to be false. We do not need to make a morally wrenching choice between more affluence and more equality; we can have both. But even if such a choice were necessary, both sides of the debate, the “altruists” who favor intervention for equalizing and the supposed “realists” who resist it, agree that inequality can be shaped by policy decisions: Wittingly or unwittingly, we choose our level of inequality.

NOTES

1. We know that in statistical models of individual status attainment much, if not most, of the variance is unaccounted for. Of the explained variance, however, the bulk is due to social environment broadly construed. Also, we believe that much of the residual, unexplained variance is attributable to unmeasured social rather than personal factors.

2. Kenneth Bock, *Human Nature Mythology* (Urbana 1994), p. 9.

3. Richard J. Herrnstein and Charles Murray, *The Bell Curve: Intelligence and Class Structure in American Life* (New York 1994).

4. See, for example, Rich Thomas, “Rising Tide Lifts the Yachts: The Gap Between Rich and Poor Has Widened, but There Are Some Comforting Twists,” *Newsweek*, May 1, 1995. See also George Will, “What's Behind Income Disparity,” *San Francisco Chronicle*, April 24, 1995.

5. Torsten Persson and Guido Tabellini, “Is Inequality Harmful for Growth?,” *American Economic Review* 84, 1994; Roberto Chang, “Income Inequality and Economic Growth: Evidence and Recent Theories,” *Economic Review* 79, 1994; George R. G. Clarke, “More Evidence on Income Distribution and Growth,” *Journal of Development Economics* 47, 1995. See also Peter H. Lindert, “The Rise of Social Spending,” *Explorations in Economic History* 31, 1994.

6. Persson and Tabellini's explanation (“Is Inequality Harmful?”) for their results is that in societies with greater earnings inequality, there is less political pressure for government redistribution; such redistribution impairs growth. However, their evidence for the explanation is thin, and Clarke's results (“More Evidence”) are inconsistent with that argument. Chang (“Income Inequality”) suggests that with more equality, lower-income families could make longer-term investment decisions. In any

event, the statistical results suggest that government intervention on behalf of equality in the market, rather than after the market, would be beneficial.

7. See, for example, Michael Porter, *Capital Choices: Changing the Way America Invests in Industry* (Washington 1992).

8. Quoted by Alan Trachtenberg, *The Incorporation of America: Culture and Society in the Gilded Age* (New York 1982), p. 75.

9. See, for example, Lee Rainwater, *What Money Buys: Inequality and the Social Meanings of Income* (New York 1974); James R. Kluegel and E. R. Smith, “Beliefs About Stratification,” *Annual Review of Sociology* 7, 1981.

10. Harry B. G. Ganzeboom, Donald J. Treiman, and Wout C. Ultee, “Comparative Intergenerational Stratification Research,” *Annual Review of Sociology* 17, 1991.

6 ■ Alan B. Krueger

Inequality, Too Much of a Good Thing

As the title of this essay suggests, I believe inequality has both positive and negative effects. On the positive side, differential rewards provide incentives for individuals to work hard, invest, and innovate. On the negative side, differences in rewards that are unrelated to productivity—those that result from racial discrimination, for example—are corrosive to civil society and cause resources to be misallocated. Even if discrimination did not exist, however, income inequality is problematic in a democratic society if those who are privileged use their economic muscle to curry favor in the political arena and thereby secure monopoly rents or other advantages. Moreover, for several reasons discussed in the next section, poverty and income inequality create negative externalities.¹ Consequently, it can be in the interest of the wealthy as well as the poor to raise the incomes of the poor, especially by using education and training as a means for redistribution.

The term *inequality* is often used rather loosely and can be a lightning rod. Some have argued that only extreme poverty is a concern. Others have argued that the gap in

income or wealth between the well off and the poor is a concern. Yet others have argued that the rapid growth in income disparity between the richest of the rich and everyone else is an issue. I will argue that, for various reasons elaborated below, all of these forms of inequality are of concern to contemporary American society, and that America has reached a point at which, on the margin, efficiently redistributing income from rich to poor is in the nation’s interest.

A theme of my contribution to this debate is that societies must strike a *balance* between the beneficial incentive effects of inequality and the harmful welfare-decreasing effects of inequality. The optimal balance will differ across societies and time, but too much inequality can be harmful in any society, just as too much equality can suppress innovation and drive. . . .

Why Care about Rising Inequality?

Philosophers have argued about income inequality and social justice for centuries. I will sidestep most of that debate. What follows is

Alan B. Krueger. “Inequality, Too Much of a Good Thing,” in *Inequality in America*, edited by Benjamin M. Friedman, pp. 1–2, 11–24, 62, 67–75. Copyright © 2003 by MIT Press.

a thumbnail sketch of reasons why I think it is in our interest for U.S. public policy to try to restore a more balanced distribution of income in the country. Because such a conclusion fundamentally rests on one's values as well as an empirical view of the world, I will touch lightly on these reasons. Suffice it to say that I hope there are enough arguments here to persuade the reader that it is worth considering using education and training as part of an overall strategy to reduce income inequality in America.

Philosophy

As Atkinson (1983) observes, “different principles of justice lead to quite different views about inequality” (p. 5). Principles of justice provide guidelines for society's welfare function, and with a welfare function for a particular society, economists could judge the distribution of inequality in that society against the optimal level. The rub, of course, is that the welfare function in a society is *not* observable and depends on philosophical arguments that are not testable. In addition, one has the Arrow impossibility theorem² with which to contend. Consequently, appealing to philosophical arguments can never be universally dispositive.

Principles of social justice can be divided into those that focus on fair exchange starting from a just distribution of endowments, and those that focus on the equality of outcomes. Rawls invites readers to arrive at a theory of justice by selecting the principles they would desire if they were choosing such principles in an original position behind a veil of ignorance, unaware of their standing in society or initial endowment of talents. He argues that in this case the social justice that would be desired would involve two principles: one protecting liberties and the other providing for an egalitarian distribution of opportunities and material goods. This leads him to a maximin welfare function in which the well-being of the worst off in society should be as high as possible. Interestingly,

Adam Smith arrived at a somewhat similar conclusion nearly two hundred years earlier, positing that “[n]o society can surely be flourishing and happy, of which the far greater part of the members are poor and miserable. It is but equity, besides, that they who feed, clothe and lodge the whole body of the people, should have such a share of the produce of their own labour as to be themselves tolerably well fed, clothed and lodged” (Smith 1776, 110–111).

Nozick (1974) questions whether a theory of justice can be based on the distribution of outcomes. Using the analogy of fans who are willing to pay a fee to watch Wilt Chamberlain play basketball—an updated analogy might substitute Shaquille O’Neal—Nozick argues that “no end-state principle or distributional patterned principle of justice can be continuously realized without continuous interference with people's lives” (p. 163). Who could complain about Wilt Chamberlain's exorbitant salary if it results from rational choices? Nozick also raises the issue of the adverse incentive effects Rawls's theory of justice would have on the acquisition of talent.

Religion

I would argue that religious beliefs provide as strong (or weak) a justification for views toward society's implicit welfare function as do philosophical reflections behind a veil of ignorance. Indeed, I would go further and say that religious tenets reflect the demand for equality among the public. If people did not adhere to the basic tenets of their religion, they would not practice or would eventually change faiths. Thus, long-standing religious views toward inequality provide something of a revealed-preference argument.³ And with regard to wealth inequality, the world's major religions are united in favoring redistribution of wealth toward the poor. Robert Nelson (1991), for example, observes that “Roman Catholicism has traditionally instilled a strong concern for the poor; in the Middle Ages the church itself provided much of the care for the indigent.

The welfare state today similarly accomplishes substantial internal redistribution with the approval of many of the wealthier contributing members of the community” (p. 326). The Jewish *Siddur* advises followers to “Be just to the poor and the orphan; Deal righteously with the afflicted and the destitute” and comments, “Happy are they who are thoughtful of the needy; In time of trouble may the Lord Keep them from harm.” And the Koran criticizes the egoism of the rich inhabitants of Mecca and urges believers in Islam to support poor people, orphans, and captives. Islam requires five major obligations of its followers, including *zakat*, an obligatory contribution to the needy (which today is implemented in the form of a tax).

Enlightened Self-Interest

Another line of argument for achieving and maintaining a minimum level of equality rests on self-interest. If wide disparities in income or education create negative externalities for a majority of people, then it clearly is in the self-interest of members of society to reduce those disparities. Individuals acting on their individual preferences (e.g., paying to see Shaquille O’Neal play basketball) will not internalize these externalities. What might such externalities be? An incomplete list would involve the following.

- More-educated voters make the democratic process work better. First, people with more education are more likely to be informed and more likely to participate in democracy. Second, more-informed citizens are likely (though not guaranteed) to make better decisions. For the latter reason, even a devout defender of free markets like Milton Friedman (1982) can be found supporting a minimum compulsory level of education.
- Available evidence suggests a link between crime and inequality (e.g., Ehrlich 1973; Freeman 1983, 1995; and Imrohorglu, Merlo, and Rupert

2001). Other things being equal, the incentive for those with limited market opportunities to commit property crimes rises as inequality increases. From the criminal’s perspective, the potential gain from crime is higher if inequality is higher, and the opportunity cost is lower. Society can devote more resources to crime prevention and incarceration, or to reducing inequality.

- Society is not willing to allow citizens to be totally destitute, to fall below some minimum level of basic consumption when it comes to food or health care (see, e.g., Pauly 1971). By providing those likely to have low incomes with skills, therefore, and thereby raising their future earnings, society can reduce the cost of providing transfer payments later on. . . .
- Nelson and Phelps (1966) and Romer (1990) model the level of education in a society as generating positive externalities for economic growth, although empirical support for this model is mixed (see Krueger and Lindahl 2001).

Low Wages, Imperfect Monitoring, and Public Safety

In an advanced economy, people are connected via markets in a myriad of ways. Employee performance can be monitored only imperfectly. If an employee performs poorly because he or she feels poorly compensated, others may suffer as a result of his or her poor performance. The tragic events of September 11th, for example, highlighted the importance of paying baggage screeners better wages.

Market Failure

More generally, market failures could lead the distribution of income in a society to be suboptimal. Credit constraints, for example, might prevent children from poor families from investing adequately in education. Monopsony power on the part of employers might enable firms to pay workers less than

the value of their marginal products.⁴ Statistical discrimination might lead to lower-than-optimal investment in education for discriminated-against groups (e.g., Lundberg and Startz 1983).

Efficient Policy Changes

Another type of externality could arise in the political arena in a society if income inequality in that society is viewed as excessive. Treaties to reduce international trade barriers offer an example. A policy that reduces trade barriers undoubtedly will increase national income in the society that enacts it. However, there will be winners and losers from such a policy. If some segments of society feel that they have not benefited from developments in the economy, then they are unlikely to support efforts to reduce trade barriers. I am not talking here just about the losers from the trade barrier reduction policy specifically, who are usually few in number and concentrated in a handful of industries. Instead, views toward free trade seem to be class related. For example, Blendon et al. (1997) find that 72 percent of those with less than a college education say one reason the U.S. economy is not doing better is that “companies are sending jobs overseas,” whereas only 53 percent of college graduates agree with the same statement, and just 6 percent of American Economic Association members. Less-educated people are also less likely than those with more education to respond in the survey that trade agreements are good for the economy. I suspect that one reason Presidents (George W.) Bush and Clinton had difficulty securing fast-track authority for the passage of trade agreements is that large segments of the public perceived that they would lose from free trade, an inference that they drew because they had seen their real incomes stagnate or decline over the previous twenty years while trade expanded. Although I suspect trade has had little to do with rising wage inequality in the United States, it is understandable why so many people would draw such an inference.

Unless the public perceives that it will benefit from more efficient policies, there is little reason to suspect it will support such policies, and with 94 percent of income growth accruing to the top 1 percent of the U.S. population since 1973, it is understandable why the American public might be a little skeptical that it gained from past changes such as expanded trade.

Money Buys Influence

Economists at least since Adam Smith have fretted that wealthy merchants and manufacturers would be led by self-interest to seek government regulation and privilege to protect their monopoly position, thereby preventing the invisible hand from working its magic. Money buys access and influence in politics. It also buys influence through think tanks. A negative consequence of the skewed distribution of income in the United States is that some individuals have much more political influence than others.

Benabou (2000) develops a formal model in which the progressivity of educational funding and taxation is endogenous. He shows that the political influence of the wealthy interacts with income inequality to block efficient progressive policies or impose inefficient regressive ones. When income inequality is high, he finds, the wealthy are more likely to block *efficiency-enhancing* programs that would improve educational opportunities for the less well off.

Growth and Income Inequality

Persson and Tabellini (1994) develop a model of economic growth in which income inequality negatively influences growth through the political process. In their model, income inequality leads to policies that do not protect property rights and therefore do not allow full private appropriation of the returns from investment. A growing body of cross-country and cross-state studies have estimated the relationship between initial income inequality and subsequent GDP growth.⁵ Although at-

tributing causality is difficult in these studies, the correlation between income inequality and growth is negative, conditional on variables like initial GDP per capita and average education. Two-stage least-squares estimates that use variables such as initial literacy and infant mortality as instruments for income inequality also show an inverse relationship between GDP growth and such inequality.

Health and Income Inequality

One common argument I will *not* make concerns health and income inequality. Wilkenson (1996), for example, argues that average health is negatively affected by the societal level of income inequality. The evidence in support of this view, however, is far from compelling (see, e.g., Smith 1999 and Deaton 2001), although Eibner and Evans (2002) provide evidence that relative deprivation affects health, and a large body of evidence finds that a person's own income level is related to his or her health.

Winner-Take-All Inefficiency in Superstar Markets

Frank and Cook (1996) argue that technological changes have facilitated a shift to superstar markets in many top-paying professions. The reward received by the one who finishes first is much greater than the reward received by the also-rans. They lament that this shift is inefficient and inequitable, causing too many students to pursue careers in law, finance, and consulting at the expense of more socially beneficial fields such as engineering, manufacturing, civil service, and teaching. The winner-take-all society engendered by this shift may create the same type of misallocation of talent that Murphy, Shleifer, and Vishny (1991) attribute to rent seeking. To some extent, income inequality probably leads to legions of tax lawyers and lobbyists who look for ways to help wealthy clients avoid taxation. Frank and Cook believe that superstar markets have led to inefficient investment and wasteful competition.

Although I think we are at little risk of becoming a nation of Tonya Hardings, there may be something to the argument that superstar salaries provide perverse incentives and unnecessary competition in some sectors and divert some workers from pursuing more socially rewarding careers.

Public Preference

Last but not least, I would surmise that a majority of the public demands a certain amount of equality and is particularly supportive of using education and training to achieve more equality of outcomes. A survey of 1,001 adults by Lake, Snell, Perry & Associates in July 2000 posed the following request to respondents: "I am going to read some different ways the government can help poor Americans find and keep good jobs. For each, please tell me if you strongly support, somewhat support, somewhat oppose, or strongly oppose this idea." Fully 90 percent supported "helping to pay for education and job training for people leaving welfare." Similarly, a Gallup poll sponsored by General Motors in May 1998 asked the following free-form question: "Just your opinion, in what ways do you think the government should help the poor?" By far, the top two responses were providing better/more affordable education (38 percent) and providing job training/skills training (29 percent). The next-highest response was providing more jobs/job opportunities, at 16 percent. Only 5 percent of those surveyed cited lowering taxes as a way the government might assist poor people.

Even when given an explicit choice of lower taxes, the public prefers education and training. A CBS News poll in September 1999, for example, asked 1,376 respondents, "Which comes closer to your view? Government should provide tools to help families better their lives, such as education and job training programs. The best thing that government can do for families is to cut taxes and allow individual families to decide for themselves how to allocate their money." Fifty-five percent of respondents said the first statement

more closely reflected their views, whereas only 42 percent replied that the second statement came closer to what they believed.

Wrapping Up

In supporting minimum schooling, Milton Friedman (1982) argued that “[a] stable and democratic society is impossible without a minimum degree of literacy and knowledge on the part of most citizens and without widespread acceptance of some common set of values” (p. 86). I would argue that inequality could grow so extreme that it eventually jeopardizes any type of “widespread acceptance” of a democratic capitalist society that might be established. This leads me to agree with Victor Fuchs (1979): “For me the key word is *balance*, both in the goals that we set and in the institutions that we nourish in order to pursue these goals. I value freedom *and* justice *and* efficiency, and economics tells me that I may have to give up a little of one goal to insure the partial achievement of others” (p. 180).

Targeted Education and Training: Part of the Solution

In a perfect world, children from all families would invest in educational resources up to the point that their marginal return equaled their discount rate, and all families would have equal access to credit and discount investments at the same prevailing rate. The evidence suggests, however, that education decisions are not made in a perfect world. Children from poor families *behave* as if they have higher discount rates. The most plausible explanations for this phenomenon are that poor families are credit constrained (i.e., cannot borrow at the same rate as everyone else), or that they discount future benefits of human capital investments at a higher than market rate because they are impatient, have a greater disutility of schooling, or fail to appreciate the benefits of education. Of these possible explanations, credit constraints have received the most attention in the literature,

because students cannot easily use the return on their future human capital as collateral to borrow for human capital investments. This may be a reason for discount rates to vary. Poor families face different borrowing costs than rich ones.

The following five observations are consistent with the view that low-income families face credit constraints when it comes to education. First, Ellwood and Kane (2000) find that when the return to college education increased in the 1980s, four-year college enrollment increased for children from all quartiles of the income distribution except the bottom one. Second, Behrman and Taubman (1990) find that the timing of parental income matters for children’s educational attainment. Using data from the Panel Study of Income Dynamics (PSID), they find that father’s income earned when children are teenagers has a stronger effect on children’s educational attainment than income earned later on. Third, Shea (2000) looks at the effect on children’s human capital of differences in parental income emanating from noncompetitive factors, such as employment in a high-paying union job or industry. Wage differences for reasons such as these arguably are independent of parents’ ability. Shea finds that family income matters for children’s human capital investment in a sample of low-income families, but not for the broader population. He concludes that this finding is consistent with the idea that the accumulation of observable skills by poverty sample fathers may have been suboptimal due to liquidity constraints. Fourth, Björklund and Jantti (1997) find stronger family income effects on children’s outcomes in the United States than in Sweden, which provides much more generous educational subsidies than the United States. Fifth, the reaction of college enrollment to changes in tuition, especially at the two-year-college level, is substantially larger than the reaction of college enrollment to equivalent, present-value changes in the payoff to education (see Kane 1999).

Although it is possible to construct complicated explanations of the above findings that are consistent with all families' having equal access to credit—and I suspect part of the association between education and parental income reflects intergenerational transmission of ability and motivation for schooling, as Cameron and Heckman (2001) argue—Occam's razor and common sense suggest that families have different access to credit. For example, some families borrow for college costs by accumulating debt on their credit cards at exorbitant rates, whereas others tap into their family finances or take out home equity loans that are given tax-preferred treatment.

One does not have to resort to theoretical assumptions or indirect tests of credit constraints, however, to support the view that redistribution of wealth via targeted education and training is desirable. *It is clear that returns to education and training are at least as big at the bottom of the income distribution as at the top.* I have presented evidence elsewhere (Krueger 2003) that the *social return* from investment in education and training for poor children, from infancy through early adulthood, is at least as great as the social return from investments in education and training in the general public.

A theme that emerges from my survey of the evidence (see Krueger 2003) is that the real rate of return from investment in various education and training programs for the disadvantaged is 6 to 11 percent. This range applies to a diverse set of programs, ranging from preschool to Job Corps to conventional K–12 public schools. To put this figure in perspective, note that the historical real rate of return on the stock market has been calculated at 6.3 percent (Burtless 1999). So investment in human capital for the disadvantaged seems to yield at least as great a return as investment in the equity market. Also, because there is not currently universal access to most of the educational and training programs, and many willing participants are thus turned away, I would argue that the re-

turns estimated from various evaluations reviewed below would approximately apply if the programs were greatly expanded to accommodate more participants. . . .

I would emphasize that I do not envision investment in human capital development as the sole component of a program to address the adverse consequences of income inequality. It is part of the solution, not the whole solution. In principle, the optimal government policy regarding income inequality would employ multiple instruments, up to the point at which the social benefit per additional dollar of cost of each instrument is equal across the instruments.

NOTES

1. In referring to “negative externalities,” Krueger is claiming that decisions resulting in poverty and inequality are made without taking into account the full costs of those decisions, just as decisions by firms to pollute are often made without those firms having to bear the full costs of such decisions (e.g., the social costs of the resulting asthma).—EDS.

2. The Arrow impossibility theorem refers in this context to the impossibility of devising a “fair” procedure for translating individual preferences about how much inequality should be allowed into a collective decision about how much inequality should be allowed.—EDS.

3. The simple point that Krueger is making here is that consumers of religion are, by virtue of their decision to consume, revealing a taste for the egalitarian doctrines of most major religions.—EDS.

4. When workers face a single employer to whom they can sell their labor, that employer has “monopsony power” that makes it possible to offer wages that are lower than would be predicted in competitive markets.—EDS.

5. The GDP (Gross Domestic Product) is the total value of the final goods and products produced in a year within a country.—EDS.

REFERENCES

- Atkinson, Anthony. 1983. *The Economics of Inequality*, 2d ed. Oxford: Oxford University Press.
- Behrman, Jere, and Paul Taubman. 1990. “The Intergenerational Correlation between Children's Adult Earnings and Their Parents' Income: Results from the Michigan Panel Survey of Income

- Dynamics." *Review of Income and Wealth* 36, no. 2: 115–127.
- Benabou, Roland. 2000. "Unequal Societies: Income Distribution and the Social Contract." *American Economic Review* 90, no. 1 (March): 96–129.
- Björklund, Anders, and Markus Jantti. 1997. "Intergenerational Income Mobility in Sweden Compared to the United States." *American Economic Review* 87, no. 5 (December): 1009–1018.
- Blendon, Robert, John Benson, Mollyann Brodie, Richard Morin, Drew Altman, Daniel Gitterman, Mario Brossard, and Matt James. 1997. "Bridging the Gap between the Public's and Economists' Views of the Economy." *Journal of Economic Perspectives* 11, no. 3 (Summer): 105–118.
- Burtless, Gary. 1999. "Risk and Returns of Stock Market Investments Held in Individual Retirement Accounts." Testimony before the House Budget Committee Task Force on Social Security Reform, May 11.
- Cameron, Stephen, and James Heckman. 2001. "The Dynamics of Educational Attainment for Black, Hispanic, and White Males." *Journal of Political Economy* 109, no. 3 (June): 455–499.
- Deaton, Angus. 2001. "Health, Inequality, and Economic Development." Research Program in Development Studies working paper, Princeton University.
- Ehrlich, Isaac. 1973. "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation." *Journal of Political Economy* 81, no. 3 (May): 521–565.
- Eibner, Christine, and William Evans. 2002. "Relative Deprivation, Poor Health Habits and Mortality." University of Maryland, College Park, mimeo.
- Ellwood, David, and Thomas Kane. 2000. "Who Is Getting a College Education? Family Background and the Growing Gaps in Enrollment." In *Securing the Future: Investing in Children from Birth to College*, Sheldon Danziger and Jane Waldfogel, eds. New York: Russell Sage.
- Frank, Robert H., and Philip J. Cook. 1996. *The Winner Take All Society*. New York: Free Press.
- Freeman, Richard. 1983. "Crime and the Labor Market." In *Crime and Public Policy*, James Wilson, ed. San Francisco: Institute for Contemporary Studies.
- Freeman, Richard. 1995. "The Labor Market." In *Crime*, James Wilson and Joan Petersilia, eds. San Francisco: Institute for Contemporary Studies.
- Friedman, Milton. 1982. *Capitalism and Freedom*. Chicago: University of Chicago Press.
- Fuchs, Victor. 1979. "Economics, Health, and Post-Industrial Society." *Health and Society* 57, no. 2 (Spring): 153–182.
- Imrohoroğlu, Ayse, Antonio Merlo, and Peter Rupert. 2001. "What Accounts for the Decline in Crime?" Penn Institute for Economic Research working paper no. 01–012.
- Kane, Thomas. 1999. *The Price of Admission: Rethinking How Americans Pay for College*. Washington, D.C.: Brookings Institution Press.
- Krueger, Alan. 2003. "Economic Considerations and Class Size." *Economic Journal*, forthcoming.
- Krueger, Alan, and Mikael Lindahl. 2001. "Education and Growth: Why and for Whom?" *Journal of Economic Literature* 39, no. 4 (December): 1101–1136.
- Lundberg, Shelly, and Richard Startz. 1983. "Private Discrimination and Social Intervention in Competitive Labor Markets." *American Economic Review* 73, no. 3 (June): 340–347.
- Murphy, Kevin M., Andrei Shleifer, and Robert Vishny. 1991. "The Allocation of Talent: Implications for Growth." *Quarterly Journal of Economics* 106, no. 2 (May): 503–530.
- Nelson, Richard, and Edmund Phelps. 1966. "Investment in Humans, Technological Diffusion, and Economic Growth." *American Economic Review* 56, no. 2 (March): 69–75.
- Nelson, Robert. 1991. *Reaching for Heaven on Earth*. Savage, Md.: Rowman & Littlefield.
- Nozick, Robert. 1974. *Anarchy, State, and Utopia*. New York: Basic.
- Pauly, Mark V. 1971. *Medical Care at Public Expense: A Study in Applied Welfare Economics*. New York: Praeger.
- Persson, Torsten, and Guido Tabellini. 1994. "Is Inequality Harmful for Growth?" *American Economic Review* 84, no. 3 (June): 600–621.
- Rawls, John. 1971. *A Theory of Justice*. Cambridge: Belknap Press of Harvard University Press.
- Romer, Paul. 1990. "Endogenous Technological Change." *Journal of Political Economy* 98, no. 5 (October): 71–102.
- Shea, John. 2000. "Does Parents' Money Matter?" *Journal of Public Economics* 77, no. 2 (August): 155–184.
- Smith, Adam. [1776]. *The Wealth of Nations*. New York: Random House.
- Smith, James P. 1999. "Healthy Bodies and Thick Wallets: The Dual Relation between Health and Economic Status." *Journal of Economic Perspectives* 13, no. 2 (Spring): 145–166.
- Wilkenson, Richard. 1996. *Unhealthy Societies: The Afflictions of Inequality*. London: Routledge.

Why Is Income Inequality Increasing?

7 ■ Martina Morris and Bruce Western

Inequality in Earnings

Trends and Implications

Ironically, as some sociologists were debating whether social classes had finally withered away (e.g., Clark and Lipset 1991), earnings inequality in the United States quietly started to grow again. Median income declined and the distribution of income grew markedly more unequal during the last three decades of the twentieth century, reversing a long trend of earnings growth and equalization. Writers in economics and the popular press described, analyzed, and debated this new trend. Sociologists, on the other hand, were, and to a large extent remain, remarkably silent about it. We have continued to address trends in the earnings “gaps” (by gender and race) or focus on poverty alone, leaving the broader trends—stagnation in earnings levels and growing polarization in earnings distributions—to others. We have continued to study how people are allocated to positions in the earnings distribution, rather than the structure of those positions. If the structure had been stable, a narrow focus on allocation might be justifiable. But this has been a period of pervasive economic restructuring, and the impact on earnings distributions both within and between groups has been

profound. Sociological theory provides a rich framework for understanding the broad changes now underway, and we offer this essay as a challenge to the field to critically evaluate the evidence and provide a sociologically informed response.

According to Simon Kuznets (1955), inequality should not be rising in the late twentieth-century United States. The Kuznets curve predicts that economic development has an inverted U-shaped relationship to inequality. Inequality first rises as capital is concentrated in the hands of investors, then falls as economic development generates widespread prosperity. For much of this century, in fact, inequality followed the Kuznets path. Earnings inequality peaked on the eve of the Great Depression, then began a long secular decline.¹ Postwar prosperity was marked by a rise in median earnings and stability in earnings inequality. The annual income of the median worker more than doubled from 1950 to 1970, with those at the bottom of the earnings scale making even greater progress. It was, as many observed in chastened hindsight, a rising tide that lifted all boats. These trends reversed in the early

This is a commissioned chapter that draws heavily on material in a previous publication (Martina Morris and Bruce Western, “Inequality in Earnings at the Close of the Twentieth Century,” *Annual Review of Sociology* 25, pp. 623–657. Copyright © 1999 by Annual Reviews, www.AnnualReviews.org. Reprinted by permission of Annual Reviews).

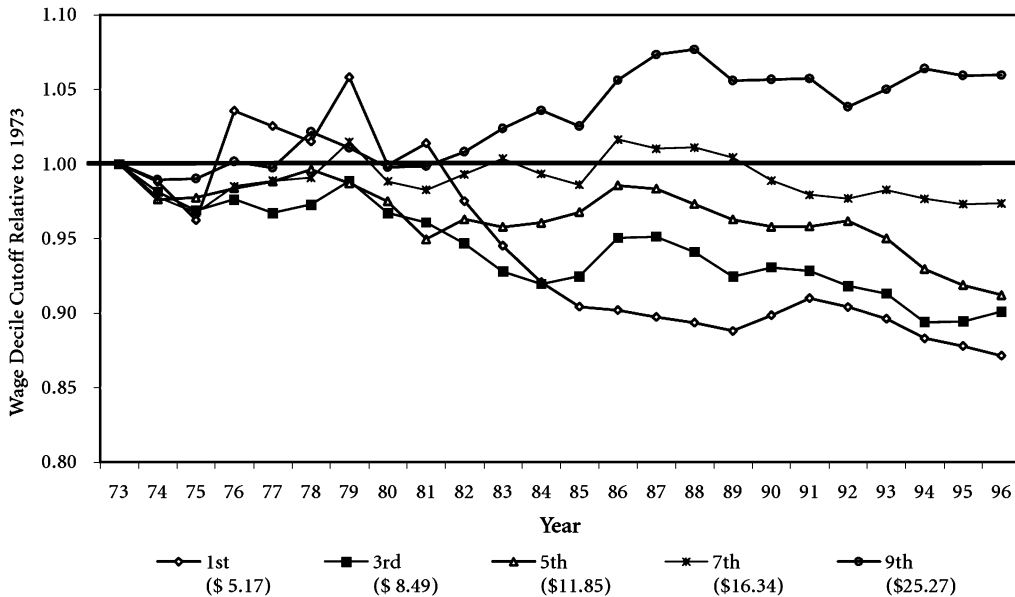


Figure 1. The lines trace the real value of wages at each decile relative to the value in 1973. Sample is all workers, from the CPS Outgoing Rotation Group File. The data can be obtained online at http://www.epinet.org/datazone/data/orghourlyxoffs_all.xls. For more analysis of these data see Bernstein and Mishel (1997).

1970s. Median earnings stabilized in 1973 and then began to decline. In 1980, earnings inequality began to rise rapidly. The net result of these two trends was that by the early 1990s, nearly 80 percent of workers earned less than their counterparts in the 1960s.² Although there appeared to be a small uptick in median wages in the last years of the 1990s, it will take decades to make up the losses.

Data from the Current Population Survey (CPS) show these trends in terms of hourly wages (Figure 1).³ Each line in the figure represents the dollar value of the wage decile cutoff indexed to its value in 1973. The line for decile 1, for example, represents the relative wage earned by the worker at the tenth percentile of the wage distribution over time. If we follow it across the graph, we can see that workers in this decile saw their real wages rise slightly from the mid-1970s to the beginning of the 1980s, fall precipitously during the 1980s, and continue to decline at a lower rate during the 1990s. By 1996,

wages for these workers had fallen about 13 percent in real terms (and by 20 percent if the sample is restricted to men). Wages at the median (fifth decile) stagnated through most of the 1970s and 1980s, then fell sharply in the 1990s, losing a total of about 10 percent over the two decades. Only those at the top (ninth decile) experienced some wage gains, a rise of about 5 percent, with most of the increase coming in the late 1980s.

Comparative data suggest that among industrialized countries, only the United Kingdom experienced a similar growth in wage inequality. In France, Germany, and Italy, the earnings ratio of the first to the fifth decile among men declined or remained stable. And while British wage inequality grew, it was driven by gains at the top, not losses at the bottom.⁴ The broad stagnation of living standards coupled with a collapse of wages at the bottom is thus a peculiarly American disease, at least among the industrialized democracies.

So the American story since the 1970s is not that the rich got richer and the poor got

poorer, but that virtually everyone lost ground, and those at the bottom lost the most. Although mean trends differ by race and sex, all groups experienced a rise in within-group inequality. This leads to a set of research questions that breaks out of the usual “wage-gap” framework and begins to address the general issue of changes in labor market dynamics. We review below supply-side, demand-side, and institutional accounts of change.

Supply-Side Accounts

On the supply side, the labor force expanded by more than 44 million workers from 1950 to 1980, a 70 percent increase. The baby boom, rising women’s labor force participation, and growing immigration all increased the share of workers with little work experience, tenure or education. Although a rising supply of “unskilled” workers seems like a natural explanation for falling wages at the bottom of the distribution, the evidence suggests little role for most of these supply-side changes:

1. The baby boom generation, born between 1946 and 1964, entered the workforce from 1964 to 1982, a period when earnings of low-wage workers held relatively steady. The dramatic decline in wages at the bottom occurred during the entry of the subsequent “baby bust” cohort, which should instead have benefited from the shrinking labor supply.
2. From 1950 to 1994, the fraction of women working for pay increased from 34 percent to 59 percent. Despite this increase in supply, women’s wages rose at every decile of the distribution. Occupational segregation also remained high, indicating the persistence of largely separate labor markets for men and women and suggesting that the impact of women’s entry on men’s wages was probably small.

3. Immigration, and in particular the rapid influx of Asians and Latin Americans following the 1965 Amendments to the Immigration and Nationality Act, is the one supply-side effect that finds some support in the data. Studies find a substantial impact of immigration on the rise in the wage differential between high school dropouts and graduates.

On the boundary between supply and demand explanations lies the debate over the role of education and “skill.” Despite a rising supply of workers with a college education,⁵ their relative wages began to rise in the mid-1980s. This came to be called a “rise” in the college premium, but it was almost entirely driven by the *collapse* in the earnings of high school graduates and dropouts. Initially, the growth in the education wage gap was attributed to a trade-induced decline in domestic demand for high school–educated workers driven by the relocation of manufacturing to less-developed countries. However, conflicting findings led to the rise of an alternative hypothesis: a rising demand for skill driven by “skill-biased technological change.” The proponents of this argument claim that computerization of the workplace and other technological advances increased the demand for high-skill workers well beyond the increase in supply. One study reported that computer use at work raised earnings by as much as 19 percent. Critics countered with an analysis showing that the use of pencils and “sitting while working” had similarly strong effects. Despite the continuing weakness of the empirical evidence, skill-biased technological change remains the dominant explanation among labor economists for the increase in earnings inequality.

Demand-Side Accounts

The structure of the labor market also changed substantially during the postwar

boom years. Restructuring took two forms: continuing decline in manufacturing employment leading to the rise of a “service economy” and an increase in market-mediated employment relations that featured outsourcing, subcontracting, and temporary, contingent, and part-time work contracts. The trend toward growing employment in the service sector and shrinking employment in manufacturing has continued almost linearly since the 1940s, and by the mid-1990s the service sector accounted for more than 80 percent of total employment. Service sector jobs have traditionally paid less, offered fewer benefits, and relied more on part-time employment. Still, there was no sharp increase in deindustrialization in the 1970s or 1980s that would account for the sharp changes in earnings during these decades. In addition, inequality is growing within the sectors, not simply between them, and declines in manufacturing wages were also substantial. So the simpler versions of the deindustrialization thesis have been discarded.

The “good jobs–bad jobs” debate has therefore been refocused in recent years, noting that high- and low-road employment strategies are being used together, not only within industries, but even within firms. This shifts attention to the trends in the system of employment relations often referred to as the “internal labor market.” Cost reduction became an important basis of competition during the 1970s, and internal labor markets became a prime target because they are inflexible and expensive to maintain. The wave of “downsizing” that took place during the late 1980s and 1990s heralded this change. Still, the size of the shift to market-mediated employment is unclear. One study finds that contingent work rose from about 25 percent of employment in 1980 to 35 percent by the end of the decade. Other estimates have been more conservative, ranging from 5 percent to 17 percent of the labor force by the end of the decade. Although the shift to market-mediated employment may not be as rapid

or revolutionary as some scholars have suggested, the rise in job instability is well documented now, and it has been linked to the declining wages of workers who are not college educated.

Institutional Accounts

Market explanations have dominated research on rising inequality, but supply and demand are clearly mediated by institutional constraints. Relatively fewer studies have examined the role of institutional changes, and those that have focus chiefly on the minimum wage and unions. Both appear to explain part of the decline in wages at the bottom of the distribution. The federal minimum wage was frozen at \$3.35/hour from 1980 to 1990. Although less than 10 percent of the working population typically makes the minimum wage, one estimate attributes about 17 percent of the growth in the gap between the wages of college graduates and high school dropouts to the freeze. The 1990 increase is estimated to have reduced the previous decade’s growth in wage inequality by about 30 percent. The decline in unionization may also have played a role. In 1970, unions represented about 27 percent of all wage and salary earners in the United States; by 1993 only 15 percent. When a job change is accompanied by the loss of union status, the wage penalty is on the order of 20 percent after controlling for other worker characteristics. Some studies suggest that the decline in union density may account for about 20 percent of the overall rise in male wage inequality and as much as 50 percent of the rise for male blue-collar workers. As with other potential explanatory factors, however, there is also evidence of increasing earnings inequality among union members.

An increasingly “globalized” economy provides the context in which all of these trends have emerged. Imports from less developed countries (LDCs) are typically produced by low-skill labor, and LDC imports increased

by 75 percent between 1978 and 1990. Low-wage imports may be an important cause of declining demand for low-skill U.S. workers, but here again the empirical evidence is mixed. Capital mobility may also raise inequality by diverting investment to countries with low-wage labor, thus reducing labor demand in higher-wage countries. In contrast to the theory, however, spending on new plants and equipment overseas by U.S. multinationals has been a declining proportion of total U.S. economic activity.

The Future of Inequality

What does the future hold for earnings inequality in America? That will depend in part on the underlying causal factors and in part on the public response to the current trends. As the review above suggests, the trends are at least now well documented, but there is little consensus on the causal forces. What is clear is that the simplistic Kuznets curve cannot adequately represent the trends in contemporary inequality, if, indeed, it ever was adequate. Local earnings dynamics increasingly reflect the intersection of national political forces and the growing global reach of capital. At the most basic level, it is a question of whether markets or politics will have the upper hand in earnings determination. If market forces dominate, we would expect that the lower and middle classes will continue to lose the benefits of unionization, internal labor markets, and minimum wage laws, thus producing an increasingly bipolar society as the inequality-reducing institutional reforms of New Deal capitalism erode. Under this scenario, the United States is not becoming like Western Europe, which is likely to bear the continuing imprint of its more deeply institutionalized welfare capitalism, but rather like South America and Eastern Europe, with their (sometimes militarily enforced) *laissez-faire* systems.

The relocation of manufacturing jobs to less-industrialized and lower-wage countries

has led some to predict that the United States and other highly industrialized countries will evolve into homogeneously high-skill (and, presumably, low-inequality) labor markets.

However, this view could only be seen from an ivory tower. Over 80 percent of the U.S. labor force is already employed in the service sector. If this sector holds the clues to our future, then wage stagnation and polarization will clearly continue. Many of the lowest paid, least challenging jobs—cashiers, sales persons, truck drivers, receptionists, and information clerks—are in the service sector. They will never be exported. These particular jobs are also in the top 10 in terms of projected employment growth, led, at number one, by cashiers.

Cross-national variation in earnings inequality, however, makes it clear that distributional trends are not driven by a single logic of, say, capitalist development or post-industrialism. Other countries face similar pressures from global markets and structural change, but have maintained a level of equality that is unknown in the United States. The difference is in politics, not markets. It is possible, therefore, to use political institutions to reverse the trends in wage decline and polarization. To the extent that rising inequality in U.S. earnings can be traced to institutional change—and these are the explanations with the most consistent empirical support—inequality is likely to persist. Institutional change tends to be slow, and trends in institutional change are reversed even more slowly. The prospects for a reversal in declining union organization, for example, are dim. Job growth in the nonunion sector of the economy is so rapid that the maintenance of current levels of private sector unionization would require organizing efforts unprecedented in the post-war period. Despite the massive redistribution of income and wealth to those at the top, there seems to be little political support for institutional changes that would reverse these trends.

Why is it, then, that the decline and polarization of earnings has occasioned so little political response among the population at large? There are many possible reasons. One is that most people now rely on two incomes to support a family, so that individual earnings losses are masked by income pooling. Another is the explosion of consumer credit, which makes it possible for many to buy what they can no longer afford. Finally, political apathy may be one of the by-products of a high level of inequality. As those at the bottom of the income distribution feel increasingly remote from the mainstream, alienation from politics grows, election turnout declines, and the public sphere shrinks to include just the affluent. The constituency for redistribution vanishes. In this scenario, graphically detailed by Massey in his provocative 1996 presidential address to the Population Association of America, the persistence of high inequality seems especially likely. The analysis may seem bleak, but it is supported by research that shows that political participation, political information, membership in voluntary associations, and a host of other indicators of political efficacy depend closely on income. The rise in inequality could thus become self-sustaining.

Sociology clearly has an important role to play in understanding the recent trends in economic inequality. It is not just a question of earnings, but of fundamental changes in politics, markets, and life chances. The current earnings trends challenge us to rethink our disciplinary perspective and reintegrate our theoretical and empirical agenda, or else lose the heart of our field to other disciplines.

NOTES

1. Here and throughout, detailed keys to the literature are provided by Morris and Western (1999). Other summaries of the literature include Levy and Murnane (1992) and Danziger and Gottschalk (1993; 1995). Comparative trends in earnings inequality are described in OECD (1996, ch. 6).

2. The precise figure depends on the measure used to adjust for inflation, and this is a hotly contested issue.

3. There are at least four different measures of economic well-being that can be examined here: hourly wages, annual earnings, household total earnings, and wealth. All show the same basic pattern, with pronounced rises in inequality. We show hourly wages here because they do not confound labor supply components, such as hours worked and income pooling, with labor pricing. Wages thus better represent the job structure.

4. Inequality does appear to be on the rise in post-socialist economies, although the quantitative evidence here is less reliable. The forces making for such change are, at least on the surface, very different than those at work in the United States, except insofar as one understands them as proceeding from “marketization” in its various forms (e.g., de-unionization, deregulation of wages).

5. About 25 percent of the 25–34-year-old population had a four-year college degree in 1995.

REFERENCES

- Bernstein, Jared, and Lawrence Mishel. 1997. “Has Wage Inequality Stopped Growing?” *Monthly Labor Review* 120:3–16.
- Clark, Terry N., and Seymour M. Lipset. 1991. “Are Social Classes Dying?” *International Sociology* 6:397–410.
- Danziger, Sheldon, and Peter Gottschalk, eds. 1993. *Uneven Tides: Rising Inequality in America*. New York.
- Danziger, Sheldon, and Peter Gottschalk. 1995. *America Unequal*. Cambridge: Harvard University Press.
- Kuznets, Simon. 1955. “Economic Growth and Income Inequality (Presidential Address).” *American Economic Review* 45:1–28.
- Levy, Frank, and Robert Murnane. 1992. “U.S. Earnings Levels and Earnings Inequality: A Review of Recent Trends and Proposed Explanations.” *Journal of Economic Literature* 30: 1333–81.
- Massey, Douglas S. 1996. “The Age of Extremes: Concentrated Affluence and Poverty in the Twenty-First Century.” *Demography* 33:395–412.
- Morris, Martina, and Bruce Western. 1999. “Inequality in Earnings at the Close of the Twentieth Century.” *Annual Review of Sociology* 25:623–57.
- Organisation for Economic Cooperation and Development (OECD). 1996. *Employment Outlook*. Paris: OECD.

8 ■ Thomas Piketty and Emmanuel Saez

The Evolution of Top Incomes

A Historical and International Perspective

This chapter summarizes the main findings and perspectives emerging from a collective research project on the dynamics of income and wealth distribution. The primary objective of this project is to construct a high-quality, long-run, international database on income and wealth concentration, using historical tax statistics. The resulting database now includes annual series covering most of the twentieth century for a number of (mostly Western) countries.

The main motivation for this project comes from a general dissatisfaction with existing income inequality databases. Existing international databases display little homogeneity over time or across countries. They cover only a few isolated years per country, generally restricted to the post-1970 or post-1980 period. They almost never offer any decomposition of income inequality into a labor-income and a capital-income component. Economic mechanisms can be very different for the distribution of labor income (demand and supply of skills, labor market institutions, etc.) and the distribution of capital income (capital accumulation, credit constraints, inheritance law and taxation, etc.), so that it is difficult to test any of these mechanisms using existing data. The fact that existing data are not long run is also problematic, because structural changes in income and wealth distribution often span several decades. In order to properly understand such changes, one needs to

be able to put them into broader historical perspective.

Our database also suffers from important limitations. In particular, our long-run series are generally confined to top income and wealth shares and contain little information about bottom segments of the distribution. However, our data are annual, long-run time series; they are fairly homogenous across countries; and they are also broken down by income source. Hence, they offer a unique opportunity to understand the dynamics of income and wealth distribution and the interplay between inequality and growth.

Constructing a New Database

Household income surveys did not exist prior to 1950. The only data source that is consistently available on a long-run basis is tax data. Most Western countries have published annual tax statistics since the beginning of income taxation, about a century ago. Those statistics generally report the corresponding number of taxpayers for a large number of income brackets, as well as their total income and tax liability. The statistics usually provide the breakdown by income source: capital income, wage income, business income, etc. Using population census data, one can estimate the total number of tax units, had everybody been required to file a tax return, and determine the number of returns in top fractile groups such as the top

10 percent, 1 percent, etc. One can then use standard Pareto interpolation techniques to compute thresholds and average incomes corresponding to each of those top groups. Top income shares series are obtained by dividing top income series by total personal income. This denominator can be computed using aggregate income sources (national accounts and their ancestors). The denominator is adjusted to represent, as closely as possible, total personal income reported on tax returns had everybody been required to file. There are two important limitations with such data.

First, prior to World War II, the proportion of individuals subject to progressive income taxation in most countries rarely exceeded 10 percent, so that one can compute only top-decile income share series (and above) for the entire period.

Second, the data are based on income reported for tax purposes, so they might not reflect real income because of tax evasion (fraudulent underreporting or nonreporting) or tax avoidance (using legal means to repackage reported incomes in order to reduce tax liability). The extent of tax evasion or tax avoidance is possibly related to the level of taxes, the enforcement of the tax law, and the legal tax environment. Therefore, when using tax data to study top incomes, it is necessary to analyze the tax structure at the same time in order to tell real changes in income concentration from changes due to tax avoidance or evasion following a tax reform. This aspect is likely to be even more important in the case of developing countries where tax enforcement is weaker and tax avoidance and evasion are likely to be key determinants of reported top incomes.

Top Income Shares Results

Figure 1, panel A, presents the income share of the top decile in the United States from 1917 to 2002. The overall pattern of the top decile share over the century is U-shaped. The share

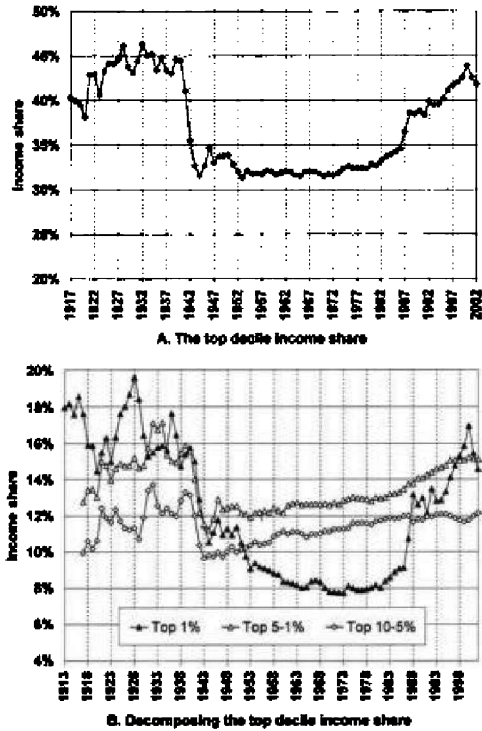


Figure 1. Top Income Shares in the United States, 1913–2002

Note: Income is defined as market income, excluding government transfers and realized capital gains.

Source: Piketty and Saez (2003) and Atkinson and Piketty (2006).

of the top decile fluctuates around 40 to 45 percent during the interwar period. It declines substantially to just above 30 percent during World War II and stays flat at 31 to 32 percent until the 1970s. Such an abrupt decline cannot easily be reconciled with a Kuznets-type process. After decades of stability in the post-war period, the top decile share has increased dramatically over the last 25 years and is now at a level close to the prewar level.

Figure 1, panel B, decomposes the top decile into the top percentile (top 1 percent), the next 4 percent (top 5–1 percent), and the bottom half of the top decile (top 10–5 percent). Most fluctuations of the top decile are due to fluctuations within the top percentile. The drop in the next two groups during World War II is far less dramatic, and they re-

cover from the WWII shock relatively quickly. Their shares did not increase much in recent decades. In contrast, the top percentile has gone through enormous fluctuations along the course of the twentieth century, from about 18 percent before WWI, to only about 8 percent during the 1960s and 1970s, and back to almost 17 percent by 2000. The top percentile share declined during WWI, recovered during the 1920s and declined again during the Great Depression and WWII. This very specific timing, together with the fact that very high incomes account for a disproportionate share of the total decline in inequality, strongly suggests the shocks incurred by capital owners from 1914 to 1945 (Depression and wars) have played a key role.¹

Looking at very top incomes and their composition provides additional evidence. Figure 2 displays the share and composition of income from 1916 to 2000 for the top 0.01 percent in the United States. Until the 1970s, very top incomes were composed primarily of capital income (mostly dividend income) and to a smaller extent business income, the wage income share being very modest. Figure 2 confirms that the large decline of top incomes observed during the 1914–1960 period is predominantly a capital income phenomenon.

Figure 2 shows the income composition pattern at the very top has changed considerably between 1960 and 2000. Salary income has been driving up top incomes and has now become the main source of income at the very top. The dramatic evolution of the composition of top incomes seems robust. National accounts data show the share of capital income in aggregate personal income has been stable in the long run. Therefore, the secular decline of top capital incomes is the consequence of a decreased concentration of capital income and not of a decline in the share of capital income in the economy as a whole. Estimates of wealth concentration from estate tax returns for the 1916–2000 period in the

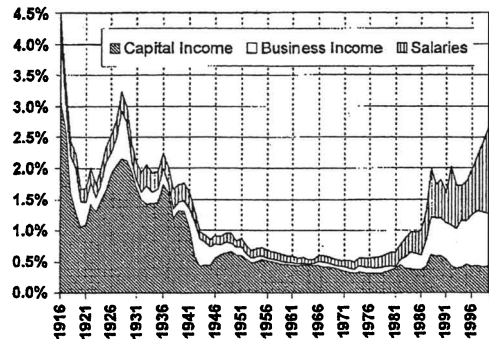


Figure 2. The Top-0.1-Percent Income Share and Composition in the United States, 1916–2000

Note: The figure displays the top-0.1-percent income share (top curve) and its composition (excluding capital gains).

Source: Piketty and Saez (2003).

United States, constructed by Wojciech Kopczuk and Saez (2004) show a precipitous decline in the first part of the century with only very modest increases in recent decades (Figure 4, panel B). This evidence is consistent with the income share series and shows that the dramatic recent increase in income concentration is primarily a labor income phenomenon, and this has not yet translated into a dramatic increase in wealth concentration.

The dramatic decline in top-income shares in the first part the twentieth century took place in almost all countries that have been studied. Figure 3, panel A, displays the top-0.1-percent income share in three English-speaking countries: the United States, United Kingdom, and Canada. Panel B displays the top 0.1 percent for France and Japan. All countries experience a sharp drop in the first part of the century. The timing and size of the decline varies across countries. For example, the decline in Japan is entirely concentrated in the immediate prewar and war years.² As in the United States, income composition series for each of those countries shows that this decrease is primarily a capital income phenomenon due to the fall of top capital incomes.

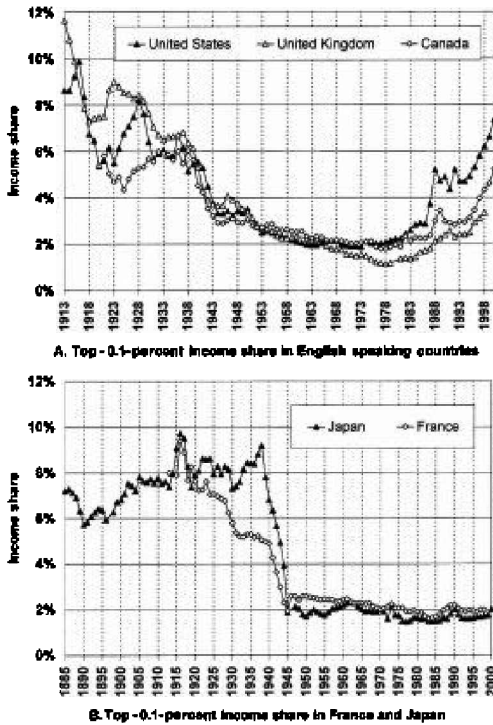


Figure 3. Top-0.1-Percent Income Shares Across Countries

Source: United States, Piketty and Saez (2003); United Kingdom, Atkinson (2006); Canada, Saez and Veall (2005); France, Piketty (2003); Japan, Moriguchi and Saez (2005).

Figure 3 shows a sharp contrast between English-speaking countries and others in recent decades. The United States, United Kingdom, and Canada display a substantial increase in the top-0.1-percent income share over the last 25 years. This increase is largest in the United States, but the timing is remarkably similar across the three countries. In contrast, France and Japan do not experience any noticeable increase in the top-0.1-percent income share. As a result, income concentration is much lower in those countries than in the English-speaking countries.

Explaining the Results and Future Research

The fact that the drop in income concentration in the first part of the twentieth century

is primarily due to the fall in top capital incomes, and that the fall took place mostly during wartime and the Great Depression in most of those countries, suggests an obvious explanation. For the most part, income inequality dropped because capital owners incurred severe shocks to their capital holdings during the 1914 to 1945 period such as destruction, inflation, bankruptcies, and fiscal shocks for financing the wars. This interpretation is confirmed by available wealth and estate data for countries such as France, the United States, and Japan. Note that the idea that capital owners incurred large shocks during the 1914–1945 period and that this had a big impact on income distribution is certainly not new (Kuznets, 1953). What is new is there is not much else going on.

The more challenging part needing explanation is the nonrecovery of top capital incomes during the post-1945 period. The proposed explanation is that the 1914 to 1945 capital shocks had a permanent impact because the introduction of progressive income and estate taxation (there was virtually no tax progressivity prior to 1914 and top rates increased enormously between 1914 and 1945) made it impossible for top capital holders to recover fully. Simple simulations suggest the long-run impact of tax progressivity on wealth concentration is large enough to explain the magnitude of the observed changes (Piketty, 2003).

Those explanations about the dynamics of capital income concentration could possibly be tested by looking at the case of countries that either did not experience large pre-1945 shocks and/or did not implement significant and sustained progressive income tax systems. Switzerland stayed out of the wars and never implemented very progressive wealth or income taxation. As displayed in Figure 4, in contrast to other countries such as the United States, top wealth shares in Switzerland hardly declined from 1913 to the 1960s.

It would be interesting to consider (a) countries that avoided the war or Depression

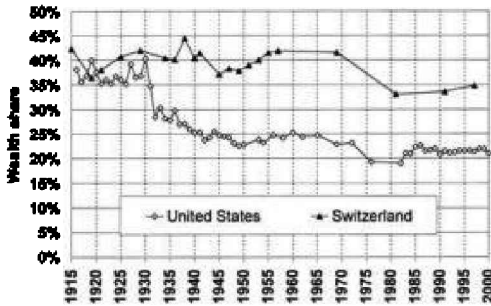


Figure 4. Top-1-Percent Wealth Share in Switzerland and the United States

(Source: *Source*: United States, Kopczuk and Saez (2004); Switzerland, Atkinson and Piketty (2006).)

shocks but developed progressive taxation, such as Ireland and Sweden; and (b) countries that experienced the shocks but did not develop progressive taxation. No European country falls clearly into this latter category.³ Atkinson and Piketty (2006) show that there was no significant drop in top income shares in Ireland during World War II, and top income shares were quite similar in the early 1920s and the late 1940s. Top income shares did fall significantly in the postwar decades, however, when Ireland implemented progressive taxation with very high top rates. Those results suggest the large war shocks may not be necessary to drive down top income shares, and the change in the tax structure might be the most important determinant of long-run income concentration. In future work, it should be possible to develop precise series of tax burdens by income fractiles and income sources for each country and use regression analysis to provide more convincing tests of those explanations. Other factors, such as fertility or the norms about estate division between heirs, could also have significant impacts on long-run wealth distribution.

During the post-1970 period, one observes a major divergence between rich countries. While top income shares have remained fairly stable in continental European countries and Japan over the past three decades,

they have increased enormously in the United States and other English-speaking countries. This rise of top income shares is due not to the revival of top capital incomes, but rather to the very large increases in top wages (especially top executive compensation). As a consequence, top executives (the “working rich”) replaced top capital owners (the “rentiers”) at the top of the income hierarchy during the twentieth century. Understanding why top wages have surged in English-speaking countries in recent decades, but not in continental Europe or Japan, remains a controversial question, with three broad views. First, the free market view claims that technological progress has made managerial skills more general and less firm-specific, increasing competition for the best executives from segregated within-firm markets to a single economywide market. While this view can account for U.S. trends, it cannot explain why executive pay has not changed in other countries such as Japan and France, which have gone through similar technological changes. A second view claims impediments to free markets due to labor market regulations, unions, or social norms regarding pay inequality can keep executive pay below market. Such impediments have been largely removed in the United States, but still exist in Europe and Japan. Under this view, the surge in executive compensation actually represents valuable efficiency gains. Finally, a third view claims the surge in top compensation in the United States is due to the increased ability of executives to set their own pay and extract rents at the expense of shareholders, perhaps for the same reasons as under the second view. In this case, however, there might not be any associated efficiency gains.

The relationship one might want to test, ultimately using our database, is the impact of inequality on growth. Casual examination of the series constructed suggests income concentration and growth are not systematically related. Many countries (such as France,

the United States, and Japan) grew fastest in the postwar decades when income concentration was at its lowest. Thus, one can safely conclude that the enormous decline in wealth concentration that took place between 1914 and 1945 did not prevent high growth from occurring. It seems that in recent decades, however, growth and increases in inequality have been positively correlated: the United States and the United Kingdom have grown faster than continental Europe and Japan. Although cross-country analysis will always suffer from severe identification problems, our hope is that the database will renew the analysis of the interplay between inequality and growth.

NOTES

1. The negative effect of the wars on top incomes can be explained, in part, by the large tax increases enacted to finance the wars. During both wars, the corporate income tax (as well as the individual income tax) was drastically increased and this reduced, mechanically, the distributions to stockholders.

2. The case of Japan is interesting because series start in 1885, at the very beginning of the industrialization of Japan. The series show that income concentration in Japan was high before industrialization began and did not increase much during the industrialization process from 1885 to 1935.

3. Japan and Germany experienced a dramatic decline in income concentration during WWII. Top income shares did not recover at all in Japan (Chiaki Moriguchi and Saez, 2005) but did so to some extent in Germany (Atkinson and Piketty, 2006). A systematic comparison of tax systems in Japan and Germany in the postwar period could thus be informative.

REFERENCES

- Atkinson, Anthony and Piketty, Thomas. *Top Incomes Over the Twentieth Century*. Oxford: Oxford University Press, 2006.
- Kopczuk, Wojciech and Saez, Emmanuel. "Top Wealth Shares in the United States, 1916–2000: Evidence from Estate Tax Returns." *National Tax Journal*, 2004, 57(2), pp. 445–87.
- Kuznets, Simon. *Shares of Upper Income Groups in Income and Savings*. New York: National Bureau of Economic Research, 1953.
- Moriguchi, Chiaki and Saez, Emmanuel. "The Evolution of Income Concentration in Japan, 1885–2002." Unpublished Paper, 2005.
- Piketty, Thomas. "Income Inequality in France, 1901–1998." *Journal of Political Economy*, 2003, 111(5), pp. 1004–42.
- Piketty, Thomas and Saez, Emmanuel. "Income Inequality in the United States, 1913–1998." *Quarterly Journal of Economics*, 2003, 118(1), pp. 1–39.
- Saez, Emmanuel and Veall, Michael R. "The Evolution of High Incomes in Northern America: Lessons from Canadian Evidence." *American Economic Review*, 2005, 95(3), pp. 831–49.

PART III

The Structure of Contemporary Inequality

Theories of Class

■ MARX AND POST-MARXISTS ■

9 ■ Karl Marx

Alienation and Social Classes

We shall begin from a *contemporary* economic fact. The worker becomes poorer the more wealth he produces and the more his production increases in power and extent. The worker becomes an ever cheaper commodity the more goods he creates. The *devaluation* of the human world increases in direct relation with the *increase in value* of the world of things. Labour does not only create goods; it also produces itself and the worker as a *commodity*, and indeed in the same proportion as it produces goods.

This fact simply implies that the object produced by labour, its product, now stands opposed to it as an *alien being*, as a *power independent* of the producer. The product of labour is labour which has been embodied in an object and turned into a physical thing; this product is an *objectification* of labour. The performance of work is at the same time its objectification. The performance of work appears in the sphere of political economy as a *vitiation* of the worker, objectification as a *loss* and as *servitude to the object*, and appropriation as *alienation*.

So much does the performance of work appear as vitiation that the worker is vitiated to

the point of starvation. So much does objectification appear as loss of the object that the worker is deprived of the most essential things not only of life but also of work. Labour itself becomes an object which he can acquire only by the greatest effort and with unpredictable interruptions. So much does the appropriation of the object appear as alienation that the more objects the worker produces the fewer he can possess and the more he falls under the domination of his product, of capital.

All these consequences follow from the fact that the worker is related to the *product of his labour* as to an *alien* object. For it is clear on this presupposition that the more the worker expends himself in work the more powerful becomes the world of objects which he creates in face of himself, the poorer he becomes in his inner life, and the less he belongs to himself. It is just the same as in religion. The more of himself man attributes to God the less he has left in himself. The worker puts his life into the object, and his life then belongs no longer to himself but to the object. The greater his activity, therefore, the less he possesses. What is embodied in the product of his labour is no longer his own. The greater this

Karl Marx. "The Economic and Philosophical Manuscripts," in *Karl Marx: Early Writings*, edited and translated by T. B. Bottomore, pp. 121–131. Copyright © 1963 by McGraw-Hill. Reprinted by permission of McGraw-Hill. "The Holy Family: A Critique of Critical Criticism," in *The Marx-Engels Reader*, edited and translated by Robert C. Tucker, pp. 133–135. Copyright © 1978, 1972 by W. W. Norton & Company, Inc. Used by permission of W. W. Norton & Company.

product is, therefore, the more he is diminished. The *alienation* of the worker in his product means not only that his labour becomes an object, assumes an *external* existence, but that it exists independently, *outside himself*, and alien to him, and that it stands opposed to him as an autonomous power. The life which he has given to the object sets itself against him as an alien and hostile force. . . .

So far we have considered the alienation of the worker only from one aspect; namely, *his relationship with the products of his labour*. However, alienation appears not merely in the result but also in the *process of production*, within *productive activity* itself. How could the worker stand in an alien relationship to the product of his activity if he did not alienate himself in the act of production itself? The product is indeed only the *résumé* of activity, of production. Consequently, if the product of labour is alienation, production itself must be active alienation—the alienation of activity and the activity of alienation. The alienation of the object of labour merely summarizes the alienation in the work activity itself.

What constitutes the alienation of labour? First, that the work is *external* to the worker, that it is not part of his nature; and that, consequently, he does not fulfil himself in his work but denies himself, has a feeling of misery rather than well-being, does not develop freely his mental and physical energies but is physically exhausted and mentally debased. The worker, therefore, feels himself at home only during his leisure time, whereas at work he feels homeless. His work is not voluntary but imposed, *forced labour*. It is not the satisfaction of a need, but only a *means* for satisfying other needs. Its alien character is clearly shown by the fact that as soon as there is no physical or other compulsion it is avoided like the plague. External labour, labour in which man alienates himself, is a labour of self-sacrifice, of mortification. Finally, the external character of work for the worker is shown by the fact that it is not his own work but work for someone else, that in

work he does not belong to himself but to another person.

Just as in religion the spontaneous activity of human fantasy, of the human brain and heart, reacts independently as an alien activity of gods or devils upon the individual, so the activity of the worker is not his own spontaneous activity. It is another's activity and a loss of his own spontaneity.

We arrive at the result that man (the worker) feels himself to be freely active only in his animal functions—eating, drinking and procreating, or at most also in his dwelling and in personal adornment—while in his human functions he is reduced to an animal. The animal becomes human and the human becomes animal.

Eating, drinking and procreating are of course also genuine human functions. But abstractly considered, apart from the environment of human activities, and turned into final and sole ends, they are animal functions.

We have now considered the act of alienation of practical human activity, labour, from two aspects: (1) the relationship of the worker to the *product of labour* as an alien object which dominates him. This relationship is at the same time the relationship to the sensuous external world, to natural objects, as an alien and hostile world; (2) the relationship of labour to the *act of production* within *labour*. This is the relationship of the worker to his own activity as something alien and not belonging to him, activity as suffering (passivity), strength as powerlessness, creation as emasculation, the *personal* physical and mental energy of the worker, his personal life (for what is life but activity?), as an activity which is directed against himself, independent of him and not belonging to him. This is *self-alienation* as against the above-mentioned alienation of the *thing*.

We have now to infer a third characteristic of *alienated labour* from the two we have considered.

Man is a species-being not only in the sense that he makes the community (his own as

well as those of other things) his object both practically and theoretically, but also (and this is simply another expression for the same thing) in the sense that he treats himself as the present, living species, as a *universal* and consequently free being.

Species-life, for man as for animals, has its physical basis in the fact that man (like animals) lives from inorganic nature, and since man is more universal than an animal so the range of inorganic nature from which he lives is more universal. Plants, animals, minerals, air, light, etc. constitute, from the theoretical aspect, a part of human consciousness as objects of natural science and art; they are man's spiritual inorganic nature, his intellectual means of life, which he must first prepare for enjoyment and perpetuation. So also, from the practical aspect, they form a part of human life and activity. In practice man lives only from these natural products, whether in the form of food, heating, clothing, housing, etc. The universality of man appears in practice in the universality which makes the whole of nature into his inorganic body: (1) as a direct means of life; and equally (2) as the material object and instrument of his life activity. Nature is the inorganic body of man; that is to say nature, excluding the human body itself. To say that man *lives* from nature means that nature is his *body* with which he must remain in a continuous interchange in order not to die. The statement that the physical and mental life of man, and nature, are interdependent means simply that nature is interdependent with itself, for man is a part of nature.

Since alienated labour: (1) alienates nature from man; and (2) alienates man from himself, from his own active function, his life activity; so it alienates him from the species. It makes *species-life* into a means of individual life. In the first place it alienates species-life and individual life, and secondly, it turns the latter, as an abstraction, into the purpose of the former, also in its abstract and alienated form.

For labour, *life activity, productive life*, now appear to man only as *means* for the satisfaction of a need, the need to maintain his physical existence. Productive life is, however, species-life. It is life creating life. In the type of life activity resides the whole character of a species, its species-character; and free, conscious activity is the species-character of human beings. Life itself appears only as a *means of life*.

The animal is one with its life activity. It does not distinguish the activity from itself. It is *its activity*. But man makes his life activity itself an object of his will and consciousness. He has a conscious life activity. It is not a determination with which he is completely identified. Conscious life activity distinguishes man from the life activity of animals. Only for this reason is he a species-being. Or rather, he is only a self-conscious being, i.e. his own life is an object for him, because he is a species-being. Only for this reason is his activity free activity. Alienated labour reverses the relationship, in that man because he is a self-conscious being makes his life activity, his *being*, only a means for his *existence*.

The practical construction of an *objective world*, the *manipulation* of inorganic nature, is the confirmation of man as a conscious species-being, i.e. a being who treats the species as his own being or himself as a species-being. Of course, animals also produce. They construct nests, dwellings, as in the case of bees, beavers, ants, etc. But they only produce what is strictly necessary for themselves or their young. They produce only in a single direction, while man produces universally. They produce only under the compulsion of direct physical needs, while man produces when he is free from physical need and only truly produces in freedom from such need. Animals produce only themselves, while man reproduces the whole of nature. The products of animal production belong directly to their physical bodies, while man is free in face of his product. Animals construct only in accordance with the standards and needs of the

species to which they belong, while man knows how to produce in accordance with the standards of every species and knows how to apply the appropriate standard to the object. Thus man constructs also in accordance with the laws of beauty.

It is just in his work upon the objective world that man really proves himself as a *species-being*. This production is his active species-life. By means of it nature appears as *his* work and his reality. The object of labour is, therefore, the *objectification of man's species-life*; for he no longer reproduces himself merely intellectually, as in consciousness, but actively and in a real sense, and he sees his own reflection in a world which he has constructed. While, therefore, alienated labour takes away the object of production from man, it also takes away his *species-life*, his real objectivity as a species-being, and changes his advantage over animals into a disadvantage in so far as his inorganic body, nature, is taken from him.

Just as alienated labour transforms free and self-directed activity into a means, so it transforms the species-life of man into a means of physical existence.

Consciousness, which man has from his species, is transformed through alienation so that species-life becomes only a means for him. (3) Thus alienated labour turns the *species-life of man*, and also nature as his mental species-property, into an *alien* being and into a *means* for his *individual existence*. It alienates from man his own body, external nature, his mental life and his *human* life. (4) A direct consequence of the alienation of man from the product of his labour, from his life activity and from his species-life, is that *man* is *alienated* from other *men*. When man confronts himself he also confronts *other* men. What is true of man's relationship to his work, to the product of his work and to himself, is also true of his relationship to other men, to their labour and to the objects of their labour.

In general, the statement that man is alienated from his species-life means that each man

is alienated from others, and that each of the others is likewise alienated from human life.

Human alienation, and above all the relation of man to himself, is first realized and expressed in the relationship between each man and other men. Thus in the relationship of alienated labour every man regards other men according to the standards and relationships in which he finds himself placed as a worker.

We began with an economic fact, the alienation of the worker and his production. We have expressed this fact in conceptual terms as *alienated labour*, and in analysing the concept we have merely analysed an economic fact.

Let us now examine further how this concept of alienated labour must express and reveal itself in reality. If the product of labour is alien to me and confronts me as an alien power, to whom does it belong? If my own activity does not belong to me but is an alien, forced activity, to whom does it belong? To a being *other* than myself. And who is this being? The *gods*? It is apparent in the earliest stages of advanced production, e.g. temple building, etc. in Egypt, India, Mexico, and in the service rendered to gods, that the product belonged to the gods. But the gods alone were never the lords of labour. And no more was *nature*. What a contradiction it would be if the more man subjugates nature by his labour, and the more the marvels of the gods are rendered superfluous by the marvels of industry, the more he should abstain from his joy in producing and his enjoyment of the product for love of these powers.

The *alien* being to whom labour and the product of labour belong, to whose service labour is devoted, and to whose enjoyment the product of labour goes, can only be *man* himself. If the product of labour does not belong to the worker, but confronts him as an alien power, this can only be because it belongs to *a man other than the worker*. If his activity is a torment to him it must be a

source of *enjoyment* and pleasure to another. Not the gods, nor nature, but only man himself can be this alien power over men.

Consider the earlier statement that the relation of man to himself is first *realized, objectified*, through his relation to other men. If he is related to the product of his labour, his objectified labour, as to an *alien*, hostile, powerful and independent object, he is related in such a way that another alien, hostile, powerful and independent man is the lord of this object. If he is related to his own activity as to unfree activity, then he is related to it as activity in the service, and under the domination, coercion and yoke, of another man. . . .

Thus, through alienated labour the worker creates the relation of another man, who does not work and is outside the work process, to this labour. The relation of the worker to work also produces the relation of the capitalist (or whatever one likes to call the lord of labour) to work. *Private property* is, therefore, the product, the necessary result, of *alienated labour*, of the external relation of the worker to nature and to himself.

Private property is thus derived from the analysis of the concept of *alienated labour*; that is, alienated man, alienated labour, alienated life, and estranged man.

We have, of course, derived the concept of *alienated labour (alienated life)* from political economy, from an analysis of the *movement of private property*. But the analysis of this concept shows that although private property appears to be the basis and cause of alienated labour, it is rather a consequence of the latter, just as the gods are *fundamentally*

not the cause but the product of confusions of human reason. At a later stage, however, there is a reciprocal influence.

Only in the final stage of the development of private property is its secret revealed, namely, that it is on one hand the *product* of alienated labour, and on the other hand the *means* by which labour is alienated, *the realization of this alienation*.

The Economic and Philosophical Manuscripts,
pp. 121–131

The possessing class and the proletarian class represent one and the same human self-alienation. But the former feels satisfied and affirmed in this self-alienation, experiences the alienation as a sign *of its own power*, and possesses in it the *appearance* of a human existence. The latter, however, feels destroyed in this alienation, seeing in it its own impotence and the reality of an inhuman existence. To use Hegel's expression, this class is, within depravity, an *indignation* against this depravity, an indignation necessarily aroused in this class by the contradiction between its human *nature* and its life-situation, which is a blatant, outright and all-embracing denial of that very nature.

Within the antagonism as a whole, therefore, private property represents the *conservative* side and the proletariat the *destructive* side. From the former comes action aimed at preserving the antagonism; from the latter, action aimed at its destruction.

The Holy Family: A Critique of Critical Criticism,
pp. 133–134

■ Karl Marx

Classes in Capitalism and Pre-Capitalism

The history of all hitherto existing society¹ is the history of class struggles.

Freeman and slave, patrician and plebeian, lord and serf, guild-master² and journeyman, in a word, oppressor and oppressed, stood in constant opposition to one another, carried on an uninterrupted, now hidden, now open fight, a fight that each time ended, either in a revolutionary re-constitution of society at large, or in the common ruin of the contending classes.

In the earlier epochs of history, we find almost everywhere a complicated arrangement of society into various orders, a manifold gradation of social rank. In ancient Rome we have patricians, knights, plebeians, slaves; in the Middle Ages, feudal lords, vassals, guild-masters, journeymen, apprentices, serfs; in almost all of these classes, again, subordinate gradations.

The modern bourgeois society that has sprouted from the ruins of feudal society has not done away with class antagonisms. It has but established new classes, new conditions of oppression, new forms of struggle in place of the old ones.

Our epoch, the epoch of the bourgeoisie, possesses, however, this distinctive feature: it has simplified the class antagonisms. Society as a whole is more and more splitting up into two great hostile camps, into two great classes directly facing each other: Bourgeoisie and Proletariat.

From the serfs of the Middle Ages sprang the chartered burghers of the earliest towns. From these burgesses the first elements of the bourgeoisie were developed.

The discovery of America, the rounding of the Cape, opened up fresh ground for the rising bourgeoisie. The East-Indian and Chinese markets, the colonisation of America, trade with the colonies, the increase in the means of exchange and in commodities generally, gave to commerce, to navigation, to industry, an impulse never before known, and thereby, to the revolutionary element in the tottering feudal society, a rapid development.

The feudal system of industry, under which industrial production was monopolised by closed guilds, now no longer sufficed for the growing wants of the new markets. The manufacturing system took its place. The guild-masters were pushed on one side by the manufacturing middle class; division of labour between the different corporate guilds vanished in the face of division of labour in each single workshop.

Meantime the markets kept ever growing, the demand ever rising. Even manufacture no longer sufficed. Thereupon, steam and machinery revolutionised industrial production. The place of manufacture was taken by the giant, Modern Industry, the place of the industrial middle class, by industrial millionaires, the leaders of whole industrial armies, the modern bourgeois.

Karl Marx. "The Communist Manifesto," in *Selected Works*, Vol. I (Moscow: Progress Publishers, 1964), pp. 108–119. Reprinted by permission of Progress Publishers. *The Poverty of Philosophy* (New York: International Publishers, 1963), pp. 172–175. Reprinted by permission of International Publishers. "The Eighteenth Brumaire of Louis Bonaparte," in *Selected Works*, Vol. I (Moscow: Progress Publishers, 1963), pp. 478–479. Reprinted by permission of Progress Publishers. *Capital*, Vol. III (Moscow: Progress Publishers, 1967), pp. 885–886. Reprinted by permission of Progress Publishers.

Modern industry has established the world-market, for which the discovery of America paved the way. This market has given an immense development to commerce, to navigation, to communication by land. This development has, in its turn, reacted on the extension of industry; and in proportion as industry, commerce, navigation, railways extended, in the same proportion the bourgeoisie developed, increased its capital, and pushed into the background every class handed down from the Middle Ages.

We see, therefore, how the modern bourgeoisie is itself the product of a long course of development, of a series of revolutions in the modes of production and of exchange.

Each step in the development of the bourgeoisie was accompanied by a corresponding political advance of that class. An oppressed class under the sway of the feudal nobility, an armed and self-governing association in the mediaeval commune³; here independent urban republic (as in Italy and Germany), there taxable “third estate” of the monarchy (as in France), afterwards, in the period of manufacture proper, serving either the semi-feudal or the absolute monarchy as a counterpoise against the nobility, and, in fact, cornerstone of the great monarchies in general, the bourgeoisie has at last, since the establishment of Modern Industry and of the world-market, conquered for itself, in the modern representative State, exclusive political sway. The executive of the modern State is but a committee for managing the common affairs of the whole bourgeoisie.

The bourgeoisie, historically, has played a most revolutionary part. The bourgeoisie, wherever it has got the upper hand, has put an end to all feudal, patriarchal, idyllic relations. It has pitilessly torn asunder the motley feudal ties that bound man to his “natural superiors,” and has left remaining no other nexus between man and man than naked self-interest, than callous “cash payment.” It has drowned the most heavenly ecstasies of religious fervour, of chivalrous enthusiasm, of hilistine sentiment-

alism, in the icy water of egotistical calculation. It has resolved personal worth into exchange value, and in place of the numberless indefeasible chartered freedoms, has set up that single, unconscionable freedom—Free Trade. In one word, for exploitation, veiled by religious and political illusions, it has substituted naked, shameless, direct, brutal exploitation.

The bourgeoisie has stripped of its halo every occupation hitherto honoured and looked up to with reverent awe. It has converted the physician, the lawyer, the priest, the poet, the man of science, into its paid wage-labourers.

The bourgeoisie has torn away from the family its sentimental veil, and has reduced the family relation to a mere money relation.

The bourgeoisie has disclosed how it came to pass that the brutal display of vigour in the Middle Ages, which Reactionists so much admire, found its fitting complement in the most slothful indolence. It has been the first to show what man’s activity can bring about. It has accomplished wonders far surpassing Egyptian pyramids, Roman aqueducts, and Gothic cathedrals; it has conducted expeditions that put in the shade all former Exoduses of nations and crusades.

The bourgeoisie cannot exist without constantly revolutionising the instruments of production, and thereby the relations of production, and with them the whole relations of society. Conservation of the old modes of production in unaltered form, was, on the contrary, the first condition of existence for all earlier industrial classes. Constant revolutionising of production, uninterrupted disturbance of all social conditions, everlasting uncertainty and agitation distinguish the bourgeois epoch from all earlier ones. All fixed, fast-frozen relations, with their train of ancient and venerable prejudices and opinions, are swept away, all new-formed ones become antiquated before they can ossify. All that is solid melts into air, all that is holy is profaned, and man is at last compelled to face with sober senses, his real conditions of life, and his relations with his kind.

The need of a constantly expanding market for its products chases the bourgeoisie over the whole surface of the globe. It must nestle everywhere, settle everywhere, establish connexions everywhere.

The bourgeoisie has through its exploitation of the world-market given a cosmopolitan character to production and consumption in every country. To the great chagrin of Reactionists, it has drawn from under the feet of industry the national ground on which it stood. All old-established national industries have been destroyed or are daily being destroyed. They are dislodged by new industries, whose introduction becomes a life and death question for all civilised nations, by industries that no longer work up indigenous raw material, but raw material drawn from the remotest zones; industries whose products are consumed, not only at home, but in every quarter of the globe. In place of the old wants, satisfied by the productions of the country, we find new wants, requiring for their satisfaction the products of distant lands and climes. In place of the old local and national seclusion and self-sufficiency, we have intercourse in every direction, universal inter-dependence of nations. And as in material, so also in intellectual production. The intellectual creations of individual nations become common property. National one-sidedness and narrow-mindedness become more and more impossible, and from the numerous national and local literatures, there arises a world literature.

The bourgeoisie, by the rapid improvement of all instruments of production, by the immensely facilitated means of communication, draws all, even the most barbarian, nations into civilisation. The cheap prices of its commodities are the heavy artillery with which it batters down all Chinese walls, with which it forces the barbarians' intensely obstinate hatred of foreigners to capitulate. It compels all nations, on pain of extinction, to adopt the bourgeois mode of production; it compels them to introduce what it calls civilisation into their midst, *i.e.*, to become bour-

geois themselves. In one word, it creates a world after its own image.

The bourgeoisie has subjected the country to the rule of the towns. It has created enormous cities, has greatly increased the urban population as compared with the rural, and has thus rescued a considerable part of the population from the idiocy of rural life. Just as it has made the country dependent on the towns, so it has made barbarian and semi-barbarian countries dependent on the civilised ones, nations of peasants on nations of bourgeois, the East on the West.

The bourgeoisie keeps more and more doing away with the scattered state of the population, of the means of production, and of property. It has agglomerated population, centralised means of production, and has concentrated property in a few hands. The necessary consequence of this was political centralisation. Independent, or but loosely connected provinces, with separate interests, laws, governments and systems of taxation, became lumped together into one nation, with one government, one code of laws, one national class-interest, one frontier and one customs-tariff.

The bourgeoisie, during its rule of scarce one hundred years, has created more massive and more colossal productive forces than have all preceding generations together. Subjection of Nature's forces to man, machinery, application of chemistry to industry and agriculture, steam-navigation, railways, electric telegraphs, clearing of whole continents for cultivation, canalisation of rivers, whole populations conjured out of the ground—what earlier century had even a presentiment that such productive forces slumbered in the lap of social labour?

We see then: the means of production and of exchange, on whose foundation the bourgeoisie built itself up, were generated in feudal society. At a certain stage in the development of these means of production and of exchange, the conditions under which feudal society produced and exchanged, the feudal

organisation of agriculture and manufacturing industry, in one word, the feudal relations of property became no longer compatible with the already developed productive forces; they became so many fetters. They had to be burst asunder; they were burst asunder.

Into their place stepped free competition, accompanied by a social and political constitution adapted to it, and by the economical and political sway of the bourgeois class.

A similar movement is going on before our own eyes. Modern bourgeois society with its relations of production, of exchange and of property, a society that has conjured up such gigantic means of production and of exchange, is like the sorcerer, who is no longer able to control the powers of the nether world whom he has called up by his spells. For many a decade past the history of industry and commerce is but the history of the revolt of modern productive forces against modern conditions of production, against the property relations that are the conditions for the existence of the bourgeoisie and of its rule. It is enough to mention the commercial crises that by their periodical return put on its trial, each time more threateningly, the existence of the entire bourgeois society. In these crises a great part not only of the existing products, but also of the previously created productive forces, are periodically destroyed. In these crises there breaks out an epidemic that, in all earlier epochs, would have seemed an absurdity—the epidemic of over-production. Society suddenly finds itself put back into a state of momentary barbarism; it appears as if a famine, a universal war of devastation had cut off the supply of every means of subsistence; industry and commerce seem to be destroyed; and why? Because there is too much civilisation, too much means of subsistence, too much industry, too much commerce. The productive forces at the disposal of society no longer tend to further the development of the conditions of bourgeois property; on the contrary, they have become too powerful for these conditions, by which they are fettered, and so soon as they overcome

these fetters, they bring disorder into the whole of bourgeois society, endanger the existence of bourgeois property. The conditions of bourgeois society are too narrow to comprise the wealth created by them. And how does the bourgeoisie get over these crises? On the one hand by enforced destruction of a mass of productive forces; on the other, by the conquest of new markets, and by the more thorough exploitation of the old ones. That is to say, by paving the way for more extensive and more destructive crises, and by diminishing the means whereby crises are prevented.

The weapons with which the bourgeoisie felled feudalism to the ground are now turned against the bourgeoisie itself.

But not only has the bourgeoisie forged the weapons that bring death to itself; it has also called into existence the men who are to wield those weapons—the modern working class—the proletarians.

In proportion as the bourgeoisie, *i.e.*, capital, is developed, in the same proportion is the proletariat, the modern working class, developed—a class of labourers, who live only so long as they find work, and who find work only so long as their labour increases capital. These labourers, who must sell themselves piecemeal, are a commodity, like every other article of commerce, and are consequently exposed to all the vicissitudes of competition, to all the fluctuations of the market.

Owing to the extensive use of machinery and to division of labour, the work of the proletarians has lost all individual character, and, consequently, all charm for the workman. He becomes an appendage of the machine, and it is only the most simple, most monotonous, and most easily acquired knack, that is required of him. Hence, the cost of production of a workman is restricted, almost entirely, to the means of subsistence that he requires for his maintenance, and for the propagation of his race. But the price of a commodity, and therefore also of labour, is equal to its cost of production. In proportion, therefore, as the repulsiveness of the work increases, the wage

decreases. Nay more, in proportion as the use of machinery and division of labour increases, in the same proportion the burden of toil also increases, whether by prolongation of the working hours, by increase of the work exacted in a given time or by increased speed of the machinery, etc.

Modern industry has converted the little workshop of the patriarchal master into the great factory of the industrial capitalist. Masses of labourers, crowded into the factory, are organised like soldiers. As privates of the industrial army they are placed under the command of a perfect hierarchy of officers and sergeants. Not only are they slaves of the bourgeois class, and of the bourgeois State; they are daily and hourly enslaved by the machine, by the overlooker, and, above all, by the individual bourgeois manufacturer himself. The more openly this despotism proclaims gain to be its end and aim, the more petty, the more hateful and the more embittering it is.

The less the skill and exertion of strength implied in manual labour, in other words, the more modern industry becomes developed, the more is the labour of men superseded by that of women. Differences of age and sex have no longer any distinctive social validity for the working class. All are instruments of labour, more or less expensive to use, according to their age and sex.

No sooner is the exploitation of the labourer by the manufacturer, so far, at an end, and he receives his wages in cash, than he is set upon by the other portions of the bourgeoisie, the landlord, the shopkeeper, the pawnbroker, etc.

The lower strata of the middle class—the small tradespeople, shopkeepers, and retired tradesmen generally, the handicraftsmen and peasants—all these sink gradually into the proletariat, partly because their diminutive capital does not suffice for the scale on which Modern Industry is carried on, and is swamped in the competition with the large capitalists, partly because their specialised skill is rendered worthless by new methods of

production. Thus the proletariat is recruited from all classes of the population.

The proletariat goes through various stages of development. With its birth begins its struggle with the bourgeoisie. At first the contest is carried on by individual labourers, then by the workpeople of a factory, then by the operatives of one trade, in one locality, against the individual bourgeois who directly exploits them. They direct their attacks not against the bourgeois conditions of production, but against the instruments of production themselves: they destroy imported wares that compete with their labour, they smash to pieces machinery, they set factories ablaze, they seek to restore by force the vanished status of the workman of the Middle Ages.

At this stage the labourers still form an incoherent mass scattered over the whole country, and broken up by their mutual competition. If anywhere they unite to form more compact bodies, this is not yet the consequence of their own active union, but of the union of the bourgeoisie, which class, in order to attain its own political ends, is compelled to set the whole proletariat in motion, and is moreover yet, for a time, able to do so. At this stage, therefore, the proletarians do not fight their enemies, but the enemies of their enemies, the remnants of absolute monarchy, the landowners, the non-industrial bourgeois, the petty bourgeoisie. Thus the whole historical movement is concentrated in the hands of the bourgeoisie; every victory so obtained is a victory for the bourgeoisie.

But with the development of industry the proletariat not only increases in number; it becomes concentrated in greater masses, its strength grows, and it feels that strength more. The various interests and conditions of life within the ranks of the proletariat are more and more equalised, in proportion as machinery obliterates all distinctions of labour, and nearly everywhere reduces wages to the same low level. The growing competition among the bourgeois, and the resulting commercial crises, make the wages of the

workers ever more fluctuating. The unceasing improvement of machinery, ever more rapidly developing, makes their livelihood more and more precarious; the collisions between individual workmen and individual bourgeois take more and more the character of collisions between two classes. Thereupon the workers begin to form combinations (Trades' Unions) against the bourgeois; they club together in order to keep up the rate of wages; they found permanent associations in order to make provision beforehand for these occasional revolts. Here and there the contest breaks out into riots.

Now and then the workers are victorious, but only for a time. The real fruit of their battles lies, not in the immediate result, but in the ever-expanding union of the workers. This union is helped on by the improved means of communication that are created by modern industry and that place the workers of different localities in contact with one another. It was just this contact that was needed to centralise the numerous local struggles, all of the same character, into one national struggle between classes. But every class struggle is a political struggle. And that union, to attain which the burghers of the Middle Ages, with their miserable highways, required centuries, the modern proletarians, thanks to railways, achieve in a few years.

This organisation of the proletarians into a class, and consequently into a political party, is continually being upset again by the competition between the workers themselves. But it ever rises up again, stronger, firmer, mightier. It compels legislative recognition of particular interests of the workers, by taking advantage of the divisions among the bourgeoisie itself. Thus the ten-hours' bill in England was carried.

Altogether collisions between the classes of the old society further, in many ways, the course of development of the proletariat. The bourgeoisie finds itself involved in a constant battle. At first with the aristocracy; later on, with those portions of the bour-

geoisie itself, whose interests have become antagonistic to the progress of industry; at all times, with the bourgeoisie of foreign countries. In all these battles it sees itself compelled to appeal to the proletariat, to ask for its help, and thus, to drag it into the political arena. The bourgeoisie itself, therefore, supplies the proletariat with its own elements of political and general education, in other words, it furnishes the proletariat with weapons for fighting the bourgeoisie.

Further, as we have already seen, entire sections of the ruling classes are, by the advance of industry, precipitated into the proletariat, or are at least threatened in their conditions of existence. These also supply the proletariat with fresh elements of enlightenment and progress.

Finally, in times when the class struggle nears the decisive hour, the process of dissolution going on within the ruling class, in fact within the whole range of old society, assumes such a violent, glaring character, that a small section of the ruling class cuts itself adrift, and joins the revolutionary class, the class that holds the future in its hands. Just as, therefore, at an earlier period, a section of the nobility went over to the bourgeoisie, so now a portion of the bourgeoisie goes over to the proletariat, and in particular, a portion of the bourgeois ideologists, who have raised themselves to the level of comprehending theoretically the historical movement as a whole.

Of all the classes that stand face to face with the bourgeoisie today, the proletariat alone is a really revolutionary class. The other classes decay and finally disappear in the face of Modern Industry; the proletariat is its special and essential product.

The lower middle class, the small manufacturer, the shopkeeper, the artisan, the peasant, all these fight against the bourgeoisie, to save from extinction their existence as fractions of the middle class. They are therefore not revolutionary, but conservative. Nay more, they are reactionary, for they

try to roll back the wheel of history. If by chance they are revolutionary, they are so only in view of their impending transfer into the proletariat; they thus defend not their present, but their future interests, they desert their own standpoint to place themselves at that of the proletariat.

The "dangerous class," the social scum, that passively rotting mass thrown off by the lowest layers of old society, may, here and there, be swept into the movement by a proletarian revolution; its conditions of life, however, prepare it far more for the part of a bribed tool of reactionary intrigue.

In the conditions of the proletariat, those of old society at large are already virtually swamped. The proletarian is without property; his relation to his wife and children has no longer anything in common with the bourgeois family-relations; modern, industrial labour, modern subjection to capital, the same in England as in France, in America as in Germany, has stripped him of every trace of national character. Law, morality, religion, are to him so many bourgeois prejudices, behind which lurk in ambush just as many bourgeois interests.

All the preceding classes that got the upper hand, sought to fortify their already acquired status by subjecting society at large to their conditions of appropriation. The proletarians cannot become masters of the productive forces of society, except by abolishing their own previous mode of appropriation, and thereby also every other previous mode of appropriation. They have nothing of their own to secure and to fortify; their mission is to destroy all previous securities for, and insurances of, individual property.

All previous historical movements were movements of minorities, or in the interests of minorities. The proletarian movement is the self-conscious, independent movement of the immense majority, in the interests of the immense majority. The proletariat, the lowest stratum of our present society, cannot stir, cannot raise itself up, without the whole

superincumbent strata of official society being sprung into the air.

Though not in substance, yet in form, the struggle of the proletariat with the bourgeoisie is at first a national struggle. The proletariat of each country must, of course, first of all settle matters with its own bourgeoisie.

In depicting the most general phases of the development of the proletariat, we traced the more or less veiled civil war, raging within existing society, up to the point where that war breaks out into open revolution, and where the violent overthrow of the bourgeoisie lays the foundation for the sway of the proletariat.

Hitherto, every form of society has been based, as we have already seen, on the antagonism of oppressing and oppressed classes. But in order to oppress a class, certain conditions must be assured to it under which it can, at least, continue its slavish existence. The serf, in the period of serfdom, raised himself to membership in the commune, just as the petty bourgeois, under the yoke of feudal absolutism, managed to develop into a bourgeois. The modern labourer, on the contrary, instead of rising with the progress of industry, sinks deeper and deeper below the conditions of existence of his own class. He becomes a pauper, and pauperism develops more rapidly than population and wealth. And here it becomes evident, that the bourgeoisie is unfit any longer to be the ruling class in society, and to impose its conditions of existence upon society as an overriding law. It is unfit to rule because it is incompetent to assure an existence to its slave within his slavery, because it cannot help letting him sink into such a state, that it has to feed him, instead of being fed by him. Society can no longer live under this bourgeoisie, in other words, its existence is no longer compatible with society.

The essential condition for the existence, and for the sway of the bourgeois class, is the formation and augmentation of capital; the condition for capital is wage-labour.

Wage-labour rests exclusively on competition between the labourers. The advance of industry, whose involuntary promoter is the bourgeoisie, replaces the isolation of the labourers, due to competition, by their revolutionary combination, due to association. The development of Modern Industry, therefore, cuts from under its feet the very foundation on which the bourgeoisie produces and appropriates products. What the bourgeoisie, therefore, produces, above all, is its own grave-diggers. Its fall and the victory of the proletariat are equally inevitable.

NOTES

1. That is, all *written* history. In 1847, the pre-history of society, the social organisation existing previous to recorded history, was all but unknown. [*Note by Engels to the English edition of 1888.*]

2. Guild-master, that is, a full member of a guild, a master within, not a head of a guild. [*Note by Engels to the English edition of 1888.*]

3. "Commune" was the name taken, in France, by the nascent towns even before they had conquered from their feudal lords and masters local self-government and political rights as the "Third Estate". Generally speaking, for the economical development of the bourgeoisie, England is here taken as the typical country; for its political development, France. [*Note by Engels to the English edition of 1888.*]

This was the name given their urban communities by the townsmen of Italy and France, after they had purchased or wrested their initial rights of self-government from their feudal lords. [*Note by Engels to the German edition of 1890.*]

The Communist Manifesto, pp. 108–119

The first attempts of workers to *associate* among themselves always take place in the form of combinations.

Large-scale industry concentrates in one place a crowd of people unknown to one another. Competition divides their interests. But the maintenance of wages, this common interest which they have against their boss, unites them in a common thought of resistance—*combination*. Thus combination always has a

double aim, that of stopping competition among the workers, so that they can carry on general competition with the capitalist. If the first aim of resistance was merely the maintenance of wages, combinations, at first isolated, constitute themselves into groups as the capitalists in their turn unite for the purpose of repression, and in face of always united capital, the maintenance of the association becomes more necessary to them than that of wages. This is so true that English economists are amazed to see the workers sacrifice a good part of their wages in favour of associations, which, in the eyes of these economists, are established solely in favour of wages. In this struggle—a veritable civil war—all the elements necessary for a coming battle unite and develop. Once it has reached this point, association takes on a political character.

Economic conditions had first transformed the mass of the people of the country into workers. The combination of capital has created for this mass a common situation, common interests. This mass is thus already a class as against capital, but not yet for itself. In the struggle, of which we have noted only a few phases, this mass becomes united, and constitutes itself as a class for itself. The interests it defends become class interests. But the struggle of class against class is a political struggle.

In the bourgeoisie we have two phases to distinguish: that in which it constituted itself as a class under the regime of feudalism and absolute monarchy, and that in which, already constituted as a class, it overthrew feudalism and monarchy to make society into a bourgeois society. The first of these phases was the longer and necessitated the greater efforts. This too began by partial combinations against the feudal lords.

Much research has been carried out to trace the different historical phases that the bourgeoisie has passed through, from the commune up to its constitution as a class.

But when it is a question of making a precise study of strikes, combinations and other

forms in which the proletarians carry out before our eyes their organization as a class, some are seized with real fear and others display a *transcendental* disdain.

An oppressed class is the vital condition for every society founded on the antagonism of classes. The emancipation of the oppressed class thus implies necessarily the creation of a new society. For the oppressed class to be able to emancipate itself it is necessary that the productive powers already acquired and the existing social relations should no longer be capable of existing side by side. Of all the instruments of production, the greatest productive power is the revolutionary class itself. The organization of revolutionary elements as a class supposes the existence of all the productive forces which could be engendered in the bosom of the old society.

Does this mean that after the fall of the old society there will be a new class domination culminating in a new political power? No.

The condition for the emancipation of the working class is the abolition of every class, just as the condition for the liberation of the third estate, of the bourgeois order, was the abolition of all estates¹ and all orders.

The working class, in the course of its development, will substitute for the old civil society an association which will exclude classes and their antagonism, and there will be no more political power properly so-called, since political power is precisely the official expression of antagonism in civil society.

Meanwhile the antagonism between the proletariat and the bourgeoisie is a struggle of class against class, a struggle which carried to its highest expression is a total revolution. Indeed, is it at all surprising that a society founded on the opposition of classes should culminate in brutal *contradiction*, the shock of body against body, as its final *dénouement*?

Do not say that social movement excludes political movement. There is never a political movement which is not at the same time social.

It is only in an order of things in which there are no more classes and class antagonisms that *social evolutions* will cease to be *political revolutions*. Till then, on the eve of every general reshuffling of society, the last word of social science will always be:

*“Le combat ou la mort; la lutte sanguinaire ou le néant. C’est ainsi que la question est invinciblement posée.”*²

NOTES

1. Estates here in the historical sense of the estates of feudalism, estates with definite and limited privileges. The revolution of the bourgeoisie abolished the estates and their privileges. Bourgeois society knows only *classes*. It was, therefore, absolutely in contradiction with history to describe the proletariat as the “fourth estate.” [Note by F. Engels to the German edition, 1885.]

2. “Combat or death; bloody struggle or extinction. It is thus that the question is inexorably put.” George Sand, *Jean Ziska*.

The Poverty of Philosophy, pp. 172–175

The small-holding peasants form a vast mass, the members of which live in similar conditions but without entering into manifold relations with one another. Their mode of production isolates them from one another instead of bringing them into mutual intercourse. The isolation is increased by France’s bad means of communication and by the poverty of the peasants. Their field of production, the small holding, admits of no division of labour in its cultivation, no application of science and, therefore, no diversity of development, no variety of talent, no wealth of social relationships. Each individual peasant family is almost self-sufficient; it itself directly produces the major part of its consumption and thus acquires its means of life more through exchange with nature than in intercourse with society. A small holding, a peasant and his family; alongside them another small holding, another peasant and another family. A few score of these make up

a village, and a few score of villages make up a Department. In this way, the great mass of the French nation is formed by simple addition of homologous magnitudes, much as potatoes in a sack form a sack of potatoes. In so far as millions of families live under economic conditions of existence that separate their mode of life, their interests and their culture from those of the other classes, and put them in hostile opposition to the latter, they form a class. In so far as there is merely a local interconnection among these small-holding peasants, and the identity of their interests begets no community, no national bond and no political organisation among them, they do not form a class. They are consequently incapable of enforcing their class interests in their own name, whether through a parliament or through a convention. They cannot represent themselves, they must be represented. Their representative must at the same time appear as their master, as an authority over them, as an unlimited governmental power that protects them against the other classes and sends them rain and sunshine from above. The political influence of the small-holding peasants, therefore, finds its final expression in the executive power subordinating society to itself.

The Eighteenth Brumaire of Louis Bonaparte,
pp. 478–479

The owners merely of labour-power, owners of capital, and landowners, whose respective sources of income are wages, profit and ground-rent, in other words, wage-labourers, capitalists and landowners, constitute then three big classes of modern society based upon the capitalist mode of production.

In England, modern society is indisputably most highly and classically developed in economic structure. Nevertheless, even here the stratification of classes does not appear in its pure form. Middle and intermediate strata

even here obliterate lines of demarcation everywhere (although incomparably less in rural districts than in the cities). However, this is immaterial for our analysis. We have seen that the continual tendency and law of development of the capitalist mode of production is more and more to divorce the means of production from labour, and more and more to concentrate the scattered means of production into large groups, thereby transforming labour into wage-labour and the means of production into capital. And to this tendency, on the other hand, corresponds the independent separation of landed property from capital and labour, or the transformation of all landed property into the form of landed property corresponding to the capitalist mode of production.

The first question to be answered is this: What constitutes a class?—and the reply to this follows naturally from the reply to another question, namely: What makes wage-labourers, capitalists and landlords constitute the three great social classes?

At first glance—the identity of revenues and sources of revenue. There are three great social groups whose members, the individuals forming them, live on wages, profit and ground-rent respectively, on the realisation of their labour-power, their capital, and their landed property.

However, from this standpoint, physicians and officials, e.g., would also constitute two classes, for they belong to two distinct social groups, the members of each of these groups receiving their revenue from one and the same source. The same would also be true of the infinite fragmentation of interest and rank into which the division of social labour splits labourers as well as capitalists and landlords—the latter, e.g., into owners of vineyards, farm owners, owners of forests, mine owners and owners of fisheries.

[Here the manuscript breaks off.]

■ Karl Marx

Ideology and Class

The ideas of the ruling class are in every epoch the ruling ideas, i.e. the class which is the ruling *material* force of society, is at the same time its ruling *intellectual* force. The class which has the means of material production at its disposal, has control at the same time over the means of mental production, so that thereby, generally speaking, the ideas of those who lack the means of mental production are subject to it. The ruling ideas are nothing more than the ideal expression of the dominant material relationships, the dominant material relationships grasped as ideas; hence of the relationships which make the one class the ruling one, therefore, the ideas of its dominance. The individuals composing the ruling class possess among other things consciousness, and therefore think. Insofar, therefore, as they rule as a class and determine the extent and compass of an epoch, it is self-evident that they do this in its whole range, hence among other things rule also as thinkers, as producers of ideas, and regulate the production and distribution of the ideas of their age: thus their ideas are the ruling ideas of the epoch. For instance, in an age and in a country where royal power, aristocracy, and bourgeoisie are contending for mastery and where, therefore, mastery is shared, the doctrine of the separation of powers proves to be the dominant idea and is expressed as an “eternal law.”

The division of labour manifests itself in the ruling class as the division of mental and material labour, so that inside this class one

part appears as the thinkers of the class (its active, conceptive ideologists, who make the perfecting of the illusion of the class about itself their chief source of livelihood), while the others’ attitude to these ideas and illusions is more passive and receptive, because they are in reality the active members of this class and have less time to make up illusions and ideas about themselves. Within this class this cleavage can even develop into a certain opposition and hostility between the two parts, which, however, in the case of a practical collision, in which the class itself is endangered, automatically comes to nothing, in which case there also vanishes the semblance that the ruling ideas were not the ideas of the ruling class and had a power distinct from the power of this class. The existence of revolutionary ideas in a particular period presupposes the existence of a revolutionary class.

If now in considering the course of history we detach the ideas of the ruling class from the ruling class itself and attribute to them an independent existence, if we confine ourselves to saying that these or those ideas were dominant at a given time, without bothering ourselves about the conditions of production and the producers of these ideas, if we thus ignore the individuals and world conditions which are the source of the ideas, we can say, for instance, that during the time that the aristocracy was dominant, the concepts honour, loyalty, etc. were dominant, during the dominance of the bourgeoisie the concepts freedom, equality, etc. The ruling class itself on

the whole imagines this to be so. This conception of history, which is common to all historians, particularly since the eighteenth century, will necessarily come up against the phenomenon that increasingly abstract ideas hold sway, i.e. ideas which increasingly take on the form of universality. For each new class which puts itself in the place of one ruling before it, is compelled, merely in order to carry through its aim, to represent its interest as the common interest of all the members of society, that is, expressed in ideal form: it has to give its ideas the form of universality, and represent them as the only rational, universally valid ones. The class making a revolution appears from the very start, if only because it is opposed to a *class*, not as a class but as the representative of the whole of society; it appears as the whole mass of society confronting the one ruling class.¹ It can do this because, to start with, its interest really is more connected with the common interest of all other non-ruling classes, because under the pressure of hitherto existing conditions its interest has not yet been able to develop as the particular interest of a particular class. Its victory, therefore, benefits also many individuals of the other classes which are not winning a dominant position, but only insofar as it now puts these individuals in a position to raise themselves into the ruling class. When the French bourgeoisie overthrew the power of

the aristocracy, it thereby made it possible for many proletarians to raise themselves above the proletariat, but only insofar as they become bourgeois. Every new class, therefore, achieves its hegemony only on a broader basis than that of the class ruling previously, whereas the opposition of the non-ruling class against the new ruling class later develops all the more sharply and profoundly. Both these things determine the fact that the struggle to be waged against this new ruling class, in its turn, aims at a more decided and radical negation of the previous conditions of society than could all previous classes which sought to rule.

This whole semblance, that the rule of a certain class is only the rule of certain ideas, comes to a natural end, of course, as soon as class rule in general ceases to be the form in which society is organised, that is to say, as soon as it is no longer necessary to represent a particular interest as general or the “general interest” as ruling.

NOTES

1. Universality corresponds to (1) the class versus the estate, (2) the competition, world-wide intercourse, etc., (3) the great numerical strength of the ruling class, (4) the illusion of the *common* interests (in the beginning this illusion is true), (5) the delusion of the ideologists and the division of labour. [*Marginal note by Marx.*]

10 ■ Ralf Dahrendorf

Class and Class Conflict in Industrial Society

One of the main questions which the present investigation is supposed to answer is: Do classes and class conflicts belong to that group of phenomena by which only the capitalist type of industrial society is characterized, or is their existence a consequence of industrial production itself, and are they therefore a lasting feature of industrial societies? This question will accompany us throughout the following analysis of changes in the structure of industrial societies since Marx.

Ownership and Control, or the Decomposition of Capital

Marx was right in seeking the root of social change in capitalist society in the sphere of industrial production, but the direction these changes took turned out to be directly contrary to Marx's expectations. With respect to capital, he had, in his later years, at least a vision of what was going to happen, as his brief and somewhat puzzled analysis of joint-stock companies shows. Joint-stock companies were legally recognized in Germany, England, France, and the United States in the second half of the nineteenth century. Laws often indicate the conclusion of social developments, and indeed early forms of joint-stock companies can be traced back at least to the commercial companies and trade societies of the seventeenth century. But it was in the nineteenth and early twentieth centuries that this type of enterprise first gained wide recognition and expanded into all branches

of economic activity. Today, more than two-thirds of all companies in advanced industrial societies are joint-stock companies, and their property exceeds four-fifths of the total property in economic enterprises. The enterprise owned and run by an individual, or even a family, has long ceased to be the dominant pattern of economic organization. . . .

According to the radical view, joint-stock companies involve a complete break with earlier capitalist traditions. By separating what has come to be called ownership and control, they give rise to a new group of managers who are utterly different from their capitalist predecessors. Thus for Marx, the joint-stock company involves a complete alienation of capital "from the real producers, and its opposition as alien property to all individuals really participating in production, from the manager down to the last day-laborer" (1953, Vol. III, p. 478). In other words, by separating ownership and control, the joint-stock company reduces the distance between manager and worker while at the same time removing the owners altogether from the sphere of production and thereby isolating their function as exploiters of others. It is merely a step from this kind of analysis to the thesis that, as Renner has it, the "capitalists without function" yield to the "functionaries without capital," and that this new ruling group of industry bears little resemblance to the old "full capitalists" (1953, pp. 182, 198). Burnham, Geiger, Sering, and others followed Marx (and Renner) in this

radical interpretation of the social effects of joint-stock companies.

The conservative view, on the other hand, holds that the consequences of the apparent separation of ownership and control have been vastly overrated. It is argued that in fact owners and controllers, i.e., stockholders and managers, are a fairly homogeneous group. There are often direct connections between them, and where this is not the case, their outlook is sufficiently similar to justify insisting on the old assumption of a homogeneous class of capitalists opposed to an equally homogeneous class of laborers. This view is not often heard in the West nowadays, although traces of it are evident in the work of C. Wright Mills (1954, 1956). It may be added that this conservative view is clearly contrary to Marx's own analysis. . . .

There is little reason to follow Marx and describe the condition of separation of ownership and control as a transitional form of historical development. It is no more transitional than any other stage of history, and it has already proven quite a vital pattern of social and economic structure. But I think that we can follow Marx in his radical interpretation of this phenomenon. The separation of ownership and control has replaced one group by two whose positions, roles, and outlooks are far from identical. In taking this view, one does of course agree with Marx against himself. For it follows from this that the homogeneous capitalist class predicted by Marx has in fact not developed. Capital—and thereby capitalism—has dissolved and given way in the economic sphere, to a plurality of partly agreed, partly competing, and partly simply different groups. The effect of this development on class conflict is threefold: first, the replacement of capitalists by managers involves a change in the composition of the groups participating in conflict; second, and as a consequence of this change in recruitment and composition, there is a change in the nature of the issues that cause conflicts, for the interests of the functionaries without capital differ

from those of full-blown capitalists, and so therefore do the interests of labor vis-à-vis their new opponents; and third, the decomposition of capital involves a change in the patterns of conflict. One might question whether this new conflict, in which labor is no longer opposed to a homogeneous capitalist class, can still be described as a class conflict at all. In any case, it is different from the division of the whole society into two great and homogeneous hostile camps with which Marx was concerned. While I would follow the radical view of the separation of ownership and control in industry to this point, there is one thing to be said in favor of the conservative view. Changes in the composition of conflict groups, of the issues, and of patterns of conflict do not imply the abolition of conflict or even of the specific conflict between management and labor in industry. Despite the effects of the decomposition of capital on class structure, we have no reason to believe that antagonisms and clashes of interest have now been banned from industrial enterprises.

Skill and Stratification, or the Decomposition of Labor

While Marx had at least a premonition of things to come with respect to capital, he remained unaware of developments affecting the unity and homogeneity of labor. Yet in this respect, too, the sphere of production which loomed so large in Marx's analyses became the starting point of changes that clearly refute his predictions. The working class of today, far from being a homogeneous group of equally unskilled and impoverished people, is in fact a stratum differentiated by numerous subtle and not-so-subtle distinctions. Here, too, history has dissolved one position, or role, and has substituted for it a plurality of roles that are endowed with diverging and often conflicting expectations. . . .

Analysis of industrial conditions suggests quite clearly that within the labor force of advanced industry we have to distinguish at least

three skill groups: a growing stratum of highly skilled workmen who increasingly merge with both engineers and white-collar employees, a relatively stable stratum of semiskilled workers with a high degree of diffuse as well as specific industrial experience, and a dwindling stratum of totally unskilled laborers who are characteristically either newcomers to industry (beginners, former agricultural laborers, immigrants) or semi-unemployables. It appears, furthermore, that these three groups differ not only in their level of skill, but also in other attributes and determinants of social status. The semiskilled almost invariably earn a higher wage than the unskilled, whereas the skilled are often salaried and thereby participate in white-collar status. The hierarchy of skill corresponds exactly to the hierarchy of responsibility and delegated authority within the working class. From numerous studies it would seem beyond doubt that it also correlates with the hierarchy of prestige, at the top of which we find the skilled man whose prolonged training, salary, and security convey special status, and at the bottom of which stands the unskilled man who is, according to a recent German investigation into workers' opinions, merely "working" without having an "occupation" proper (see Kluth, 1955, p. 67). Here as elsewhere Marx was evidently mistaken. "Everywhere, the working class differentiates itself more and more, on the one hand into occupational groups, on the other hand into three large categories with different, if not contradictory, interests: the skilled craftsmen, the unskilled laborers, and the semiskilled specialist workers" (Philip, 1955, p. 2).

In trying to assess the consequences of this development, it is well to remember that, for Marx, the increasing uniformity of the working class was an indispensable condition of that intensification of the class struggle which was to lead, eventually, to its climax in a revolution. The underlying argument of what for Marx became a prediction appears quite plausible. For there to be a revolution, the conflicts within a society have to become extremely in-

tense. For conflicts to be intense, one would indeed expect its participants to be highly unified and homogeneous groups. But neither capital nor labor have developed along these lines. Capital has dissolved into at least two, in many ways distinct, elements, and so has labor. The proletariat, the impoverished slave of industry who is indistinguishable from his peers in terms of his work, his skill, his wage, and his prestige, has left the scene. What is more, it appears that by now he has been followed by his less depraved, but equally alienated successor, the worker. In modern industry, "the worker" has become precisely the kind of abstraction which Marx quite justly resented so much. In his place, we find a plurality of status and skill groups whose interests often diverge. Demands of the skilled for security may injure the semiskilled; wage claims of the semiskilled may raise objections by the skilled; and any interest on the part of the unskilled is bound to set their more highly skilled fellow workmen worrying about differentials.

Again, as in the case of capital, it does not follow from the decomposition of labor that there is no bond left that unites most workers—at least for specific goals; nor does it follow that industrial conflict has lost its edge. But here, too, a change of the issues and, above all, of the patterns of conflict is indicated. As with the capitalist class, it has become doubtful whether speaking of the working class still makes much sense. Probably Marx would have agreed that class "is a force that unites into groups people who differ from one another, by overriding the differences between them" (Marshall, 1950, p. 114), but he certainly did not expect the differences to be so great, and the uniting force so precarious as it has turned out to be in the case both of capital and of labor. . . .

The Institutionalization of Class Conflict

A historian might argue that all the tendencies of change here described as changes in

the structure of industrial societies since Marx had in fact begun before and in some cases long before Marx died in 1883. . . . There is, however, one line of social development in industrial societies which has both originated and spread since about the time of Marx's death, and which is directly relevant to our problem. Geiger, who has described this change as the "institutionalization of class conflict," says: "The tension between capital and labor is recognized as a principle of the structure of the labor market and has become a legal institution of society. . . . The methods, weapons, and techniques of the class struggle are recognized—and are thereby brought under control. The struggle evolves according to certain rules of the game. Thereby the class struggle has lost its worst sting, it is converted into a legitimate tension between power factors which balance each other. Capital and labor struggle with each other, conclude compromises, negotiate solutions, and thereby determine wage levels, hours of work, and other conditions of work" (1949, p. 184).

Marx displayed a certain sociological naïveté when he expressed his belief that capitalist society would be entirely unable to cope with the class conflict generated by its structure. In fact, every society is capable of coping with whatever new phenomena arise in it, if only by the simple yet effective inertia which can be described, a little pretentiously, as the process of institutionalization. In the case of class conflict, institutionalization assumed a number of successive and complementary forms. It began with the painful process of recognition of the contending parties as legitimate interest groups. Within industry, a "secondary system of industrial citizenship" (Marshall, 1950, p. 68) enabled both workers and entrepreneurs to associate and defend their interests collectively. Outside industry, the primary system of political citizenship had the same effect. And while, in the stage of organization, conflict may develop a greater visible intensity,

organization has at least two side effects which operate in the opposite direction. Organization presupposes the legitimacy of conflict groups, and it thereby removes the permanent and incalculable threat of guerrilla warfare. At the same time, it makes systematic regulations of conflicts possible. Organization is institutionalization, and whereas its manifest function is usually an increasingly articulate and outspoken defense of interests, it invariably has the latent function also of inaugurating routines of conflict which contribute to reducing the violence of clashes of interest. . . .

Nobody can, of course, ever be sure that a given pattern of conflict regulation will always prove successful. There are still strikes, and for all we know they will continue to occur. But it has proved possible for industrial society to get along with the clashes of interest arising from its industrial and political structure—and it has proved possible for interest groups to get along with industrial society. Instead of a battlefield, the scene of group conflict has become a kind of market in which relatively autonomous forces contend according to certain rules of the game, by virtue of which nobody is a permanent winner or loser. This course of development must naturally be bitter for the orthodox and the dogmatic, but theirs is the kind of bitterness which makes liberal minds rejoice. . . .

Power and Authority

One of the central theses of this study consists in the assumption that the differential distribution of authority invariably becomes the determining factor of systematic social conflicts of a type that is germane to class conflicts in the traditional (Marxian) sense of this term. The structural origin of such group conflicts must be sought in the arrangement of social roles endowed with expectations of domination or subjection. Wherever there are such roles, group conflicts of the type in question are to be ex-

pected. Differentiation of groups engaged in such conflicts follows the lines of differentiation of roles that are relevant from the point of view of the exercise of authority. Identification of variously equipped authority roles is the first task of conflict analysis;¹ conceptually and empirically all further steps of analysis follow from the investigation of distributions of power and authority.

“Unfortunately, the concept of power is not a settled one in the social sciences, either in political science or in sociology” (Parsons, 1957, p. 139). Max Weber (1947), Pareto (1955), Mosca (1950), later Russell (1938), Bendix (1952), Lasswell (1936), and others have explored some of the dimensions of this category; they have not, however, reached such a degree of consensus as would enable us to employ the categories of power and authority without at least brief conceptual preliminaries. So far as the terms “power” and “authority” and their distinction are concerned, I shall follow in this study the useful and well-considered definitions of Max Weber. For Weber, power is the “probability that one actor within a social relationship will be in a position to carry out his own will despite resistance, regardless of the basis on which this probability rests”; whereas authority (*Herrschaft*) is the “probability that a command with a given specific content will be obeyed by a given group of persons” (1947, p. 28). The important difference between power and authority consists in the fact that whereas power is essentially tied to the personality of individuals, authority is always associated with social positions or roles. The demagogue has power over the masses to whom he speaks or whose actions he controls; but the control of the officer over his men, the manager over his workers, the civil servant over his clientele is authority, because it exists as an expectation independent of the specific person occupying the position of officer, manager, civil servant. It is only another way of putting this difference if we say—as does Max Weber—that while power is

merely a factual relation, authority is a legitimate relation of domination and subjection. In this sense, authority can be described as legitimate power.

In the present study we are concerned exclusively with relations of authority, for these alone are part of social structure and therefore permit the systematic derivation of group conflicts from the organization of total societies and associations within them. The significance of such group conflicts rests with the fact that they are not the product of structurally fortuitous relations of power but come forth wherever authority is exercised—and that means in all societies under all historical conditions. (1) Authority relations are always relations of super- and subordination. (2) Where there are authority relations, the superordinate element is socially expected to control, by orders and commands, warnings and prohibitions, the behavior of the subordinate element. (3) Such expectations attach to relatively permanent social positions rather than to the character of individuals; they are in this sense legitimate. (4) By virtue of this fact, they always involve specification of the persons subject to control and of the spheres within which control is permissible. Authority, as distinct from power, is never a relation of generalized control over others. (5) Authority being a legitimate relation, noncompliance with authoritative commands can be sanctioned; it is indeed one of the functions of the legal system (and of course of quasi-legal customs and norms) to support the effective exercise of legitimate authority.

Alongside the term “authority,” we shall employ in this study the terms “domination” and “subjection.” These will be used synonymously with the rather clumsy expressions “endowed with authority” or “participating in the exercise of authority” (domination), and “deprived of authority” or “excluded from the exercise of authority” (subjection).

It seems desirable for purposes of conflict analysis to specify the relevant unit of social

organization in analogy to the concept of social system in the analysis of integration. To speak of specification here is perhaps misleading. “Social system” is a very general concept applicable to all types of organization; and we shall want to employ an equally general concept which differs from that of social system by emphasizing a different aspect of the same organizations. It seems to me that Max Weber’s category “imperatively coordinated association” (*Herrschaftsverband*) serves this purpose despite its clumsiness. . . .

Empirically it is not always easy to identify the border line between domination and subjection. Authority has not remained unaffected by the modern process of division of labor. But even here, groups or aggregates can be identified which do not participate in the exercise of authority other than by complying with given commands or prohibitions. Contrary to all criteria of social stratification, authority does not permit the construction of a scale. So-called hierarchies of authority (as displayed, for example, in organization charts) are in fact hierarchies of the “plus-side” of authority, i.e., of the differentiation of domination; but there is, in every association, also a “minus-side” consisting of those who are subjected to authority rather than participate in its exercise.

In two respects this analysis has to be specified, if not supplemented. First, for the individual incumbent of roles, domination in one association does not necessarily involve domination in all others to which he belongs, and subjection, conversely, in one association does not mean subjection in all. The dichotomy of positions of authority holds for specific associations only. In a democratic state, there are both mere voters and incumbents of positions of authority such as cabinet ministers, representatives, and higher civil servants. But this does not mean that the “mere voter” cannot be incumbent of a position of authority in a different context, say, in an industrial enterprise; conversely, a cabinet minister may be,

in his church, a mere member, i.e., subject to the authority of others. Although empirically a certain correlation of the authority positions of individuals in different associations seems likely, it is by no means general and is in any case a matter of specific empirical conditions. It is at least possible, if not probable, that if individuals in a given society are ranked according to the sum total of their authority positions in all associations, the resulting pattern will not be a dichotomy but rather like scales of stratification according to income or prestige. For this reason it is necessary to emphasize that in the sociological analysis of group conflict the unit of analysis is always a specific association and the dichotomy of positions within it.

As with respect to the set of roles associated with an individual, total societies, also, do not usually present an unambiguously dichotomic authority structure. There are a large number of imperatively coordinated associations in any given society. Within every one of them we can distinguish the aggregates of those who dominate and those who are subjected. But since domination in industry does not necessarily involve domination in the state, or a church, or other associations, total societies can present the picture of a plurality of competing dominant (and, conversely, subjected) aggregates. This, again, is a problem for the analysis of specific historical societies and must not be confounded with the clearer lines of differentiation within any one association. Within the latter, the distribution of authority always sums up to zero, i.e., there always is a division involving domination and subjection.

I need hardly emphasize that from the point of view of “settling” the concepts of power and authority, the preceding discussion has raised more problems than it has solved. I believe, however, that for the purposes of this study, and of a sociological theory of conflict, little needs to be added to what has been stated here. In order somewhat to substantiate this perhaps rather bold assertion, it seems useful to recapitulate

briefly the heuristic purpose and logical status of the considerations of this section.

I have introduced, as a structural determinant of conflict groups, the category of authority as exercised in imperatively coordinated associations. While agreeing with Marx that source and level of income—even socioeconomic status—cannot usefully be conceived as determinants of conflict groups, I have added to this list of erroneous approaches Marx's own in terms of property in the means of production. Authority is both a more general and a more significant social relation. The former has been shown in our critique of Marx; the latter will have to be demonstrated elsewhere (see Dahrendorf 1959). The concept of authority is used, in this context, in a specific sense. It is differentiated from power by what may roughly be referred to as the element of legitimacy; and it has to be understood throughout in the restricted sense of authority as distributed and exercised in imperatively coordinated associations. While its "disruptive" or conflict-generating consequences are not the only aspect of authority, they are the one relevant in terms of the coercion model of society. Within the frame of reference of this model, (1) the distribution of authority in associations is the ultimate "cause" of the formation of conflict groups, and (2) being dichotomous, it is, in any given association, the cause of the formation of two, and only two, conflict groups.

NOTES

1. To facilitate communication, I shall employ in this study a number of abbreviations. These must

not however be misunderstood. Thus, "conflict analysis" in this context stands for "analysis of group conflicts of the class type, class being understood in the traditional sense." At no point do I want to imply a claim for a generalized theory of social conflict.

BIBLIOGRAPHY

- Reinhard Bendix. 1952. "Bureaucracy and the Problem of Power." *Reader in Bureaucracy*. Edited by R.K. Merton, A.P. Gray, B. Hockey, and H.C. Selvin. Glencoe.
- Ralf Dahrendorf. 1959. *Class and Class Conflict in Industrial Society*. Stanford.
- Theodor Geiger. 1949. *Die Klassengesellschaft im Schmelztiegel*. Cologne and Hagen.
- Heinz B. Kluth. 1955. "Arbeiterjugend—Begriff und Wirklichkeit," in Helmut Schelsky, ed., *Arbeiterjugend—gestern und heute*. Heidelberg.
- Harold Lasswell. 1936. *Politics—Who Gets What, When, and How?* New York.
- T.H. Marshall. 1950. *Citizenship and Social Class*. Cambridge.
- Karl Marx. 1953. *Das Kapital*. Vols. I, III. New ed. Berlin.
- C.W. Mills. 1954. *The New Men of Power*. New York.
- C.W. Mills. 1956. *The Power Elite*. New York.
- Gaetano Mosca. 1950. *Die herrschende Klasse*. Bern.
- Vilfredo Pareto. 1955. *Allgemeine Soziologie*. Translated and edited by C. Brinkmann. Tübingen.
- Talcott Parsons. 1957. "The Distribution of Power in American Society," *World Politics* X, No. 1, October.
- André Philip. 1955. *La Démocratie Industrielle*. Paris.
- Karl Renner. 1953. *Wandlungen der Modernen Gesellschaft: Zwei Abhandlungen Über die Probleme der Nachkriegszeit*. Vienna.
- Bertrand Russell. 1938. *Power: A New Social Analysis*. London.
- Paul Sering. 1947. *Jenseits des Kapitalismus*. Nürnberg.
- Max Weber. 1947. *Wirtschaft und Gesellschaft* (Grundriss der Sozialökonomik, section III). 4th ed. Tübingen.

11 ■ Erik Olin Wright

A General Framework for the Analysis of Class Structure

The Point of Departure: Neo-Marxist Analyses of Class Structure

At the heart of the recent resurgence of Marxist theorizing on the problem of class has been what might be termed the “embarrassment” of the middle class. For all of their disagreements, all Marxists share a basic commitment to a polarized abstract concept of class relations. Yet, at least at first glance, the concrete class structures of contemporary advanced capitalist societies look anything but polarized. This empirical evidence of a large middle class has provided critics of Marxism with one of their principal arguments against Marxist class theory. In response, a variety of solutions to the problem of the middle class have been proposed in the recent Marxist debates.

Without going into any detail, it is possible to identify four broadly different strategies that Marxists have adopted to deal with the conceptual problem of nonpolarized class positions within a logic of polarized class relations.¹ First, the class structure of advanced capitalist societies really *is* polarized; the “middle class” is strictly an ideological illusion. This position deals with the problem of the middle class by denying the problem itself. Second, the middle class should be viewed as a *segment* of some other class, typically a “new petty bourgeoisie” or “new working class.”² In

this strategy the basic class map of capitalism remains intact, but significant internal differentiations within classes are added to the analysis of class structure. Third, the middle class is really a new class in its own right, completely distinct from either the bourgeoisie, the proletariat, or the petty bourgeoisie. Sometimes this class is given a specific name, such as the Professional Managerial Class,³ sometimes it is simply called “the New Class.”⁴ By adding entirely new classes to the class structure, this approach more radically alters the class map of capitalism than the class-segment strategy. Fourth, the positions aggregated under the popular rubric “middle class” are not really in *a* class at all. Rather they should be viewed as locations that are simultaneously in more than one class, positions that I have characterized as “contradictory locations within class relations.”⁵ Managers, for example, should be viewed as simultaneously in the working class (in so far as they are wage laborers dominated by capitalists) and in the capitalist class (in so far as they control the operation of production and the labor of workers). This strategy departs most from the traditional Marxist vision of class structure since the very meaning of a “location” is altered: there is no longer a one-to-one correspondence between structural locations filled by individuals and classes.

Eric Olin Wright. “A General Framework for the Analysis of Class Structure,” *Politics and Society* 13 (1984), pp. 383–397, 399–402, 417–422. Copyright © 1984 by Sage Publications, Inc. Reprinted by permission of Sage Publications, Inc.

I no longer feel that this fourth solution is satisfactory. Specifically, it suffers from two important problems that it shares with most other neo-Marxist conceptualizations of class structure: it tends to shift the analysis of class relations from exploitation to domination; and it implicitly regards socialism—a society within which the working class is the “ruling class”—as the only possible alternative to capitalism.

Domination Versus Exploitation

Throughout the development of the concept of contradictory class locations I have insisted that this was a reformulation of a distinctively Marxist class concept. As part of the rhetoric of such an enterprise, I affirmed the relationship between class and exploitation. Nevertheless, in practice the concept of contradictory locations within class relations rested almost exclusively on relations of *domination* rather than exploitation. Reference to exploitation functioned more as a background concept to the discussion of classes than as a constitutive element of the analysis of class structures. Managers, for example, were basically defined as a contradictory location because they were simultaneously dominators and dominated. Domination relations were also decisive in defining the class character of “semiautonomous employees”—locations that, I argued, were simultaneously petty bourgeois and proletarian by virtue of their self-direction within the labor process—since “autonomy” defines a condition with respect to domination. This same tendency of substituting domination for exploitation at the core of the concept of class is found in most other neo-Marxist conceptualizations of class structure.

For some people, of course, marginalizing the concept of exploitation is a virtue, not a sin. My own view, however, is that this is a serious weakness. The marginalization of exploitation both undermines claims that classes have “objective” interests and erodes the centrality Marxists have accorded class in social theory.

The concept of domination does not in and of itself imply any specific interest of actors. Parents dominate small children, but this does not imply that they have intrinsically opposed interests to their children. What would make those interests antagonistic is if the relation of parents to children were exploitative as well. Exploitation, unlike domination, intrinsically implies a set of opposing material interests. If we wish to retain some sense in which the interests of individuals as members of classes are not simply whatever interests those individuals subjectively hold, then the shift to a domination-centered concept renders this more difficult.

Domination-centered concepts of class also tend to slide into what can be termed “the multiple oppressions” approach to understanding society. Societies, in this view, are characterized by a plurality of oppressions each rooted in a different form of domination—sexual, racial, national, economic—none of which has any explanatory priority over any other. Class, then, becomes just one of many oppressions, with no particular centrality for social and historical analysis. How important class is in a given society becomes an historically contingent question.

Again, this displacement of class from the center stage may be viewed as an achievement rather than a problem. It may be that class should not occupy a privileged place in social theory. But if one believes, as Marxists traditionally have believed, that only by giving class this central place is it possible to develop a scientific theory of the trajectory of historical development, and in particular, a theory of the real historical alternatives to capitalism, then the domination-centered concept of class risks eroding the theoretical justification for Marxian class analysis itself.

Classes in Postcapitalist Societies

Classical Marxism was absolutely unequivocal about the historical prognosis for capitalism: socialism—and ultimately communism—was the future of capitalist societies. The bearer of

that necessary future was the working class. The polarized class structure *within* capitalism between the bourgeoisie and the proletariat thus paralleled the polarized historical alternatives *between* capitalism and socialism.

The actual historical experience of the twentieth century has called into question, although not unambiguously refuted, this historical vision. As I have argued elsewhere, it is necessary to at least entertain the possibility of postcapitalist class structures.⁶ The difficulty is that with very few exceptions, the conceptual frameworks adopted by Marxists for analyzing capitalist class relations do not contain adequate criteria for understanding postcapitalist classes.⁷ In particular, all of the class categories in my analysis of contradictory locations within class relations were either situated firmly within capitalist relations (bourgeoisie, managers, workers) or in contradictory locations involving basically precapitalist relations (semiautonomous employees, the petty bourgeoisie, small employers). There were no elements within this analysis of class relations in capitalist society that could point the direction for the analysis of postcapitalist classes. The result is a tendency for discussions of postcapitalist class structures—the class structures of “actually existing socialism”—to have a very ad hoc character to them.

Given these conceptual problems—the shift from exploitation to domination and the lack of a conceptual basis for analyzing postcapitalist classes—there are really two theoretical alternatives that could be pursued. One possibility is to celebrate the shift to a domination-centered concept and use this new class concept as the basis for analyzing both capitalist and postcapitalist society. This would lead class analysis firmly in the direction of Dahrendorf’s analysis of classes as positions within authority relations.⁸ A second alternative is to attempt to restore exploitation as the center of class analysis in such a way that it can both accommodate the empirical complexities of the middle class

within capitalism and the historical reality of postcapitalist class structures. It is this second course of action that I will pursue in the rest of this paper.

The basis for this reconstruction of an exploitation-centered concept of class comes from the recent work of John Roemer.⁹ While Roemer himself has not been particularly concerned with problems of empirical investigation or the elaboration of concrete maps of class structures, nevertheless his work does provide a rich foundation for such endeavors. As I will attempt to show, with suitable modification and extension, his strategy of analysis can provide a rigorous basis for resolving the problems in the concept of contradictory class locations.

Roemer’s Account of Class and Exploitation

The Concept of Exploitation

We observe inequalities in the distribution of incomes, the real consumption packages available to individuals, families, groups. The concept of exploitation is a particular way of analyzing such inequalities. To describe an inequality as reflecting exploitation is to make the claim that there exists a particular kind of causal relationship between the incomes of different actors. More concretely, we will say that the rich exploit the poor when two things can be established: that the welfare of the rich causally depends on the deprivations of the poor—the rich are rich *because* the poor are poor; and that the welfare of the rich depends upon the *effort* of the poor—the rich, through one mechanism or another, appropriate part of the fruits of labor of the poor. The first of these criteria by itself defines *economic oppression*, but not exploitation. Unemployed workers, in these terms, are economically oppressed but not exploited. Exploitation implies both economic oppression and appropriation of at least part of the social surplus by the oppressor.

The traditional Marxist concept of exploitation is clearly a special case of this general concept. In Marxian exploitation one class appropriates the surplus labor performed by another class through various mechanisms. The income of the exploiting class comes from the labor performed by the exploited class. There is thus a straightforward causal linkage between the poverty and effort of the exploited and the affluence of the exploiter. The latter benefits at the expense of the former.

Roemer has attempted to elaborate this view of exploitation using two strategies. The first of these involves studying through a series of formal mathematical models the flows of “surplus labor” from one category of actors to another in the course of various exchange relations; the second involves adopting a kind of game-theory approach to specifying different forms of exploitation. Let us briefly examine each of these in turn.

The Labor-Transfer Approach

The analysis of labor transfers is an extension of the traditional Marxist view of exploitation, although Roemer self-consciously does not rely on the labor theory of value in order to explore such labor transfers. The main target of his analysis is the view, commonly held by Marxists, that a necessary condition for the exploitation of labor in a market economy is the institution of wage labor. Roemer demonstrates two basic propositions. First, Roemer demonstrates that exploitation can occur in an economy in which all producers own their own means of production and in which there is no market in labor power and no credit market (that is, no borrowing). The only things that are traded are products. In such an economy if different producers own different amounts of productive assets such that different producers have to work different numbers of hours to produce the exchange-equivalent of their own subsistence, then free trade among these producers will lead to exploitation of the asset poor by the

asset rich. What Roemer shows in this simple economy is not simply that some producers work less than others for the same subsistence, but that the workers who work less are able to do so *because* the less-endowed producers have to work more. The critical proof in this example is that if the asset-poor person simply stopped producing—died—and the asset-rich person took over the asset-poor’s assets, then the asset-rich producer would have to work longer hours than before to maintain the same subsistence. There is thus not merely an inequality among the producers in this economy, but exploitation as well.

Second, Roemer demonstrates that there is complete symmetry in the structure of exploitation in a system in which capital hires wage laborers and in a system in which workers rent capital (that is, systems with credit and labor markets). For this analysis, he compares the class structures and patterns of exploitation on the two imaginary islands, “labor-market island” and “credit-market island.” On both islands some people own no means of production and other people own varying amounts of the means of production. The distribution of these assets is identical on the two islands. And on both islands people have the same motivations: they all seek to minimize the amount of labor-time they must expend to achieve a common level of subsistence. The two islands differ in only one respect: on the labor-market island people are allowed to sell their labor power, whereas on the credit-market island people are prohibited from selling their labor power but are allowed to borrow, at some interest rate, the means of production. Roemer shows that on each island there is a strict correspondence between class location (derived from ownership of differing amounts of means of production, including no means of production) and exploitation status (having one’s surplus labor appropriated by someone else). This is what he terms the “Class-Exploitation Correspondence Principle.” He

also shows that the two class structures are completely isomorphic: every individual on one island would be in exactly the same exploitation status on the other island.

The upshot of these two propositions (and others that Roemer explores) is the claim that market-based exploitation is strictly a consequence of inequalities in the distribution of the means of production. However, while this may typically play itself out through a labor market, this is only one concrete institutional form for such exploitation: it is not the necessary condition for the exploitation to occur.

The Game-Theory Approach

While the labor-transfer analyses of exploitation were primarily designed to reveal the underlying logic of exploitation in market exchanges, the game-theory approach is used by Roemer to compare different systems of exploitation. The idea is to compare different systems of exploitation by treating the organization of production as a “game” and asking if a coalition of players would be better off if they withdrew from the game under certain specified procedures. Different types of exploitation are defined by the withdrawal rules that would make certain agents better off.

More formally, Roemer argues that a coalition of actors S can be said to be exploited, and another coalition S' (the complement of S) can be said to be exploiting, if “there is no alternative, which we may conceive of as hypothetically feasible, in which S would be better off than in its present situation, [and if,] under this alternative, the complement to S . . . would be worse off than at present.”¹⁰ The counterfactual in these two conditions is meant to convey the sense in which the welfare of S' is causally dependent upon the deprivation of S .

Roemer uses this strategy to define three kinds of exploitation: feudal exploitation, capitalist exploitation, and what he refers to as socialist exploitation. Let's begin with cap-

italist exploitation. Workers own no physical assets (means of production) and sell their labor power to capitalists for a wage. Are workers exploited under capitalism? The answer to this question, in the game theoretic formulation, requires posing an alternative game to the game of capitalism within which the two conditions specified above hold. What is the alternative? It is a game within which each worker receives his/her *per capita share of society's total productive assets*. What Roemer demonstrates is that if the coalition of all wage-earners were to leave the game of capitalism with their per capita share of society's assets, then they would be better off than staying in capitalism, and capitalists would be worse off. The “withdrawal rule” in this case—leaving the game with per capita shares of physical assets—then becomes the formal “test” of whether or not a particular social system involves capitalistic exploitation.

In contrast, the withdrawal rule to specify feudal exploitation is leaving the game with one's *personal assets* (rather than one's per capita share of total social assets). This is equivalent to the feudal serf being freed from all obligations based on personal bondage. Peasants would be better off under such circumstances; feudal lords would be worse off.¹¹

The concept of the socialist exploitation is the least systematically worked out in Roemer's analysis. The withdrawal rule in this case is leaving the game with one's *per capita share of inalienable assets* (skills). A coalition will be said to be socialistically exploited if it would improve its position by leaving with its per capita skills while its complement would be worse off under such circumstances. This implies that people with high levels of skills in the game receive high income not simply because they have high skills, but because of the differentials in skill levels across actors. The highly skilled would become worse off if the unskilled obtained skills; they thus have an interest in maintaining skill differentials, and this is what underpins the claim that their income reflects ex-

ploitation.¹² If a skilled person's income reflected no more than the amount of time and resources it takes to obtain the skill, then there would be no skill-based exploitation. The higher incomes would simply be reimbursement for real costs incurred. The argument behind skill exploitation is that people with scarce skills receive incomes above the costs of producing those skills, a "rent" component to their income; it is this element that constitutes exploitation.

Class and Exploitation

The central message of both of Roemer's strategies for analyzing exploitation is that the material basis of exploitation is inequalities in distributions of productive assets, or what is usually referred to as property relations. On the one hand, inequalities of assets are sufficient to account for transfers of labor surplus; on the other hand, different forms of asset inequality specify different systems of exploitation. Classes are then defined as positions within the social relations of production derived from these relations of exploitation.¹³

These conclusions have led Roemer to challenge directly the tendency of Marxists (like myself) to define class relations primarily in terms of domination relations within production. Of course, exploiting classes dominate exploited classes in the sense of preventing the exploited classes from taking the exploiting class's productive assets. But domination *within* production, Roemer insists, is not a central part of defining class relations as such.

In previous work I have criticized Roemer's position on this issue.¹⁴ I argued that class relations intrinsically involved domination *at the point of production*, not simply in the repressive protection of the property relations as such. I now think that Roemer is correct on this point. That capitalists boss workers around within production is unquestionably an important feature of most historic forms of capitalist production and may play an im-

portant role in explaining the forms of class organization and class conflict within production. However, the basis of the capital-labor relation should be identified with relations of effective control (that is, real economic ownership) over productive assets as such.

One of the reasons why I resisted Roemer's conceptualization of classes in terms of property relations is that it seemed to blur the difference between Marxist definitions of class and Weberian definitions. Weberian definitions, as I construed them, were "market based" definitions of class, whereas Marxist definitions were "production based." The reputed advantage of the latter was that production was more "fundamental" than exchange, and therefore production-based class concepts had more explanatory power than market-based concepts.

What now seems clear to me is that definitions of classes in terms of property relations should not be identified with strictly market-based definitions. Property-relations accounts of classes do not define classes by income shares, by the results of market transactions, but by the productive assets that classes control, which lead them to adopt certain strategies within exchange relations and which thereby determine the outcomes of those market transactions.

Toward a General Framework of Class Analysis

Extending Roemer's Analysis

The heart of Roemer's analysis is the linkage between the distribution of productive assets of various sorts and exploitation. Different mechanisms of exploitation are defined by different kinds of assets, and different class systems are defined by which of these assets is most important for shaping the patterns of exploitation in the society.

In Roemer's own explicit formulation, only two kinds of assets are formally considered: physical assets (alienable assets in his terminology) and skill assets (inalienable

assets). The distinction between exploitation in feudalism and exploitation in capitalism revolves around the nature of the withdrawal rules with respect to physical assets (withdrawing with one's personal assets to define feudal exploitation versus withdrawing with one's per capita share of assets to define capitalist exploitation). The feudal case, however, can be characterized in a somewhat different way. Labor power is a productive asset.¹⁵ In capitalist societies everyone owns one unit of this asset, namely themselves. In feudalism, on the other hand, ownership rights over labor power are unequally distributed: feudal lords have more than one unit, serfs have less than one unit. To be sure, it is not typical of feudalism for serfs to own no labor power—they are generally not slaves divested of all ownership rights in their own labor power—but they do not have complete effective control over their own persons as productive actors, and this is what it means to “own” one's own labor power assets. The withdrawal rule that defines feudal exploitation can then be specified as leaving the feudal game with one's per capita share of society's assets in labor power, namely one unit. Feudal exploitation is thus exploitation (transfers of labor) that results from inequalities in the distribution of assets in labor power.

Reformulating feudal exploitation in this manner makes the game-theory specification of different exploitations in Roemer's analysis symmetrical: feudal exploitation is based on inequalities generated by ownership of labor-power assets; capitalist exploitation on inequalities generated by ownership of alienable assets; socialist exploitation on inequalities generated by ownership of inalienable assets. And corresponding to each of these exploitation-generating inequalities of assets, there is a specific class relation: lords and serfs in feudalism, bourgeoisie and proletariat in capitalism, experts and workers in socialism.

But how, it might be asked, should “actually existing socialist societies” be theorized within these categories? The anticapitalist rev-

olution in Russia resulted in the virtual elimination of private property in the means of production: individuals cannot own means of production, they cannot inherit them or dispose of them on a market, and so on. And yet it seems unsatisfactory to characterize such societies simply in terms of skill-based exploitation. Experts do not appear to be the “ruling class” in those societies, and the dynamic of the societies does not seem to revolve around skill inequalities as such.

Roemer recognized this problem and introduced what he termed “status exploitation” to deal with it. The exploitation exercised by bureaucrats is the prototypical example. “If these positions,” Roemer writes, “required special skills, then one might be justified in calling the differential remuneration to these positions an aspect of socialist [skill-based] exploitation. . . . [However] there is some extra remuneration to holders of those positions which accrues solely by virtue of the position and not by virtue of the skill necessary to carry out the tasks associated with it. These special payments to positions give rise to *status exploitation*.”¹⁶

Roemer's concept of status exploitation is unsatisfactory for two principal reasons. First, it is outside of the logic of the rest of his analysis of exploitation. In each of the other cases, exploitation is rooted in relations to the forces of production. Each of the other forms of exploitation is “materialist” not only because the concept is meant to explain material distribution, but also because it is based on the relation to the material conditions of production. “Status” exploitation has no necessary relationship to production at all. Second, it is hard to rigorously distinguish status exploitation from feudal exploitation. The “lord” receives remuneration strictly because of an incumbency in a position, not because of skills or ownership of capital. Yet, it hardly seems reasonable to consider the logic of exploitation and class in the contemporary Soviet Union and in fourteenth-century feudal Europe as being essentially the same.

The problems with the concept of status exploitation can be solved by analyzing exploitation based on a fourth element in the inventory of productive assets, an asset that can be referred to as “organization.” As both Adam Smith and Marx noted, the technical division of labor among producers is itself a source of productivity. The way the production process is organized is a productive resource independent of the expenditure of labor power, the use of means of production, or the skills of the producer. Of course there is an interrelationship between organization and these other assets, just as there is an interdependence between means of production and skills. But organization—the conditions of coordinated cooperation among producers in a complex division of labor—is a productive resource in its own right.

How is this asset distributed in different kinds of societies? In contemporary capitalism, organization assets are generally controlled by managers and capitalists: managers control the organization assets within specific firms under constraints imposed by the ownership of the capital assets by capitalists. Entrepreneurial capitalists directly control both kinds of assets (and probably skill assets as well); pure rentier capitalists (“coupon clippers”) only own capital assets. Because of the anarchy of the capitalist market, no set of actors controls the technical division of labor across firms.

In state bureaucratic socialism, organization assets assume a much greater importance. Controlling the technical division of labor—the coordination of productive activities within and across labor processes—becomes a societal task organized at the center. The control over organization assets is no longer simply the task of firm-level managers but extends into the central organs of planning within the state. Exploitation in such societies is thus based on bureaucratic power: the control over organization assets defines the material basis for class relations and exploitation.

This notion of organization assets bears a close relation to the problem of authority and hierarchy. The asset is organization. The activity of using that asset is coordinated decision making over a complex technical division of labor. When that asset is distributed unequally, so some positions have effective control over much more of the asset than others, then the social relation with respect to that asset takes the form of hierarchical authority. Authority, however, is not the asset as such; organization is the asset and is controlled through a hierarchy of authority.

The claim that effective control over organization assets is a basis of exploitation is equivalent to saying that nonmanagers would be better off and managers/bureaucrats worse off if nonmanagers were to withdraw with their per capita share of organization assets (or equivalently, if organizational control were democratized); and that by virtue of effectively controlling organization assets managers/bureaucrats control part or all of the socially produced surplus.¹⁷

A Typology of Class Structures, Assets, and Exploitation

If we add organization assets to the list in Roemer’s analysis, we generate the more complex typology presented in Table 1. Let us briefly look at each row of this table and examine its logic. Feudalism is a class system based on unequal distribution of ownership rights in labor power. What “personal bondage” means is that feudal lords have partial effective economic control over vassals. The empirical manifestation of this unequal distribution of ownership rights over labor power in classical feudalism is the coercive extraction of labor dues from serfs. When *corvée* labor is commuted to rents in kind and eventually money rents, the feudal character of the exploitation relation is reflected in legal prohibitions on the movement of peasants off the land. The “flight” of a peasant to the city is, in effect, a form of theft: the peasant is stealing part of the labor power owned by the

TABLE I
Assets, Exploitation, and Classes

| Type of class structure | Form of asset that is unequally distributed | Mechanism of exploitation | Classes | Central task of revolutionary transformation |
|------------------------------|---|--|-------------------------|--|
| Feudalism | Labor power | Coerced extraction | Lords and serfs | Individual liberty |
| Capitalism | Means of production | | Capitalists and workers | Socializing means of production |
| State bureaucratic socialism | Organizational assets | | Managers and workers | Democratization of organizational control |
| Socialism | Skills | Negotiated redistribution of surplus from workers to experts | Experts and workers | Substantive equality |

P/u from
Social Strat. 2e, p. 124

lord. Feudal lords may also have more means of production than serfs, more organizational assets, and more productive skills (although this is unlikely), and thus they may be exploiters with respect to these assets as well. What defines the society as “feudal,” however, is the primacy of the distinctively feudal mechanisms of exploitation. Accordingly, feudal class relations will be the primary structural basis of class struggle.

The bourgeois revolutions radically redistributed productive assets in people: everyone, at least in principle, owns one unit. This is what is meant by “bourgeois freedoms,” and in this sense capitalism can be regarded as an historically progressive force. But capitalism raises the second type of exploitation, exploitation based on property relations in means of production, to an unprecedented level.

The typical institutional form of capitalist class relations is capitalists having full ownership rights in the means of production and workers none. Other possibilities, however, have existed historically. Cottage industries in early capitalism involved workers owning some of their means of production, but not having sufficient assets to actually produce commodities without the assistance of merchant capitalists. Such workers were still being capitalistically exploited even though

there was no formal labor market with wages. In all capitalist exploitation, the mediating mechanism is market exchanges. Unlike in feudalism, surplus is not directly appropriated from workers in the form of coerced labor. Rather, it is appropriated through market exchanges: workers are paid a wage that covers the costs of production of their labor power; capitalists receive an income from the sale of the commodities produced by workers. The difference in these quantities constitutes the exploitative surplus appropriated by capitalists.

Anticapitalist revolutions attempt to eliminate the distinctively capitalist form of exploitation, exploitation based on private ownership of the means of production. The nationalization of the principal means of production is, in effect, a radical equalization of ownership of capital: everyone owns one citizen-share. Such revolutions, however, do not eliminate, and indeed may considerably strengthen and deepen, inequalities of effective control over organization assets. Whereas in capitalism the control over organization assets does not extend beyond the firm, in state bureaucratic socialism the coordinated integration of the division of labor extends to the whole society through institutions of central state planning. The mechanism by which

this generates exploitative transfers of surplus involves the centrally planned bureaucratic appropriation and distribution of the surplus along hierarchical principles. The corresponding class relation is therefore between managers/bureaucrats—people who control organization assets—and nonmanagers.

The historical task of revolutionary transformation of state bureaucratic socialism revolves around the equalization of effective economic control over organization assets, or, equivalently, the democratization of bureaucratic apparatuses of production. This does not imply total direct democracy, where all decisions of any consequence are directly made in democratic assemblies. There will still inevitably be delegated responsibilities, and there certainly can be representative forms of democratic control. But it does mean that the basic parameters of planning and coordinating social production are made through democratic mechanisms and that incumbency within delegated positions of responsibility does not give incumbents any personal claims on the social surplus. Such equalization, however, would not necessarily affect exploitation based on skills/credentials. Such exploitation would remain a central feature of socialism.

“Skill” in this context is not a trivial concept. The mere possession of enhanced laboring capabilities acquired through training is not sufficient to generate relations of exploitation, since the income of such trained labor may simply reflect the costs of acquiring the training. In such cases there is neither a transfer of surplus, nor would the untrained be better off under the game-theory specification of exploitation. For a skill to be the basis of exploitation, therefore, it has to be in some sense scarce relative to its demand, and there must be a mechanism through which individual owners of scarce skills are able to translate that scarcity into higher incomes.

There are basically three ways that skills can become scarce: first, they may require special *talents* that are naturally scarce in a

population; second, access to the training needed to develop the skill may be restricted through various mechanisms, creating an artificial scarcity of trained people; third, a certification system may be established that prohibits uncertified people from being employed to use the skill even if they have it. In all of these cases, the exploitation comes from the skilled/certified individual receiving an income that is above the costs of production of the skills by virtue of the scarcity of the availability of the skill.

In this conceptualization of socialism, a socialist society is essentially a kind of democratic technocracy. Experts control their own skills and knowledge within production, and by virtue of such control are able to appropriate some of the surplus out of production. However, because of the democratization of organization assets, actual planning decisions will not be made under the direct control of experts but will be made through some kind of democratic procedure (this is in effect what democratization of organization assets means: equalizing control over the planning and coordinating of social production). This means that the actual class power of a socialist technocratic exploiting class will be much weaker than the class power of exploiting classes in other class systems. Their ownership rights extend to only a limited part of the social surplus.

This much more limited basis of *domination* implied by skill-based exploitation is consistent with the spirit, if not the letter, of Marx’s claim that socialism is the “lower stage” of “communism,” since classes are already in a partial state of dissolution in a society with only skill-based exploitation. Communism itself, then, would be understood as a society within which skill-based exploitation itself had “withered away,” that is, in which ownership rights in skills had been equalized. This does not mean, it must be stressed, that all individuals would actually *possess* the same skills in communism, any more than eliminating property rights in

means of production implies that all individuals would actively use the same amount of physical capital. What is equalized is effective control over skills as a productive resource and claims to differential incomes resulting from differential use of skills.¹⁸ . . .

The Middle Classes and Contradictory Locations

The framework in Table 1 enables us to pose the problem of middle classes in a new way. Two different kinds of nonpolarized class locations can be defined in the logic of this framework:

1. There are class locations that are neither exploiters nor exploited, that is, people who have precisely the per capita level of the relevant asset. A petty bourgeois, self-employed producer with average capital stock, for example, would be neither exploiter nor exploited within capitalist relations. These kinds of positions are what can be called the "traditional" or "old"

middle class of a particular kind of class system.

2. Since concrete societies are rarely, if ever, characterized by a single mode of production, the actual class structures of given societies will be characterized by complex patterns of intersecting exploitation relations. There will therefore tend to be some positions that are exploiting along one dimension of exploitation relations and are exploited along another. Highly skilled wage-earners (for example, professionals) in capitalism are a good example: they are capitalistically exploited because they lack assets in capital, and yet they are skill exploiters. Such positions are what are typically referred to as the "new middle class" of a given system.

Table 2 presents a schematic typology of such complex class locations for capitalism. The typology is divided into two segments: one for owners of the means of production and one for nonowners. Within the wage-earner

TABLE 2
Basic Typology of Exploitation and Class

| Owners (%) | | Nonowners (wage-earners) (%) | | |
|----------------------|--|-----------------------------------|--------------------------|----------------------------|
| 1. Bourgeoisie | | Areas in the means of production | | |
| US 1.8 | | 4. Expert manager | 7. Semi-abstract manager | 10. Abstract labor manager |
| Sweden 3.7 | | US 3.9 | US 6.2 | US 2.3 |
| 2. Small employer | | P/u from Social Strat. 2e, p. 126 | | |
| US 6.0 | | 5. Petty bourgeois | | 6. Petty bourgeoisie |
| Sweden 4.8 | | US 6.9 | | US 39.9 |
| 3. Petty bourgeoisie | | 12. Educated worker | | |
| US 6.9 | | US 14 | US 12.2 | US 39.9 |
| Sweden 5.4 | | Sweden 6.7 | Sweden 17.8 | Sweden 43.5 |

United States, N = 1487
Sweden, N = 1179

Note: Excludes self-employed people working in the labor force, thus excluding essentially all nonwage-earner professionals, and those in Cooperative Production (COP) Structures and Class Intersections.

section of the typology, locations are distinguished by the two subordinate relations of exploitation characteristic of capitalist society—organization assets and skill/credential assets. It is thus possible within this framework to distinguish a whole terrain of class locations in capitalist *society* that are distinct from the polarized classes of the capitalist *mode of production*: expert managers, nonmanagerial experts, nonexpert managers, and so on.¹⁹

What is the relationship between this heterogeneous exploitation definition of the middle class and my previous conceptualization of such positions as contradictory locations within class relations? There is still a sense in which such positions could be characterized as “contradictory locations,” for they will typically hold contradictory interests with respect to the primary forms of class struggle in capitalist society, the struggle between labor and capital. On the one hand, they are like workers, in being excluded from ownership of the means of production. On the other hand, they have interests opposed to workers because of their effective control of organization and skill assets. Within the struggles of capitalism, therefore, these new middle classes do constitute contradictory locations, or more precisely, contradictory locations within exploitation relations.

This conceptualization of the middle classes also suggests that historically the principal forms of contradictory locations will vary depending upon the particular combinations of exploitation relations in a given society. These principal contradictory locations are presented in Table 3. In feudalism, the critical contradictory location is consti-

tuted by the bourgeoisie, the rising class of the successor mode of production. Within capitalism, the central contradictory location within exploitation relations is constituted by managers and state bureaucrats. They embody a principle of class organization that is quite distinct from capitalism and that potentially poses an alternative to capitalist relations. This is particularly true for state managers who, unlike corporate managers, are less likely to have their careers tightly integrated with the interests of the capitalist class. Finally, in state bureaucratic socialism, the “intelligentsia” broadly defined constitutes the pivotal contradictory location.

One of the upshots of this reconceptualization of the middle class is that it is no longer axiomatic that the proletariat is the unique, or perhaps even the central, rival to the capitalist class for class power in capitalist society. That classical Marxist assumption depended upon the thesis that there were no other classes within capitalism that could be viewed as the “bearers” of an historical alternative to capitalism. Socialism (as the transition to communism) was the only possible future for capitalism. What Table 3 suggests is that there are other class forces within capitalism that potentially pose an alternative to capitalism. This does not imply that there is any inevitability to the sequence feudalism—capitalism—state bureaucratic socialism—socialism—communism; state bureaucrats are not inevitably destined to be the future ruling class of present-day capitalisms. But it does suggest that the process of class formation and class struggle is considerably more complex and indeterminate than the traditional Marxist story has allowed.

TABLE 3
Basic Classes and Contradictory Locations
in Successive Modes of Production

| | | |
|--|---|--|
| | <p>P/u from Social Strat. 2e, p. 127</p> | |
|--|---|--|

NOTES

1. For a more detailed review of these alternatives, see E.O. Wright, “Varieties of Marxist Concepts of Class Structure,” *Politics and Society*, vol. 9, no. 3 (1980).
2. The leading proponent of the concept of the “new petty bourgeoisie” is N. Poulantzas, *Classes in Contemporary Capitalism* (London: Verso, 1975).

For the new-working-class concept, see S. Mallet, *La Nouvelle Classe Ouvrière* (Paris: Seuil, 1963).

3. B. Ehrenreich and J. Ehrenreich, "The Professional and Managerial Class," *Radical America*, vol. 11, no. 2 (1977).

4. A. Gouldner, *The Future of Intellectuals and the Rise of the New Class* (New York: Seabury Press, 1979); and G. Konrad and I. Szelenyi, *Intellectuals on the Road to Class Power* (New York: Harcourt, Brace, Jovanovich, 1979).

5. E.O. Wright, "Class Boundaries in Advanced Capitalist Societies," *New Left Review*, no. 98 (1976); and *Class, Crisis and the State* (London: New Left Books, 1978). See also G. Carchedi, *The Economic Identification of Social Classes* (London: Routledge and Kegan Paul, 1977).

6. E.O. Wright, "Capitalism's Futures," *Socialist Review*, no. 68 (1983).

7. A partial exception to this can be found in arguments for the existence of a "new class" of intellectuals and/or bureaucrats in capitalist and postcapitalist society. See: A. Gouldner, *The Future of Intellectuals*; and I. Szelenyi and W. Martin, *New Class Theory and Beyond* (unpublished book manuscript, Department of Sociology, University of Wisconsin, 1985).

8. R. Dahrendorf, *Class and Class Conflict in Industrial Society* (Palo Alto: Stanford University Press, 1959).

9. Roemer is a Marxist economist engaged in a long-term project of elaborating what he calls the "microfoundations" of Marxist theory. His most important work is entitled *A General Theory of Exploitation and Class* (Cambridge: Harvard University Press, 1982).

10. Roemer, *A General Theory*, pp. 194–95.

11. But note: workers in capitalism are *not* feudalistically exploited; they would be worse off, not better off, if they withdrew from the game of capitalism with only their personal assets. As Roemer argues, the claim by neoclassical theorists that wage earners in capitalism are not exploited is generally equivalent to the claim that they are not feudalistically exploited, that is, that they are not subjected to surplus extraction based on relations of personal bondage. See Roemer, *A General Theory*, p. 206.

12. The asset-exploitation nexus thus depends upon the capacity of asset-holders to deprive others of that asset. The social basis of exploitation, understood in this way, is quite similar to Frank Parkin's characterization of Weber's concept of social closure as "the process by which social collectivities seek to maximize rewards by restricting access to resources and opportunities to a limited circle of eligibles." F.

Parkin, *Marxism and Class Theory: A Bourgeois Critique* (New York: Columbia University Press, 1979). While Parkin's central concern is with the kinds of attributes that serve as the basis for closure—race, religion, language—Roemer's is with the nature of the resources (productive assets) over which closure is organized.

13. Roemer's conceptualization of the relationship between class and exploitation is similar in certain aspects to Alvin Gouldner's, although Roemer is unaware of Gouldner's work. Gouldner defines the "New Class" as a *cultural* bourgeoisie defined by its control over "cultural capital," where "capital" is defined as "any produced object used to make saleable utilities, thus providing its possessor with *incomes*, or claims to incomes defined as legitimate because of their imputed contribution to economic productivity." (*Future of Intellectuals*, p. 21). While Gouldner does not characterize this income allocation process in terms of exploitation, Roemer's exploitation concept would fit comfortably within Gouldner's general approach.

14. E.O. Wright, "The Status of the Political in the Concept of Class Structure," *Politics and Society*, vol. 11, no. 3 (1982).

15. See G.A. Cohen, *Karl Marx's Theory of History: A Defense* (Princeton: Princeton University Press, 1978), pp. 40–41, for a discussion of why labor power should be considered part of the forces of production (that is, a productive asset).

16. Roemer, *A General Theory*, p. 243.

17. This "control of the surplus," it must be noted, is *not* the equivalent of the *actual* personal consumption income of managers and bureaucrats, any more than capitalist profits or feudal rents are the equivalent of the personally consumed income of capitalists and feudal lords. It is historically variable both within and between types of societies what fraction of the surplus effectively controlled by exploiting classes is used for personal consumption and what portion is used for other purposes (feudal military expenditures, capitalist accumulation, organization growth). The claim that managers-bureaucrats would be "worse off" under conditions of a redistribution of organization assets refers to the amount of income they effectively control, which is therefore potentially available for personal appropriation, not simply the amount they personally consume.

18. It may be utopian to imagine a society without skill-based exploitation, or even a society without organization-asset exploitation, particularly if we reject the claim that a future society will ever exist in a state of absolute abundance. In the absence of absolute abundance, all societies will face dilemmas and trade-offs around the problem of distribution of

consumption, and such dilemmas may pose intractable incentive problems in the absence of exploitation. For a careful exposition of the problem of utopian fantasies in Marxist theory, see A. Nove, *The Economics of Feasible Socialism* (Hemel Hempstead: George Allen and Unwin, 1983).

19. The labor-force data in this table come from the comparative project on class structure and class consciousness, University of Wisconsin. Details of the coding of categories and the operationalization of variables can be found in E.O. Wright, *Classes* (London: Verso, 1985), appendix 2.

12 ■ Immanuel Wallerstein

Class Conflict in the Capitalist World Economy

What is capitalism as a mode of production? This is not an easy question, and for that reason is not in fact a widely discussed one. It seems to me that there are several elements that combine to constitute the ‘model’. Capitalism is the *only* mode of production in which the *maximization* of surplus creation is rewarded *per se*. In every historical system, there has been *some* production for *use*, and *some* production for *exchange*, but only in capitalism are all producers rewarded primarily in terms of the exchange value they produce and penalized to the extent they neglect it. The ‘rewards’ and ‘penalties’ are mediated through a structure called the ‘market’. It is a structure but not an institution. It is a structure molded by *many* institutions (political, economic, social, even cultural), and it is the principal arena of economic struggle.

Not only is surplus maximized for its own sake, but those who use the surplus to accumulate more capital to produce still more surplus are further rewarded. Thus the pressure is for constant expansion, although the individualistic premise of the system simultaneously renders *constant* expansion impossible.

How does the search for profit operate? It operates by creating legal protections for individual firms (which can range in size from individuals to quite large organizations, including parastatal agencies) to appropriate the surplus value created by the labor of the primary producers. Were all or most of this surplus value however consumed by the few who owned or controlled the ‘firms’, we would not have capitalism. This is in fact approximately what had happened in various pre-capitalist systems.

Capitalism involves in addition structures and institutions which reward primarily that subsegment of the owners and controllers who use the surplus value only *in part* for their own consumption, and in another (usually larger) part for further investment. The structure of the market ensures that those who do not accumulate capital (but merely consume surplus value) lose out economically over time to those who do accumulate capital.

We may thereupon designate as the bourgeoisie those who receive a part of the surplus value they do not themselves create and use some of it to accumulate capital. What

defines the bourgeois is not a particular profession and not even the legal status of proprietor (although this was historically important) but the fact that the bourgeois obtains, either as an individual or a member of some collectivity, a part of the surplus that he did not create and is in the position to invest (again either individually or as part of a collectivity) some of this surplus in capital goods.

There is a very large gamut of organizational arrangements which can permit this, of which the classic model of the 'free entrepreneur' is only one. Which organizational arrangements prevail at particular moments of time in particular states (for these arrangements are dependent on the legal framework) is a function of the state of development of the world-economy as a whole (and the role of a particular state in that world-economy) on the one hand, and the consequent forms of class struggle in the world-economy (and within the particular state) on the other. Hence, like all other social constructs, the 'bourgeoisie' is not a static phenomenon. It is the designation of a class in the process of perpetual re-creation and hence of constant change of form and composition. . . .

The fundamental role of the state as an institution in the capitalist world-economy is to augment the advantage of some against others in the market—that is, to *reduce* the 'freedom' of the market. Everyone is in favor of this, as long as one is the beneficiary of the 'distortion', and everyone opposed to the extent that one loses. It is all a matter of whose ox is being gored.

The modes of augmenting advantage are many. The state can transfer income by taking it from some and giving it to others. The state can restrict access to the market (of commodities or of labor) which favor those who thereby share in the oligopoly or oligopsony. The state can restrain persons from organizing to change the actions of the state. And, of course, the state can act

not only within its jurisdiction but beyond it. This may be licit (the rules concerning transit over boundaries) or illicit (interference in the internal affairs of another state). Warfare is of course one of the mechanisms used.

What is crucial to perceive is that the state is a special kind of organization. Its 'sovereignty', a notion of the modern world, is the claim to the monopolization (regulation) of the legitimate use of force within its boundaries, and it is in a relatively strong position to interfere effectively with the flow of factors of production. Obviously also it is possible for particular social groups to alter advantage by altering state boundaries; hence both movements for secession (or autonomy) and movements for annexation (or federation).

It is this realistic ability of states to interfere with the flow of factors of production that provides the political underpinnings of the structural division of labor in the capitalist world-economy as a whole. Normal market considerations may account for recurring initial thrusts to specialization (natural or socio-historical advantages in the production of one or another commodity), but it is the state system which encrusts, enforces, and exaggerates the patterns, and it has regularly required the use of state machinery to revise the pattern of the world-wide division of labor.

Furthermore, the ability of states to interfere with flows becomes differentiated. That is, core states become *stronger* than peripheral states, and use this differential power to maintain a differential degree of interstate freedom of flow. Specifically, core states have historically arranged that world-wide and over time, money and goods have flowed more 'freely' than labor. The reason for doing this is that core states have thereby received the advantages of 'unequal exchange'.

In effect, unequal exchange is simply a part of the world-wide process of the appropriation of surplus. We analyze falsely if we try to take literally the model of *one* proletarian re-

lating to *one* bourgeois. In fact, the surplus value that the producer creates passes through a series of persons and firms. It is therefore the case that *many* bourgeois *share* the surplus value of *one* proletarian. The exact share of different groups in the chain (property owner, merchants, intermediate consumers) is subject to much historical change and is itself a principal analytical variable in the functioning of the capitalist world-economy.

This chain of the transfer of surplus value frequently (often? almost always?) traverses national boundaries and, when it does, state operations intervene to tilt the sharing among bourgeois towards those bourgeois located in core states. This is unequal exchange, a mechanism in the overall process of the appropriation of surplus value.

One of the socio-geographic consequences of this system is the uneven distribution of the bourgeoisie and proletariat in different states, core states containing a higher percentage nationally of bourgeois than peripheral states. In addition, there are systematic differences in *kinds* of bourgeois and proletarians located in the two zones. For example, the percentage of wage-earning proletarians is systematically higher in core states.

Since states are the primary arena of political conflict in a capitalist world-economy, and since the functioning of the world-economy is such that national class composi-

tion varies widely, it is easy to perceive why the politics of states differentially located in relation to the world-economy should be so dissimilar. It is also then easy to perceive that using the political machinery of a given state to change the social composition and world-economic function of national production does not *per se* change the capitalist world-system as such.

Obviously, however, these various national thrusts to a change in structural position (which we misleadingly often call 'development') do in fact affect, indeed over the long run do in fact transform, the world-system. But they do so via the intervening variable of their impact on world-wide class consciousness of the proletariat.

Core and periphery then are simply phrases to locate one crucial part of the system of surplus appropriation by the bourgeoisie. To oversimplify, capitalism is a system in which the surplus value of the proletarian is appropriated by the bourgeois. When this proletarian is located in a different country from this bourgeois, one of the mechanisms that has affected the process of appropriation is the manipulation of controlling flows over state boundaries. This results in patterns of 'uneven development' which are *summarized* in the concepts of core, semiperiphery, and periphery. This is an intellectual tool to help analyze the multiple forms of class conflict in the capitalist world-economy.

■ WEBER AND POST-WEBERIANS ■

13 ■ Max Weber

Class, Status, Party

Economically Determined Power and the Social Order

Law exists when there is a probability that an order will be upheld by a specific staff of men who will use physical or psychical compulsion with the intention of obtaining conformity with the order, or of inflicting sanctions for infringement of it.¹ The structure of every legal order directly influences the distribution of power, economic or otherwise, within its respective community. This is true of all legal orders and not only that of the state. In general, we understand by ‘power’ the chance of a man or of a number of men to realize their own will in a communal action even against the resistance of others who are participating in the action.

‘Economically conditioned’ power is not, of course, identical with ‘power’ as such. On the contrary, the emergence of economic power may be the consequence of power existing on other grounds. Man does not strive for power only in order to enrich himself economically. Power, including economic power, may be valued ‘for its own sake.’ Very frequently the striving for power is also conditioned by the social ‘honor’ it entails. Not all power, however, entails social honor: The typical American Boss, as well as the typical

big speculator, deliberately relinquishes social honor. Quite generally, ‘mere economic’ power, and especially ‘naked’ money power, is by no means a recognized basis of social honor. Nor is power the only basis of social honor. Indeed, social honor, or prestige, may even be the basis of political or economic power, and very frequently has been. Power, as well as honor, may be guaranteed by the legal order, but, at least normally, it is not their primary source. The legal order is rather an additional factor that enhances the chance to hold power or honor; but it cannot always secure them.

The way in which social honor is distributed in a community between typical groups participating in this distribution we may call the ‘social order.’ The social order and the economic order are, of course, similarly related to the ‘legal order.’ However, the social and the economic order are not identical. The economic order is for us merely the way in which economic goods and services are distributed and used. The social order is of course conditioned by the economic order to a high degree, and in its turn reacts upon it.

Now: ‘classes,’ ‘status groups,’ and ‘parties’ are phenomena of the distribution of power within a community.

Max Weber. “Class, Status, Party,” in *From Max Weber: Essays in Sociology*, edited by H. H. Gerth and C. Wright Mills, and translated by H. H. Gerth and C. Wright Mills, pp. 180–195. Translation copyright © 1946, 1958 by H. H. Gerth and C. Wright Mills. Used by permission of Oxford University Press, Inc.

Determination of Class-Situation by Market-Situation

In our terminology, 'classes' are not communities; they merely represent possible, and frequent, bases for communal action. We may speak of a 'class' when (1) a number of people have in common a specific causal component of their life chances, in so far as (2) this component is represented exclusively by economic interests in the possession of goods and opportunities for income, and (3) is represented under the conditions of the commodity or labor markets. [These points refer to 'class situation,' which we may express more briefly as the typical chance for a supply of goods, external living conditions, and personal life experiences, in so far as this chance is determined by the amount and kind of power, or lack of such, to dispose of goods or skills for the sake of income in a given economic order. The term 'class' refers to any group of people that is found in the same class situation.]

It is the most elemental economic fact that the way in which the disposition over material property is distributed among a plurality of people, meeting competitively in the market for the purpose of exchange, in itself creates specific life chances. According to the law of marginal utility this mode of distribution excludes the non-owners from competing for highly valued goods; it favors the owners and, in fact, gives to them a monopoly to acquire such goods. Other things being equal, this mode of distribution monopolizes the opportunities for profitable deals for all those who, provided with goods, do not necessarily have to exchange them. It increases, at least generally, their power in price wars with those who, being propertyless, have nothing to offer but their services in native form or goods in a form constituted through their own labor, and who above all are compelled to get rid of these products in order barely to subsist. This mode of distribution gives to the propertied a monopoly on the possibility of transferring property from the sphere of use as a 'fortune,' to the

sphere of 'capital goods'; that is, it gives them the entrepreneurial function and all chances to share directly or indirectly in returns on capital. All this holds true within the area in which pure market conditions prevail. 'Property' and 'lack of property' are, therefore, the basic categories of all class situations. It does not matter whether these two categories become effective in price wars or in competitive struggles.

Within these categories, however, class situations are further differentiated: on the one hand, according to the kind of property that is usable for returns; and, on the other hand, according to the kind of services that can be offered in the market. Ownership of domestic buildings; productive establishments; warehouses; stores; agriculturally usable land, large and small holdings—quantitative differences with possibly qualitative consequences; ownership of mines; cattle; men (slaves); disposition over mobile instruments of production, or capital goods of all sorts, especially money or objects that can be exchanged for money easily and at any time; disposition over products of one's own labor or of others' labor differing according to their various distances from consumability; disposition over transferable monopolies of any kind—all these distinctions differentiate the class situations of the propertied just as does the 'meaning' which they can and do give to the utilization of property, especially to property which has money equivalence. Accordingly, the propertied, for instance, may belong to the class of rentiers or to the class of entrepreneurs.

Those who have no property but who offer services are differentiated just as much according to their kinds of services as according to the way in which they make use of these services, in a continuous or discontinuous relation to a recipient. But always this is the generic connotation of the concept of class: that the kind of chance in the *market* is the decisive moment which presents a common condition for the individual's fate. 'Class situation' is, in this sense, ultimately 'market situation.' The effect of naked possession *per se*, which

among cattle breeders gives the nonowning slave or serf into the power of the cattle owner, is only a forerunner of real 'class' formation. However, in the cattle loan and in the naked severity of the law of debts in such communities, for the first time mere 'possession' as such emerges as decisive for the fate of the individual. This is very much in contrast to the agricultural communities based on labor. The creditor-debtor relation becomes the basis of 'class situations' only in those cities where a 'credit market,' however primitive, with rates of interest increasing according to the extent of dearth and a factual monopolization of credits, is developed by a plutocracy. Therewith 'class struggles' begin.

Those men whose fate is not determined by the chance of using goods or services for themselves on the market, e.g. slaves, are not, however, a 'class' in the technical sense of the term. They are, rather, a 'status group.'

Communal Action Flowing from Class Interest

According to our terminology, the factor that creates 'class' is unambiguously economic interest, and indeed, only those interests involved in the existence of the 'market.' Nevertheless, the concept of 'class-interest' is an ambiguous one: even as an empirical concept it is ambiguous as soon as one understands by it something other than the factual direction of interests following with a certain probability from the class situation for a certain 'average' of those people subjected to the class situation. The class situation and other circumstances remaining the same, the direction in which the individual worker, for instance, is likely to pursue his interests may vary widely, according to whether he is constitutionally qualified for the task at hand to a high, to an average, or to a low degree. In the same way, the direction of interests may vary according to whether or not a *communal* action of a larger or smaller portion of those commonly affected by the 'class situation,' or even an association among them,

e.g. a 'trade union,' has grown out of the class situation from which the individual may or may not expect promising results. [Communal action refers to that action which is oriented to the feeling of the actors that they belong together. Societal action, on the other hand, is oriented to a rationally motivated adjustment of interests.] The rise of societal or even of communal action from a common class situation is by no means a universal phenomenon.

The class situation may be restricted in its effects to the generation of essentially *similar* reactions, that is to say, within our terminology, of 'mass actions.' However, it may not have even this result. Furthermore, often merely an amorphous communal action emerges. For example, the 'murmuring' of the workers known in ancient oriental ethics: the moral disapproval of the work-master's conduct, which in its practical significance was probably equivalent to an increasingly typical phenomenon of precisely the latest industrial development, namely, the 'slow down' (the deliberate limiting of work effort) of laborers by virtue of tacit agreement. The degree in which 'communal action' and possibly 'societal action,' emerges from the 'mass actions' of the members of a class is linked to general cultural conditions, especially to those of an intellectual sort. It is also linked to the extent of the contrasts that have already evolved, and is especially linked to the *transparency* of the connections between the causes and the consequences of the 'class situation.' For however different life chances may be, this fact in itself, according to all experience, by no means gives birth to 'class action' (communal action by the members of a class). The fact of being conditioned and the results of the class situation must be distinctly recognizable. For only then the contrast of life chances can be felt not as an absolutely given fact to be accepted, but as a resultant from either (1) the given distribution of property, or (2) the structure of the concrete economic order. It is only then that people may react against the class structure not only through acts of an intermittent and irrational

protest, but in the form of rational association. There have been ‘class situations’ of the first category (1), of a specifically naked and transparent sort, in the urban centers of Antiquity and during the Middle Ages; especially then, when great fortunes were accumulated by factually monopolized trading in industrial products of these localities or in foodstuffs. Furthermore, under certain circumstances, in the rural economy of the most diverse periods, when agriculture was increasingly exploited in a profit-making manner. The most important historical example of the second category (2) is the class situation of the modern ‘proletariat.’

Types of ‘Class Struggle’

Thus every class may be the carrier of any one of the possibly innumerable forms of ‘class action,’ but this is not necessarily so: In any case, a class does not in itself constitute a community. To treat ‘class’ conceptually as having the same value as ‘community’ leads to distortion. That men in the same class situation regularly react in mass actions to such tangible situations as economic ones in the direction of those interests that are most adequate to their average number is an important and after all simple fact for the understanding of historical events. Above all, this fact must not lead to that kind of pseudo-scientific operation with the concepts of ‘class’ and ‘class interests’ so frequently found these days, and which has found its most classic expression in the statement of a talented author, that the individual may be in error concerning his interests but that the ‘class’ is ‘infallible’ about its interests. Yet, if classes as such are not communities, nevertheless class situations emerge only on the basis of communalization. The communal action that brings forth class situations, however, is not basically action between members of the identical class; it is an action between members of different classes. Communal actions that directly determine the class situation of

the worker and the entrepreneur are: the labor market, the commodities market, and the capitalistic enterprise. But, in its turn, the existence of a capitalistic enterprise presupposes that a very specific communal action exists and that it is specifically structured to protect the possession of goods *per se*, and especially the power of individuals to dispose, in principle freely, over the means of production. The existence of a capitalistic enterprise is preconditioned by a specific kind of ‘legal order.’ Each kind of class situation, and above all when it rests upon the power of property *per se*, will become most clearly efficacious when all other determinants of reciprocal relations are, as far as possible, eliminated in their significance. It is in this way that the utilization of the power of property in the market obtains its most sovereign importance.

Now ‘status groups’ hinder the strict carrying through of the sheer market principle. In the present context they are of interest to us only from this one point of view. Before we briefly consider them, note that not much of a general nature can be said about the more specific kinds of antagonism between ‘classes’ (in our meaning of the term). The great shift, which has been going on continuously in the past, and up to our times, may be summarized, although at the cost of some precision: the struggle in which class situations are effective has progressively shifted from consumption credit toward, first, competitive struggles in the commodity market and, then, toward price wars on the labor market. The ‘class struggles’ of antiquity—to the extent that they were genuine class struggles and not struggles between status groups—were initially carried on by indebted peasants, and perhaps also by artisans threatened by debt bondage and struggling against urban creditors. For debt bondage is the normal result of the differentiation of wealth in commercial cities, especially in seaport cities. A similar situation has existed among cattle breeders. Debt relationships as such produced

class action up to the time of Cataline. Along with this, and with an increase in provision of grain for the city by transporting it from the outside, the struggle over the means of sustenance emerged. It centered in the first place around the provision of bread and the determination of the price of bread. It lasted throughout antiquity and the entire Middle Ages. The propertyless as such flocked together against those who actually and supposedly were interested in the dearth of bread. This fight spread until it involved all those commodities essential to the way of life and to handicraft production. There were only incipient discussions of wage disputes in antiquity and in the Middle Ages. But they have been slowly increasing up into modern times. In the earlier periods they were completely secondary to slave rebellions as well as to fights in the commodity market.

The propertyless of antiquity and of the Middle Ages protested against monopolies, pre-emption, forestalling, and the withholding of goods from the market in order to raise prices. Today the central issue is the determination of the price of labor.

This transition is represented by the fight for access to the market and for the determination of the price of products. Such fights went on between merchants and workers in the putting-out system of domestic handicraft during the transition to modern times. Since it is quite a general phenomenon we must mention here that the class antagonisms that are conditioned through the market situation are usually most bitter between those who actually and directly participate as opponents in price wars. It is not the rentier, the share-holder, and the banker who suffer the ill will of the worker, but almost exclusively the manufacturer and the business executives who are the direct opponents of workers in price wars. This is so in spite of the fact that it is precisely the cash boxes of the rentier, the share-holder, and the banker into which the more or less 'unearned' gains flow, rather than into the pockets of the

manufacturers or of the business executives. This simple state of affairs has very frequently been decisive for the role the class situation has played in the formation of political parties. For example, it has made possible the varieties of patriarchal socialism and the frequent attempts—formerly, at least—of threatened status groups to form alliances with the proletariat against the 'bourgeoisie.'

Status Honor

In contrast to classes, *status groups* are normally communities. They are, however, often of an amorphous kind. In contrast to the purely economically determined 'class situation' we wish to designate as 'status situation' every typical component of the life fate of men that is determined by a specific, positive or negative, social estimation of *honor*. This honor may be connected with any quality shared by a plurality, and, of course, it can be knit to a class situation: class distinctions are linked in the most varied ways with status distinctions. Property as such is not always recognized as a status qualification, but in the long run it is, and with extraordinary regularity. In the subsistence economy of the organized neighborhood, very often the richest man is simply the chieftain. However, this often means only an honorific preference. For example, in the so-called pure modern 'democracy,' that is, one devoid of any expressly ordered status privileges for individuals, it may be that only the families coming under approximately the same tax class dance with one another. This example is reported of certain smaller Swiss cities. But status honor need not necessarily be linked with a 'class situation.' On the contrary, it normally stands in sharp opposition to the pretensions of sheer property.

Both propertied and propertyless people can belong to the same status group, and frequently they do with very tangible consequences. This 'equality' of social esteem may, however, in the long run become quite pre-

carious. The ‘equality’ of status among the American ‘gentlemen,’ for instance, is expressed by the fact that outside the subordination determined by the different functions of ‘business,’ it would be considered strictly repugnant—wherever the old tradition still prevails—if even the richest ‘chief,’ while playing billiards or cards in his club in the evening, would not treat his ‘clerk’ as in every sense fully his equal in birthright. It would be repugnant if the American ‘chief’ would bestow upon his ‘clerk’ the condescending ‘benevolence’ marking a distinction of ‘position,’ which the German chief can never dis sever from his attitude. This is one of the most important reasons why in America the German ‘clubby-ness’ has never been able to attain the attraction that the American clubs have.

Guarantees of Status Stratification

In content, status honor is normally expressed by the fact that above all else a specific *style of life* can be expected from all those who wish to belong to the circle. Linked with this expectation are restrictions on ‘social’ intercourse (that is, intercourse which is not subservient to economic or any other of business’s ‘functional’ purposes). These restrictions may confine normal marriages to within the status circle and may lead to complete endogamous closure. As soon as there is not a mere individual and socially irrelevant imitation of another style of life, but an agreed-upon communal action of this closing character, the ‘status’ development is under way.

In its characteristic form, stratification by ‘status groups’ on the basis of conventional styles of life evolves at the present time in the United States out of the traditional democracy. For example, only the resident of a certain street (‘the street’) is considered as belonging to ‘society,’ is qualified for social intercourse, and is visited and invited. Above all, this differentiation evolves in such

a way as to make for strict submission to the fashion that is dominant at a given time in society. This submission to fashion also exists among men in America to a degree unknown in Germany. Such submission is considered to be an indication of the fact that a given man *pretends* to qualify as a gentleman. This submission decides, at least *prima facie*, that he will be treated as such. And this recognition becomes just as important for his employment chances in ‘swank’ establishments, and above all, for social intercourse and marriage with ‘esteemed’ families, as the qualification for dueling among Germans in the Kaiser’s day. As for the rest: certain families resident for a long time, and, of course, correspondingly wealthy, e.g. ‘F. F. V., i.e. First Families of Virginia,’ or the actual or alleged descendants of the ‘Indian Princess’ Pocahontas, of the Pilgrim fathers, or of the Knickerbockers, the members of almost inaccessible sects and all sorts of circles setting themselves apart by means of any other characteristics and badges . . . all these elements usurp ‘status’ honor. The development of status is essentially a question of stratification resting upon usurpation. Such usurpation is the normal origin of almost all status honor. But the road from this purely conventional situation to legal privilege, positive or negative, is easily traveled as soon as a certain stratification of the social order has in fact been ‘lived in’ and has achieved stability by virtue of a stable distribution of economic power.

‘Ethnic’ Segregation and ‘Caste’

Where the consequences have been realized to their full extent, the status group evolves into a closed ‘caste.’ Status distinctions are then guaranteed not merely by conventions and laws, but also by *rituals*. This occurs in such a way that every physical contact with a member of any caste that is considered to be ‘lower’ by the members of a ‘higher’ caste is considered as making for a ritualistic impurity

and to be a stigma which must be expiated by a religious act. Individual castes develop quite distinct cults and gods.

In general, however, the status structure reaches such extreme consequences only where there are underlying differences which are held to be 'ethnic.' The 'caste' is, indeed, the normal form in which ethnic communities usually live side by side in a 'societalized' manner. These ethnic communities believe in blood relationship and exclude exogamous marriage and social intercourse. Such a caste situation is part of the phenomenon of 'pariah' peoples and is found all over the world. These people form communities, acquire specific occupational traditions of handicrafts or of other arts, and cultivate a belief in their ethnic community. They live in a 'diaspora' strictly segregated from all personal intercourse, except that of an unavoidable sort, and their situation is legally precarious. Yet, by virtue of their economic indispensability, they are tolerated, indeed, frequently privileged, and they live in interspersed political communities. The Jews are the most impressive historical example.

A 'status' segregation grown into a 'caste' differs in its structure from a mere 'ethnic' segregation: the caste structure transforms the horizontal and unconnected coexistences of ethnically segregated groups into a vertical social system of super- and subordination. Correctly formulated: a comprehensive societalization integrates the ethnically divided communities into specific political and communal action. In their consequences they differ precisely in this way: ethnic coexistences condition a mutual repulsion and disdain but allow each ethnic community to consider its own honor as the highest one; the caste structure brings about a social subordination and an acknowledgment of 'more honor' in favor of the privileged caste and status groups. This is due to the fact that in the caste structure ethnic distinctions as such have become 'functional' distinctions within the political societalization (warriors, priests,

artisans that are politically important for war and for building, and so on). But even pariah people who are most despised are usually apt to continue cultivating in some manner that which is equally peculiar to ethnic and to status communities: the belief in their own specific 'honor.' This is the case with the Jews.

Only with the negatively privileged status groups does the 'sense of dignity' take a specific deviation. A sense of dignity is the precipitation in individuals of social honor and of conventional demands which a positively privileged status group raises for the deportment of its members. The sense of dignity that characterizes positively privileged status groups is naturally related to their 'being' which does not transcend itself, that is, it is to their 'beauty and excellence.' Their kingdom is 'of this world.' They live for the present and by exploiting their great past. The sense of dignity of the negatively privileged strata naturally refers to a future lying beyond the present, whether it is of this life or of another. In other words, it must be nurtured by the belief in a providential 'mission' and by a belief in a specific honor before God. The 'chosen people's' dignity is nurtured by a belief either that in the beyond 'the last will be the first,' or that in this life a Messiah will appear to bring forth into the light of the world which has cast them out the hidden honor of the pariah people. This simple state of affairs, and not the 'resentment' which is so strongly emphasized in Nietzsche's much admired construction in the *Genealogy of Morals*, is the source of the religiosity cultivated by pariah status groups. In passing, we may note that resentment may be accurately applied only to a limited extent; for one of Nietzsche's main examples, Buddhism, it is not at all applicable.

Incidentally, the development of status groups from ethnic segregations is by no means the normal phenomenon. On the contrary, since objective 'racial differences' are by no means basic to every subjective sentiment of an ethnic community, the ulti-

mately racial foundation of status structure is rightly and absolutely a question of the concrete individual case. Very frequently a status group is instrumental in the production of a thoroughbred anthropological type. Certainly a status group is to a high degree effective in producing extreme types, for they select personally qualified individuals (e.g. the Knighthood selects those who are fit for warfare, physically and psychically). But selection is far from being the only, or the predominant, way in which status groups are formed: Political membership or class situation has at all times been at least as frequently decisive. And today the class situation is by far the predominant factor, for of course the possibility of a style of life expected for members of a status group is usually conditioned economically.

Status Privileges

For all practical purposes, stratification by status goes hand in hand with a monopolization of ideal and material goods or opportunities, in a manner we have come to know as typical. Besides the specific status honor, which always rests upon distance and exclusiveness, we find all sorts of material monopolies. Such honorific preferences may consist of the privilege of wearing special costumes, of eating special dishes taboo to others, of carrying arms—which is most obvious in its consequences—the right to pursue certain non-professional dilettante artistic practices, e.g. to play certain musical instruments. Of course, material monopolies provide the most effective motives for the exclusiveness of a status group; although, in themselves, they are rarely sufficient, almost always they come into play to some extent. Within a status circle there is the question of intermarriage: the interest of the families in the monopolization of potential bridegrooms is at least of equal importance and is parallel to the interest in the monopolization of daughters. The daughters of the circle must be provided for. With

an increased inclosure of the status group, the conventional preferential opportunities for special employment grow into a legal monopoly of special offices for the members. Certain goods become objects for monopolization by status groups. In the typical fashion these include ‘entailed estates’ and frequently also the possessions of serfs or bondsmen and, finally, special trades. This monopolization occurs positively when the status group is exclusively entitled to own and to manage them; and negatively when, in order to maintain its specific way of life, the status group must *not* own and manage them.

The decisive role of a ‘style of life’ in status ‘honor’ means that status groups are the specific bearers of all ‘conventions.’ In whatever way it may be manifest, all ‘stylization’ of life either originates in status groups or is at least conserved by them. Even if the principles of status conventions differ greatly, they reveal certain typical traits, especially among those strata which are most privileged. Quite generally, among privileged status groups there is a status disqualification that operates against the performance of common physical labor. This disqualification is now ‘setting in’ in America against the old tradition of esteem for labor. Very frequently every rational economic pursuit, and especially ‘entrepreneurial activity,’ is looked upon as a disqualification of status. Artistic and literary activity is also considered as degrading work as soon as it is exploited for income, or at least when it is connected with hard physical exertion. An example is the sculptor working like a mason in his dusty smock as over against the painter in his salon-like ‘studio’ and those forms of musical practice that are acceptable to the status group.

Economic Conditions and Effects of Status Stratification

The frequent disqualification of the gainfully employed as such is a direct result of the principle of status stratification peculiar to the

social order, and of course, of this principle's opposition to a distribution of power which is regulated exclusively through the market. These two factors operate along with various individual ones, which will be touched upon below.

We have seen above that the market and its processes 'knows no personal distinctions': 'functional' interests dominate it. It knows nothing of 'honor.' The status order means precisely the reverse, viz.: stratification in terms of 'honor' and of styles of life peculiar to status groups as such. If mere economic acquisition and naked economic power still bearing the stigma of its extra-status origin could bestow upon anyone who has won it the same honor as those who are interested in status by virtue of style of life claim for themselves, the status order would be threatened at its very root. This is the more so as, given equality of status honor, property *per se* represents an addition even if it is not overtly acknowledged to be such. Yet if such economic acquisition and power gave the agent any honor at all, his wealth would result in his attaining more honor than those who successfully claim honor by virtue of style of life. Therefore all groups having interests in the status order react with special sharpness precisely against the pretensions of purely economic acquisition. In most cases they react the more vigorously the more they feel themselves threatened. Calderon's respectful treatment of the peasant, for instance, as opposed to Shakespeare's simultaneous and ostensible disdain of the *canaille* illustrates the different way in which a firmly structured status order reacts as compared with a status order that has become economically precarious. This is an example of a state of affairs that recurs everywhere. Precisely because of the rigorous reactions against the claims of property *per se*, the 'parvenu' is never accepted, personally and without reservation, by the privileged status groups, no matter how completely his style of life

has been adjusted to theirs. They will only accept his descendants who have been educated in the conventions of their status group and who have never besmirched its honor by their own economic labor.

As to the general *effect* of the status order, only one consequence can be stated, but it is a very important one: the hindrance of the free development of the market occurs first for those goods which status groups directly withheld from free exchange by monopolization. This monopolization may be effected either legally or conventionally. For example, in many Hellenic cities during the epoch of status groups, and also originally in Rome, the inherited estate (as is shown by the old formula for indication against spendthrifts) was monopolized just as were the estates of knights, peasants, priests, and especially the clientele of the craft and merchant guilds. The market is restricted, and the power of naked property *per se*, which gives its stamp to 'class formation,' is pushed into the background. The results of this process can be most varied. Of course, they do not necessarily weaken the contrasts in the economic situation. Frequently they strengthen these contrasts, and in any case, where stratification by status permeates a community as strongly as was the case in all political communities of antiquity and of the Middle Ages, one can never speak of a genuinely free market competition as we understand it today. There are wider effects than this direct exclusion of special goods from the market. From the contrariety between the status order and the purely economic order mentioned above, it follows that in most instances the notion of honor peculiar to status absolutely abhors that which is essential to the market: higgling. Honor abhors higgling among peers and occasionally it taboos higgling for the members of a status group in general. Therefore, everywhere some status groups, and usually the most influential, consider almost any kind of overt participation in economic acquisition as absolutely stigmatizing.

With some over-simplification, one might thus say that ‘classes’ are stratified according to their relations to the production and acquisition of goods; whereas ‘status groups’ are stratified according to the principles of their *consumption* of goods as represented by special ‘styles of life.’

An ‘occupational group’ is also a status group. For normally, it successfully claims social honor only by virtue of the special style of life which may be determined by it. The differences between classes and status groups frequently overlap. It is precisely those status communities most strictly segregated in terms of honor (*viz.* the Indian castes) who today show, although within very rigid limits, a relatively high degree of indifference to pecuniary income. However, the Brahmins seek such income in many different ways.

As to the general economic conditions making for the predominance of stratification by ‘status,’ only very little can be said. When the bases of the acquisition and distribution of goods are relatively stable, stratification by status is favored. Every technological repercussion and economic transformation threatens stratification by status and pushes the class situation into the foreground. Epochs and countries in which the naked class situation is of predominant significance are regularly the periods of technical and economic transformations. And every slowing down of the shifting of economic stratifications leads, in due course, to the growth of status structures and makes for a resuscitation of the important role of social honor.

Parties

Whereas the genuine place of ‘classes’ is within the economic order, the place of ‘status groups’ is within the social order, that is, within the sphere of the distribution of ‘honor.’ From within these spheres, classes and status groups influence one another and they influence the legal order and are in turn influenced by it. But ‘parties’ live in a house of ‘power.’

Their action is oriented toward the acquisition of social ‘power,’ that is to say, toward influencing a communal action no matter what its content may be. In principle, parties may exist in a social ‘club’ as well as in a ‘state.’ As over against the actions of classes and status groups, for which this is not necessarily the case, the communal actions of ‘parties’ always mean a societalization. For party actions are always directed toward a goal which is striven for in planned manner. This goal may be a ‘cause’ (the party may aim at realizing a program for ideal or material purposes), or the goal may be ‘personal’ (sinecures, power, and from these, honor for the leader and the followers of the party). Usually the party action aims at all these simultaneously. Parties are, therefore, only possible within communities that are societalized, that is, which have some rational order and a staff of persons available who are ready to enforce it. For parties aim precisely at influencing this staff, and if possible, to recruit it from party followers.

In any individual case, parties may represent interests determined through ‘class situation’ or ‘status situation,’ and they may recruit their following respectively from one or the other. But they need be neither purely ‘class’ nor purely ‘status’ parties. In most cases they are partly class parties and partly status parties, but sometimes they are neither. They may represent ephemeral or enduring structures. Their means of attaining power may be quite varied, ranging from naked violence of any sort to canvassing for votes with coarse or subtle means: money, social influence, the force of speech, suggestion, clumsy hoax, and so on to the rougher or more artful tactics of obstruction in parliamentary bodies.

The sociological structure of parties differs in a basic way according to the kind of communal action which they struggle to influence. Parties also differ according to whether or not the community is stratified by status or by classes. Above all else, they vary according

to the structure of domination within the community. For their leaders normally deal with the conquest of a community. They are, in the general concept which is maintained here, not only products of specially modern forms of domination. We shall also designate as parties the ancient and medieval ‘parties,’ despite the fact that their structure differs basically from the structure of modern parties. By virtue of these structural differences of domination it is impossible to say anything about the structure of parties without discussing the structural forms of social domination *per se*. Parties, which are always structures struggling for domination, are very frequently organized in a very strict ‘authoritarian’ fashion. . . .

Concerning ‘classes,’ ‘status groups,’ and ‘parties,’ it must be said in general that they necessarily presuppose a comprehensive societalization, and especially a political framework of communal action, within which they operate. This does not mean that parties would be confined by the frontiers of any individual political community. On the contrary, at all times it has been the order of the day that the societalization (even when it aims at the use of military force in common)

reaches beyond the frontiers of politics. This has been the case in the solidarity of interests among the Oligarchs and among the democrats in Hellas, among the Guelfs and among Ghibellines in the Middle Ages, and within the Calvinist party during the period of religious struggles. It has been the case up to the solidarity of the landlords (international congress of agrarian landlords), and has continued among princes (holy alliance, Karlsbad decrees), socialist workers, conservatives (the longing of Prussian conservatives for Russian intervention in 1850). But their aim is not necessarily the establishment of new international political, i.e. *territorial*, dominion. In the main they aim to influence the existing dominion.²

NOTES

1. *Wirtschaft und Gesellschaft*, part III, chap. 4, pp. 631–40. The first sentence in paragraph one and the several definitions in this chapter which are in brackets do not appear in the original text. They have been taken from other contexts of *Wirtschaft und Gesellschaft*.

2. The posthumously published text breaks off here. We omit an incomplete sketch of types of ‘warrior estates.’

■ Max Weber

Status Groups and Classes

The Concepts of Class and Class Situation

The term ‘class situation’¹ will be applied to the typical probability that a given state of (a) provision with goods, (b) external condi-

tions of life, and (c) subjective satisfaction or frustration will be possessed by an individual or a group. These probabilities define class situation in so far as they are dependent on the kind and extent of control or lack of it

Max Weber. “Status Groups and Classes,” in *The Theory of Social and Economic Organization*, translated by A. M. Henderson and Talcott Parsons, pp. 424–429. Copyright © 1947 by Talcott Parsons; copyright renewed in 1975 by Talcott Parsons. Reprinted with the permission of the Free Press, a Division of Simon and Schuster, Inc.

which the individual has over goods or services and existing possibilities of their exploitation for the attainment of income or receipts within a given economic order.

A 'class' is any group of persons occupying the same class situation. The following types of classes may be distinguished: (a) A class is a 'property class' when class situation for its members is primarily determined by the differentiation of property holdings; (b) a class is an 'acquisition class' when the class situation of its members is primarily determined by their opportunity for the exploitation of services on the market; (c) the 'social class' structure is composed of the plurality of class situations between which an interchange of individuals on a personal basis or in the course of generations is readily possible and typically observable. On the basis of any of the three types of class situation, associative relationships between those sharing the same class interests, namely, corporate class organizations may develop. This need not, however, necessarily happen. The concepts of class and class situation as such designate only the fact of identity or similarity in the typical situation in which a given individual and many others find their interests defined. In principle control over different combinations of consumer goods, means of production, investments, capital funds or marketable abilities constitute class situations which are different with each variation and combination. Only persons who are completely unskilled, without property and dependent on employment without regular occupation, are in a strictly identical class situation. Transitions from one class situation to another vary greatly in fluidity and in the ease with which an individual can enter the class. Hence the unity of 'social' classes is highly relative and variable.

The Significance of Property Classes

The primary significance of a positively privileged property class lies in the following facts: (i) Its members may be able to monopolize

the purchase of high-priced consumer goods. (ii) They may control the opportunities of pursuing a systematic monopoly policy in the sale of economic goods. (iii) They may monopolize opportunities for the accumulation of property through unconsumed surpluses. (iv) They may monopolize opportunities to accumulate capital by saving, hence, the possibility of investing property in loans and the related possibility of control over executive positions in business. (v) They may monopolize the privileges of socially advantageous kinds of education so far as these involve expenditures.

Positively privileged property classes typically live from property income. This may be derived from property rights in human beings, as with slaveowners, in land, in mining property, in fixed equipment such as plant and apparatus, in ships, and as creditors in loan relationships. Loans may consist of domestic animals, grain, or money. Finally they may live on income from securities.

Class interests which are negatively privileged with respect to property belong typically to one of the following types: (a) They are themselves objects of ownership, that is they are unfree. (b) They are 'outcasts,' that is 'proletarians' in the sense meant in Antiquity. (c) They are debtor classes and, (d) the 'poor.'

In between stand the 'middle' classes. This term includes groups who have all sorts of property, or of marketable abilities through training, who are in a position to draw their support from these sources. Some of them may be 'acquisition' classes. Entrepreneurs are in this category by virtue of essentially positive privileges; proletarians, by virtue of negative privileges. But many types such as peasants, craftsmen, and officials do not fall in this category.

The differentiation of classes on the basis of property alone is not 'dynamic,' that is, it does not necessarily result in class struggles or class revolutions. It is not uncommon for very strongly privileged property classes, such as slaveowners, to exist side by side with such far less privileged groups as peasants or even

outcasts without any class struggle. There may even be ties of solidarity between privileged property classes and unfree elements. However, such conflicts as that between land owners and outcast elements or between creditors and debtors, the latter often being a question of urban patricians as opposed to either rural peasants or urban craftsmen, may lead to revolutionary conflict. Even this, however, need not necessarily aim at radical changes in economic organization. It may, on the contrary, be concerned in the first instance only with a redistribution of wealth. These may be called 'property revolutions.'

A classic example of the lack of class antagonism has been the relation of the 'poor white trash,' originally those not owning slaves, to the planters in the Southern States of the United States. The 'poor whites' have often been much more hostile to the Negro than the planters who have frequently had a large element of patriarchal sentiment. The conflict of outcast against the property classes, of creditors and debtors, and of landowners and outcasts are best illustrated in the history of Antiquity.

The Significance of Acquisition and Social Classes

The primary significance of a positively privileged acquisition class is to be found in two directions. On the one hand it is generally possible to go far toward attaining a monopoly of the management of productive enterprises in favour of the members of the class and their business interests. On the other hand, such a class tends to insure the security of its economic position by exercising influence on the economic policy of political bodies and other groups.

The members of positively privileged acquisition classes are typically entrepreneurs. The following are the most important types: merchants, shipowners, industrial and agricultural entrepreneurs, bankers and financiers. Under certain circumstances two other types are also members of such classes,

namely, members of the 'liberal' professions with a privileged position by virtue of their abilities or training, and workers with special skills commanding a monopolistic position, regardless of how far they are hereditary or the result of training.

Acquisition classes in a negatively privileged situation are workers of the various principal types. They may be roughly classified as skilled, semi-skilled and unskilled.

In this connexion as well as the above, independent peasants and craftsmen are to be treated as belonging to the 'middle classes.' This category often includes in addition officials, whether they are in public or private employment, the liberal professions, and workers with exceptional monopolistic assets or positions.

Examples of 'social classes' are: (a) The 'working' class as a whole. It approaches this type the more completely mechanized the productive process becomes. (b) The petty bourgeoisie.² (c) The 'intelligentsia' without independent property and the persons whose social position is primarily dependent on technical training such as engineers, commercial and other officials, and civil servants. These groups may differ greatly among themselves, in particular according to costs of training. (d) The classes occupying a privileged position through property and education.

The unfinished concluding section of Karl Marx's *Kapital* was evidently intended to deal with the problem of the class unity of the proletariat, which he held existed in spite of the high degree of qualitative differentiation. A decisive factor is the increase in the importance of semi-skilled workers who have been trained in a relatively short time directly on the machines themselves, at the expense of the older type of 'skilled' labour and also of unskilled. However, even this type of skill may often have a monopolistic aspect. Weavers are said to attain the highest level of productivity only after five years' experience.

At an earlier period every worker could be said to have been primarily interested in be-

coming an independent small bourgeois, but the possibility of realizing this goal is becoming progressively smaller. From one generation to another the most readily available path to advancement both for skilled and semi-skilled workers is into the class of technically trained individuals. In the most highly privileged classes, at least over the period of more than one generation, it is coming more and more to be true that money is overwhelmingly decisive. Through the banks and corporate enterprises members of the lower middle class and the salaried groups have certain opportunities to rise into the privileged class.

Organized activity of class groups is favoured by the following circumstances: (a) The possibility of concentrating on opponents where the immediate conflict of interests is vital. Thus workers organize against management and not against security holders who are the ones who really draw income without working. Similarly peasants are not apt to organize against landlords. (b) The existence of a class situation which is typically similar for large masses of people. (c) The technical possibility of being easily brought together. This is particularly true where large numbers work together in a small area, as in the modern factory. (d) Leadership directed to readily understandable goals. Such goals are very generally imposed or at least are interpreted by persons, such as intelligentsia, who do not belong to the class in question.

Status and Status Group

The term of 'status'³ will be applied to a typically effective claim to positive or negative privilege with respect to social prestige so far as it rests on one or more of the following bases: (a) mode of living, (b) a formal process of education which may consist in empirical or rational training and the acquisition of the corresponding modes of life, or (c) on the prestige of birth, or of an occupation.

The primary practical manifestations of status with respect to social stratification are

conubium, commensality, and often monopolistic appropriation of privileged economic opportunities and also prohibition of certain modes of acquisition. Finally, there are conventions or traditions of other types attached to a status.

Status may be based on class situation directly or related to it in complex ways. It is not, however, determined by this alone. Property and managerial positions are not as such sufficient to lend their holder a certain status, though they may well lead to its acquisition. Similarly, poverty is not as such a disqualification for high status though again it may influence it.

Conversely, status may partly or even wholly determine class situation, without, however, being identical with it. The class situation of an officer, a civil servant, and a student as determined by their income may be widely different while their status remains the same, because they adhere to the same mode of life in all relevant respects as a result of their common education.

A 'status group' is a plurality of individuals who, within a larger group, enjoy a particular kind and level of prestige by virtue of their position and possibly also claim certain special monopolies.

The following are the most important sources of the development of distinct status groups: (a) The most important is by the development of a peculiar style of life including, particularly, the type of occupation pursued. (b) The second basis is hereditary charisma arising from the successful claim to a position of prestige by virtue of birth. (c) The third is the appropriation of political or hierocratic authority as a monopoly by socially distinct groups.

The development of hereditary status groups is usually a form of the hereditary appropriation of privileges by an organized group or by individual qualified persons. Every well-established case of appropriation of opportunities and abilities, especially of exercising imperative powers, has a tendency to lead to the

development of distinct status groups. Conversely, the development of status groups has a tendency in turn to lead to the monopolistic appropriation of governing powers and of the corresponding economic advantages.

Acquisition classes are favoured by an economic system oriented to market situations, whereas status groups develop and subsist most readily where economic organization is of a monopolistic and liturgical character and where the economic needs of corporate groups are met on a feudal or patrimonial basis. The type of class which is most closely related to a status group is the 'social' class, while the 'acquisition' class is the farthest removed. Property classes often constitute the nucleus of a status group.

Every society where status groups play a prominent part is controlled to a large extent by conventional rules of conduct. It thus creates economically irrational conditions of consumption and hinders the development

of free markets by monopolistic appropriation and by restricting free disposal of the individual's own economic ability. This will have to be discussed further elsewhere.

NOTES

1. Although Parsons chooses to translate *Klasse* as 'class status' in this context, to do so is potentially confusing because Weber so carefully distinguishes between the concepts of class and status. I have therefore followed the lead of Roth and Wittich (*Economy and Society*, 1968) and opted for the term 'class situation' throughout this essay.—ED.

2. I have again followed Roth and Wittich (*Economy and Society*, 1968) in translating the German term *Kleinbürgertum* as 'petty bourgeoisie,' whereas Parsons opted for the more ambiguous term 'lower middle' class.—ED.

3. For the purposes of consistency with the other selections, I have translated the term *ständische Lage* as 'status' (see Roth and Wittich, *Economy and Society*, 1968), whereas Parsons opted for the terms 'social status,' 'stratificatory status,' and the like.—ED.

■ Max Weber

Open and Closed Relationships

Social Relationships

A social relationship, regardless of whether it is communal or associative in character, will be spoken of as "open" to outsiders if and insofar as its system of order does not deny participation to anyone who wishes to join and is actually in a position to do so. A relationship will, on the other hand, be called "closed" against outsiders so far as, according to its subjective meaning and its binding rules, participation of certain persons is excluded, limited, or subjected to conditions.

Whether a relationship is open or closed may be determined traditionally, affectually, or rationally in terms of values or of expediency. It is especially likely to be closed, for rational reasons, in the following type of situation: a social relationship may provide the parties to it with opportunities for the satisfaction of spiritual or material interests, whether absolutely or instrumentally, or whether it is achieved through co-operative action or by a compromise of interests. If the participants expect that the admission of

others will lead to an improvement of their situation, an improvement in degree, in kind, in the security or the value of the satisfaction, their interest will be in keeping the relationship open. If, on the other hand, their expectations are of improving their position by monopolistic tactics, their interest is in a closed relationship.

There are various ways in which it is possible for a closed social relationship to guarantee its monopolized advantages to the parties. (a) Such advantages may be left free to competitive struggle within the group; (b) they may be regulated or rationed in amount and kind, or (c) they may be appropriated by individuals or sub-groups on a permanent basis and become more or less inalienable. The last is a case of closure within, as well as against outsiders. Appropriated advantages will be called “rights.” As determined by the relevant order, appropriation may be (1) for the benefit of the members of particular communal or associative groups (for instance, household groups), or (2) for the benefit of individuals. In the latter case, the individual may enjoy his rights on a purely personal basis or in such a way that in case of his death one or more other persons related to the holder of the right by birth (kinship), or by some other social relationship, may inherit the rights in question. Or the rights may pass to one or more individuals specifically designated by the holder. These are cases of hereditary appropriation. Finally, (3) it may be that the holder is more or less fully empowered to alienate his rights by voluntary agreement, either to other specific persons or to anyone he chooses. This is alienable appropriation. A party to a closed social relationship will be called a “member”; in case his participation is regulated in such a way as to guarantee him appropriated advantages, a privileged member (*Rechtsgenosse*). Appropriated rights which are enjoyed by individuals through inheritance or by hereditary groups, whether communal or associative, will be called the “property” of the individual or of groups in question; and, insofar as they are alienable, “free” property.

The apparently gratuitous tediousness involved in the elaborate definition of the above concepts is an example of the fact that we often neglect to think out clearly what seems to be obvious, because it is intuitively familiar.

1. a. Examples of communal relationships, which tend to be closed on a traditional basis, are those in which membership is determined by family relationship.

b. Personal emotional relationships are usually affectually closed. Examples are erotic relationships and, very commonly, relations of personal loyalty.

c. Closure on the basis of value-rational commitment to values is usual in groups sharing a common system of explicit religious belief.

d. Typical cases of rational closure on grounds of expediency are economic associations of a monopolistic or a plutocratic character.

A few examples may be taken at random. Whether a group of people engaged in conversation is open or closed depends on its content. General conversation is apt to be open, as contrasted with intimate conversation or the imparting of official information. Market relationships are in most, or at least in many, cases essentially open. In the case of many relationships, both communal and associative, there is a tendency to shift from a phase of expansion to one of exclusiveness. Examples are the guilds and the democratic city-states of Antiquity and the Middle Ages. At times these groups sought to increase their membership in the interest of improving the security of their position of power by adequate numbers. At other times they restricted their membership to protect the value of their monopolistic position. The same phenomenon is not uncommon in monastic orders and religious sects which have passed from a stage of religious proselytizing to one of restriction in the interest of the maintenance of an ethical standard or for the protection of

material interests. There is a similar close relationship between the extension of market relationships in the interest of increased turnover on the one hand, their monopolistic restriction on the other. The promotion of linguistic uniformity is today a natural result of the interests of publishers and writers, as opposed to the earlier, not uncommon, tendency for status groups to maintain linguistic peculiarities or even for secret languages to emerge.

2. Both the extent and the methods of regulation and exclusion in relation to outsiders may vary widely, so that the transition from a state of openness to one of regulation and closure is gradual. Various conditions of participation may be laid down; qualifying tests, a period of probation, requirement of possession of a share which can be purchased under certain conditions, election of new members by ballot, membership or eligibility by birth or by virtue of achievements open to anyone. Finally, in case of closure and the appropriation of rights within the group, participation may be dependent on the acquisition of an appropriated right. There is a wide variety of different degrees of closure and of conditions of participation. Thus regulation and closure are relative concepts. There are all manner of gradual shadings as between an exclusive club, a theatrical audience the members of which have purchased tickets, and a party rally to which the largest possible number has been urged to come; similarly, from a church service open to the general public through the rituals of a limited sect to the mysteries of a secret cult.

3. Similarly, closure within the group may also assume the most varied forms. Thus a caste, a guild, or a group of stock exchange brokers, which is closed to outsiders, may allow to its members a perfectly free competition for all the advantages which the group as a whole monopolizes for itself. Or it may assign every member strictly to the enjoyment of certain advantages, such as claims over customers or particular business opportunities,

for life or even on a hereditary basis. This is particularly characteristic of India. Similarly, a closed group of settlers (*Markgenossenschaft*) may allow its members free use of the resources of its area or may restrict them rigidly to a plot assigned to each individual household. A closed group of colonists may allow free use of the land or sanction and guarantee permanent appropriation of separate holdings. In such cases all conceivable transitional and intermediate forms can be found. Historically, the closure of eligibility to fiefs, benefices, and offices within the group, and the appropriation on the part of those enjoying them, have occurred in the most varied forms. Similarly, the establishment of rights to and possession of particular jobs on the part of workers may develop all the way from the “closed shop” to a right to a particular job. The first step in this development may be to prohibit the dismissal of a worker without the consent of the workers’ representatives. The development of the “works councils” [in Germany after 1918] might be a first step in this direction, though it need not be. . . .

4. The principal motives for closure of a relationship are: (a) The maintenance of quality, which is often combined with the interest in prestige and the consequent opportunities to enjoy honor, and even profit; examples are communities of ascetics, monastic orders, especially, for instance, the Indian mendicant orders, religious sects like the Puritans, organized groups of warriors, of *ministeriales* and other functionaries, organized citizen bodies as in the Greek states, craft guilds; (b) the contraction of advantages in relation to consumption needs (*Nahrungsspielraum*); examples are monopolies of consumption, the most developed form of which is a self-subsistent village community; (c) the growing scarcity of opportunities for acquisition (*Erwerbsspielraum*). This is found in trade monopolies such as guilds, the ancient monopolies of fishing rights, and so on. Usually motive (a) is combined with (b) or (c). . . .

Economic Relationships

One frequent economic determinant [of closure] is the competition for a livelihood—offices, clients and other remunerative opportunities. When the number of competitors increases in relation to the profit span, the participants become interested in curbing competition. Usually one group of competitors takes some externally identifiable characteristic of another group of (actual or potential) competitors—race, language, religion, local or social origin, descent, residence, etc.—as a pretext for attempting their exclusion. It does not matter which characteristic is chosen in the individual case: whatever suggests itself most easily is seized upon. Such group action may provoke a corresponding reaction on the part of those against whom it is directed.

In spite of their continued competition against one another, the jointly acting competitors now form an “interest group” toward outsiders; there is a growing tendency to set up some kind of association with rational regulations; if the monopolistic interests persist, the time comes when the competitors, or another group whom they can influence (for example, a political community), establish a legal order that limits competition through formal monopolies; from then on, certain persons are available as “organs” to protect the monopolistic practices, if need be, with force. In such a case, the interest group has developed into a “legally privileged group” (*Rechtsgemeinschaft*) and the participants have become “privileged members” (*Rechtsgenossen*). Such closure, as we want to call it, is an ever-recurring process; it is the source of property in land as well as of all guild and other group monopolies.

The tendency toward the monopolization of specific, usually economic opportunities is always the driving force in such cases as: “co-operative organization,” which always means closed monopolistic groups, for example, of

fishermen taking their name from a certain fishing area; the establishment of an association of engineering graduates, which seeks to secure a legal, or at least factual, monopoly over certain positions; the exclusion of outsiders from sharing in the fields and commons of a village; “patriotic” associations of shop clerks; the *ministeriales*, knights, university graduates and craftsmen of a given region or locality; ex-soldiers entitled to civil service positions—all these groups first engage in some joint action (*Gemeinschaftshandeln*) and later perhaps an explicit association. This monopolization is directed against competitors who share some positive or negative characteristics; its purpose is always the closure of social and economic opportunities to *outsiders*. Its extent may vary widely, especially so far as the group member shares in the apportionment of monopolistic advantages. . . .

This monopolistic tendency takes on specific forms when groups are formed by persons with shared qualities *acquired* through upbringing, apprenticeship and training. These characteristics may be economic qualifications of some kind, the holding of the same or of similar offices, a knightly or ascetic way of life, etc. If in such a case an association results from social action, it tends toward the *guild*. Full members make a vocation out of monopolizing the disposition of spiritual, intellectual, social and economic goods, duties and positions. Only those are admitted to the unrestricted practice of the vocation who (1) have completed a novitiate in order to acquire the proper training, (2) have proven their qualification, and (3) sometimes have passed through further waiting periods and met additional requirements. This development follows a typical pattern in groups ranging from the juvenile student fraternities, through knightly associations and craft-guilds, to the qualifications required of the modern officials and employees. It is true that the interest in guaranteeing

an efficient performance may everywhere have some importance; the participants may desire it for idealistic or materialistic reasons in spite of their possibly continuing competition with one another: local craftsmen may desire it for the sake of their business reputation, *ministeriales* and knights of a given association for the sake of their professional reputation and also their own military security, and ascetic groups for fear that the gods and demons may turn their wrath against all members because of faulty manipulations.

(For example, in almost all primitive tribes, persons who sang falsely during a ritual dance were originally slain in expiation of such an offense.) But normally this concern for efficient performance recedes behind the interest in limiting the supply of candidates for the benefices and honors of a given occupation. The novitiates, waiting periods, masterpieces and other demands, particularly the expensive entertainment of group members, are more often economic than professional tests of qualification.

14 ■ Anthony Giddens

The Class Structure of the Advanced Societies

The Weberian Critique

For the most significant developments in the theory of classes since Marx, we have to look to those forms of social thought whose authors, while being directly influenced by Marx's ideas, have attempted at the same time to criticise or to reformulate them. This tendency has been strongest, for a combination of historical and intellectual reasons, in German sociology, where a series of attempts have been made to provide a fruitful critique of Marx—beginning with Max Weber, and continuing through such authors as Geiger, Renner and Dahrendorf.¹ Weber's critique of Marx here has been of particular importance. But, especially in the English-speaking world, the real import of Weber's analysis has frequently been misrepresented. The customary procedure has been to contrast Weber's discussion of 'Class, status and party', a fragment of *Economy and Society*, with the conception of class supposedly taken by Marx, to

the detriment of the latter. Marx, so it is argued, treated 'class' as a purely economic phenomenon and, moreover, regarded class conflicts as in some way the 'inevitable' outcome of clashes of material interest. He failed to realise, according to this argument, that the divisions of economic interest which create classes do not necessarily correspond to sentiments of communal identity which constitute differential 'status'. Thus, status, which depends upon subjective evaluation, is a separate 'dimension of stratification' from class, and the two may vary independently. There is yet a third dimension, so the argument continues, which Weber recognised as an independently variable factor in 'stratification', but which Marx treated as directly contingent upon class interests. This is the factor of 'power'.²

Evaluation of the validity of this interpretation is difficult because there is no doubt that Weber himself accepted it—or certain

elements of it. What is often portrayed in the secondary literature as a critique of ‘Marx’s conception of class’ actually takes a stilted and impoverished form of crude Marxism as its main target of attack. But this sort of determinist Marxism was already current in Germany in Weber’s lifetime, and since Weber himself set out to question this determinism, the true lines of similarity and difference between his and Marx’s analysis of classes are difficult to disentangle. . . .³

In the two versions of ‘Class, status and party’ which have been embodied in *Economy and Society*,⁴ Weber provides what is missing in Marx: an explicit discussion of the concept of class. There are two principal respects in which this analysis differs from Marx’s ‘abstract model’ of classes. One is that which is familiar from most secondary accounts—the differentiation of ‘class’ from ‘status’ and ‘party’. The second, however, as will be argued below, is equally important: this is that, although Weber employs for some purposes a dichotomous model which in certain general respects resembles that of Marx, his viewpoint strongly emphasises a *pluralistic conception of classes*. Thus Weber’s distinction between ‘ownership classes’ (*Besitzklassen*) and ‘acquisition classes’ (*Erwerbsklassen*) is based upon a fusion of two criteria: ‘on the one hand . . . the kind of property that is usable for returns; and, on the other hand . . . the kind of services that can be offered on the market’, thus producing a complex typology. The sorts of property which may be used to obtain market returns, although dividing generally into two types—creating ownership (*rentier*) and acquisition (entrepreneurial) classes—are highly variable, and may produce many differential interests within dominant classes:

Ownership of dwellings; workshops; warehouses; stores; agriculturally usable land in large or small-holdings—a quantitative difference with possibly qualitative consequences; ownership of mines; cattle; men

(slaves); disposition over mobile instruments of production, or capital goods of all sorts, especially money or objects that can easily be exchanged for money; disposition over products of one’s own labour or of others’ labour differing according to their various distances from consumability; disposition over transferable monopolies of any kind—all these distinctions differentiate the class situations of the propertied . . .⁵

But the class situations of the propertyless are also differentiated, in relation both to the types and the degree of ‘monopolisation’ of ‘marketable skills’ which they possess. Consequently, there are various types of ‘middle class’ which stand between the ‘positively privileged’ classes (the propertied) and the ‘negatively privileged’ classes (those who possess neither property nor marketable skills). While these groupings are all nominally propertyless, those who possess skills which have a definite ‘market value’ are certainly in a different class situation from those who have nothing to offer but their (unskilled) labour. In acquisition classes—i.e., those associated particularly with the rise of modern capitalism—educational qualifications take on a particular significance in this respect; but the monopolisation of trade skills by manual workers is also important.

Weber insists that a clear-cut distinction must be made between class ‘in itself’ and class ‘for itself’: ‘class’, in his terminology, always refers to market interests, which exist independently of whether men are aware of them. Class is thus an ‘objective’ characteristic influencing the life-chances of men. But only under certain conditions do those sharing a common class situation become conscious of, and act upon, their mutual economic interests. In making this emphasis, Weber undoubtedly intends to separate his position from that adopted by many Marxists, involving what he calls a ‘pseudo-scientific operation’ whereby the link between class and class consciousness is treated as direct

and immediate.⁶ Such a consideration evidently also underlies the emphasis which Weber places upon ‘status groups’ (*Stände*) as contrasted to classes. The contrast between class and status group, however, is not, as often seems to be assumed, merely, nor perhaps even primarily, a distinction between subjective and objective aspects of differentiation. While class is founded upon differentials of economic interest in market relationships, Weber nowhere denies that, under certain given circumstances, a class may be a subjectively aware ‘community’. The importance of status groups—which are normally ‘communities’ in this sense—derives from the fact that they are built upon criteria of grouping other than those stemming from market situation. The contrast between classes and status groups is sometimes portrayed by Weber as one between the objective and the subjective; but it is also one between production and consumption. Whereas class expresses relationships involved in production, status groups express those involved in consumption, in the form of specific ‘styles of life’.

Status affiliations may cut across the relationships generated in the market, since membership of a status group usually carries with it various sorts of monopolistic privileges. Nonetheless, classes and status groups tend in many cases to be closely linked, through property: possession of property is both a major determinant of class situation and also provides the basis for following a definite ‘style of life’. The point of Weber’s analysis is not that class and status constitute two ‘dimensions of stratification’, but that classes and status communities represent two possible, and competing, modes of group formation in relation to the distribution of power in society. Power is *not*, for Weber, a ‘third dimension’ in some sense comparable to the first two. He is quite explicit about saying that classes, status groups and parties are all ‘phenomena of the distribution of power’.⁷ The theorem informing Weber’s position here

is his insistence that power is not to be assimilated to economic domination—again, of course, a standpoint taken in deliberate contrast to that of Marx. The party, oriented towards the acquisition or maintenance of political leadership, represents, like the class and the status group, a major focus of social organisation relevant to the distribution of power in a society. It is, however, only characteristic of the modern rational state. . . .

In his conceptual discussion of class, besides distinguishing the purely economic *Besitzklassen* and *Erwerbsklassen*, Weber also refers to what he calls ‘social classes’. A social class, in Weber’s sense, is formed of a cluster of class situations which are linked together by virtue of the fact that they involve common mobility chances, either within the career of individuals or across the generations. Thus while a worker may fairly readily move from an unskilled to a semi-skilled manual occupation, and the son of an unskilled worker may become a semi-skilled or perhaps a skilled worker, the chances of either intra- or inter-generational mobility into non-manual occupations are much less. While the conception of the ‘social class’ remains relatively undeveloped in Weber’s writings, it is of particular interest in relation to his model of capitalist development. As Weber himself points out, the notion of ‘social class’ comes much closer to that of ‘status group’ than does the conception of purely economic class (although, as with economic class situation, individuals who are in the same social class are not necessarily conscious of the fact). The notion of social class is important because it introduces a unifying theme into the diversity of cross-cutting class relationships which may stem from Weber’s identification of ‘class situation’ with ‘market position’. If the latter is applied strictly, it is possible to distinguish an almost endless multiplicity of class situations. But a ‘social class’ exists only when these class situations cluster together in such a way as to create a common nexus of social interchange between individuals. In capitalism, Weber

distinguishes four main social class groupings: the manual working class; the petty bourgeoisie; propertyless white-collar workers: 'technicians, various kinds of white-collar employees, civil servants—possibly with considerable social differences depending on the cost of their training'; and those 'privileged through property and education'.⁸ Of these social class groupings, the most significant are the working class, the propertyless 'middle class' and the propertied 'upper class'. Weber agrees with Marx that the category of small property-owners (*Kleinbürgertum*) tends to become progressively more restricted with the increasing maturity of capitalism. The result of this process, however, is not normally that they 'sink into the proletariat', but that they become absorbed into the expanding category of skilled manual or non-manual salaried workers.

To emphasise, therefore, that Weber's 'abstract model' of classes is a pluralistic one is not to hold that he failed to recognise unifying ties between the numerous combinations of class interests made possible by his conception of 'class situation'. But there is no doubt that his viewpoint drastically amends important elements of Marx's picture of the typical trend of development of the capitalist class structure. Even Weber's simplified ('social class') model of capitalism diverges significantly from the Marxian conception, in treating the propertyless 'middle class' as the category which tends to expand most with the advance of capitalism. Moreover, the social classes do not necessarily constitute 'communities', and they may be fragmented by interest divisions deriving from differentials in market position; and finally, as Weber shows in his historical writings, the relationship between class structure and the political sphere is a contingent one. . . .

Rethinking the Theory of Class

The deficiency in Weber's reinterpretation of Marx's view is that it is not sufficiently radi-

cal. While Weber recognises the unsatisfactory character of the Marxian standpoint, particularly as regards the undifferentiated category of the 'propertyless', he does not pursue the implications of his own conception far enough. Dahrendorf has suggested that we may stand the Marxian concept of property on its head in terms of its relation to authority;⁹ the implications of the Weberian analysis, however, are that the conception of property may be 'inverted' or generalised in a different way, which does not sacrifice the economic foundation of the concept of class. 'Property' refers, not to any characteristics of physical objects as such, but to rights which are associated with them, which in turn confer certain capacities upon the 'owner'. . . . In the market, of course, the significance of capital as private property is that it confers certain very definite capacities upon its possessor as compared to those who are 'propertyless'—those who do not own their means of production. But we can readily perceive that, even in the Marxian view, the notion of 'propertylessness' is something of a misnomer. For if 'property' is conceived of as a set of capacities of action with reference to the operations of the market, it is evident that the wage-labourer does possess such capacities. The 'property' of the wage-labourer is the labour-power which he brings for sale in entering into the contractual relation. While this fundamentally disadvantages him in the competitive bargaining situation in relation to the owner of capital, this is not simply a one-way power relationship: the 'property' which the wage-labourer possesses is needed by the employer, and he must pay at least some minimal attention to the demands of the worker—providing a basis for the collective withdrawal of labour as a possible sanction. It would be departing too much from usual terminology to refer to capital and to the labour-power of the worker both as 'property'; and, anyway, the point is rather that 'property' (capital) is a particular case of capacity to determine the bargaining

outcome, rather than vice versa. So I shall continue to speak below of ‘property’ (in the means of production) in a conventional sense, and shall use the term ‘market capacity’ in an inclusive manner *to refer to all forms of relevant attributes which individuals may bring to the bargaining encounter*.

It is an elementary fact that where ownership of property is concentrated in the hands of a minority and in a society in which the mass of the population is employed in industrial production, the vast majority consequently offer their labour for sale on the market. Because of his general emphasis upon ‘productive labour’, and because of his expectation that it is in the nature of modern technology to reduce productive operations to a homogeneous skill-level, Marx failed to recognise the potential significance of differentiations of market capacity which do not derive directly from the factor of property ownership. Such differentiations, it seems clear, are contingent upon the scarcity value of what the individual ‘owns’ and is able to offer on the market. As Weber indicates, possession of recognised ‘skills’—including educational qualifications—is the major factor influencing market capacity. Differentiations in market capacity may be used, as various recent authors have indicated, to secure economic returns other than income as such. These include, principally, security of employment, prospects of career advancement, and a range of ‘fringe benefits’, such as pension rights, etc.¹⁰ In the same way as the capacities which individuals bring to the bargaining process may be regarded as a form of ‘property’ which they exchange on the market, so these material returns may be regarded as forms of ‘good’ which are obtained through the sale of labour-power.

In the market structure of competitive capitalism, *all* those who participate in the exchange process are in a certain sense in (interest) conflict with one another for access to scarce returns. Conflict of interest may be created by the existence of many sorts of dif-

ferential market capacities. Moreover, the possible relationships between property and ‘propertyless’ forms of market capacity are various. Speculative investment in property may, for example, be one of the specific market advantages used by those in certain occupations (thus directors are often able to use ‘inside knowledge’ to profit from property deals). Marx himself, of course, recognised the existence of persistent conflicts of interest within property-owning groupings: notably, between financial and industrial sectors of the large bourgeoisie, and between large and petty bourgeoisie.

The difficulty of identifying ‘class’ with common market capacity has already been alluded to with reference to Weber. While Weber’s concept of ‘market situation’ successfully moves away from some of the rigidities of the Marxian scheme, it tends to imply the recognition of a cumbersome plurality of classes. There would appear to be as many ‘classes’, and as many ‘class conflicts’, as there are differing market positions. The problem here, however, is not the recognition of the diversity of the relationships and conflicts created by the capitalist market as such, but that of making the *theoretical transition from such relationships and conflicts to the identification of classes as structured forms*. The unsatisfactory and ill-defined character of the connections between ‘class position’, the typology of *Besitzklassen* and *Erwerbsklassen*, and ‘social classes’ in Weber’s work has already been mentioned. But the problem is by no means confined to Weber’s theoretical scheme. Marx was certainly conscious of the problematic character of the links between class as a latent set of characteristics generated by the capitalist system and class as an historical, dynamic entity, an ‘historical actor’. But his contrast between class ‘in itself’ and class ‘for itself’ is primarily one distinguishing between class relationships as a cluster of economic connections on the one hand and class consciousness on the other. This emphasis was very much dictated by the

nature of Marx's interests, lying as they did above all in understanding and promoting the rise of a revolutionary class consciousness within capitalism. While it would by no means be true to hold that Marx ignored this completely, it can be said that he gave only little attention to the modes in which classes, founded in a set of economic relationships, take on or 'express' themselves in definite social forms.

Nor has the matter been adequately dealt with in the writings of later authors. In fact, one of the leading dilemmas in the theory of class—which figures prominently, for example, in Aron's discussion—is that of identifying the 'reality' of class.¹¹ Not only has there been some considerable controversy over whether class is a 'real' or 'nominal' category, but many have argued that, since it is difficult or impossible to draw the 'boundaries' between classes with any degree of clarity, we should abandon the notion of class as a useful sociological concept altogether.¹² Only Dahrendorf seems to have attempted to give attention to the problem within the framework of an overall theory of class, and since his identification of class with authority divisions is unacceptable [for reasons outlined elsewhere],¹³ his analysis does not help greatly.

The major problems in the theory of class, I shall suggest, do not so much concern the nature and application of the class concept itself, as what, for want of a better word, I shall call the *structuration* of class relationships.¹⁴ Most attempts to revise class theory since Marx have sought to accomplish such a revision primarily by refining, modifying, or substituting an altogether different notion for the Marxian concept of class. While it is useful to follow and develop certain of Weber's insights in this respect, the most important blank spots in the theory of class concern the processes whereby 'economic classes' become 'social classes', and whereby in turn the latter are related to other social forms. As Marx was anxious to stress in criti-

cising the premises of political economy, all economic relationships, and any sort of 'economy', presuppose a set of social ties between producers. In arguing for the necessity of conceptualising the structuration of class relationships, I do not in any way wish to question the legitimacy of this insight, but rather to focus upon *the modes in which* 'economic' relationships become translated into 'non-economic' social structures.

One source of terminological ambiguity and conceptual confusion in the usage of the term 'class' is that it has often been employed to refer both to an economic *category* and to a specifiable cluster of social groupings. Thus Weber uses the term in both of these ways, although he seeks terminologically to indicate the difference between 'class' (as a series of 'class positions') and 'social class'. But in order to insist that the study of class and class conflict must concern itself with the interdependence of economy and society, it is not necessary to identify the term 'class' with the divisions and interests generated by the market as such. Consequently, in the remainder of this [essay], I shall use the term in the sense of Weber's 'social class'—appropriately explicated. While there may be an indefinite multiplicity of cross-cutting interests created by differential market capacities, there are only, in any given society, a limited number of classes.

It will be useful at this juncture to state what class is *not*. First, a class is not a specific 'entity'—that is to say, a bounded social form in the way in which a business firm or a university is—and a class has no publicly sanctioned identity. It is extremely important to stress this, since established linguistic usage often encourages us to apply active verbs to the term 'class'; but the sense in which a class 'acts' in a certain way, or 'perceives' elements in its environment on a par with an individual actor, is highly elliptical, and this sort of verbal usage is to be avoided wherever possible. Similarly, it is perhaps misleading to speak of 'membership' of a class, since this

might be held to imply participation in a definite ‘group’. This form of expression, however, is difficult to avoid altogether, and I shall not attempt to do so in what follows. Secondly, class has to be distinguished from ‘stratum’, and class theory from the study of ‘stratification’ as such. The latter, comprising what Ossowski terms a gradation scheme, involves a criterion or set of criteria in terms of which individuals may be ranked descriptively along a scale.¹⁵ The distinction between class and stratum is again a matter of some significance, and bears directly upon the problem of class ‘boundaries’. For the divisions between strata, for analytical purposes, may be drawn very precisely, since they may be set upon a measurement scale—as, for example, with ‘income strata’. The divisions between classes are *never* of this sort; nor, moreover, do they lend themselves to easy visualisation, in terms of any ordinal scale of ‘higher’ and ‘lower’, as strata do—although this sort of imagery cannot be escaped altogether. Finally we must distinguish clearly between class and elite. Elite theory, as formulated by Pareto and Mosca, developed in part as a conscious and deliberate repudiation of class analysis.¹⁶ In place of the concept of class relationships, the elite theorists substituted the opposition of ‘elite’ and ‘mass’; and in place of the Marxian juxtaposition of class society and classlessness they substituted the idea of the cyclical replacement of elites *in perpetuo*. . . .

The Structuration of Class Relationships

It is useful, initially, to distinguish the *mediate* from the *proximate* structuration of class relationships. By the former term, I refer to the factors which intervene between the existence of certain given market capacities and the formation of classes as identifiable social groupings, that is to say which operate as ‘overall’ connecting links between the market on the one hand and structured systems of

class relationships on the other. In using the latter phrase, I refer to ‘localised’ factors which condition or shape class formation. The mediate structuration of class relationships is governed above all by the distribution of mobility chances which pertain within a given society. Mobility has sometimes been treated as if it were in large part separable from the determination of class structure. According to Schumpeter’s famous example, classes may be conceived of as like conveyances, which may be constantly carrying different ‘passengers’ without in any way changing their shape. But, compelling though the analogy is at first sight, it does not stand up to closer examination, especially within the framework I am suggesting here.¹⁷ In general, the greater the degree of ‘closure’ of mobility chances—both intergenerationally and within the career of the individual—the more this facilitates the formation of identifiable classes. For the effect of closure in terms of intergenerational movement is to provide for the *reproduction* of common life experience over the generations; and this homogenisation of experience is reinforced to the degree to which the individual’s movement within the labour market is confined to occupations which generate a similar range of material outcomes. In general we may state that the structuration of classes is facilitated *to the degree to which mobility closure exists in relation to any specified form of market capacity*. There are three sorts of market capacity which can be said to be normally of importance in this respect: ownership of property in the means of production; possession of educational or technical qualifications; and possession of manual labour-power. In so far as it is the case that these tend to be tied to closed patterns of inter- and intragenerational mobility, this yields the foundation of a *basic three-class system* in capitalist society: an ‘upper’, ‘middle’, and ‘lower’ or ‘working’ class. But as has been indicated previously, it is an intrinsic characteristic of the development of the cap-

italist market that there exist no legally sanctioned or formally prescribed limitations upon mobility, and hence it must be emphasised that there is certainly never anything even approaching complete closure. In order to account for the emergence of structured classes, we must look in addition at the proximate sources of structuration.

There are three, related, sources of proximate structuration of class relationships: the division of labour within the productive enterprise; the authority relationships within the enterprise; and the influence of what I shall call 'distributive groupings'. I have already suggested that Marx tended to use the notion of 'division of labour' very broadly, to refer both to market relationships and to the allocation of occupational tasks within the productive organisation. Here I shall use the term only in this second, more specific, sense. In capitalism, the division of labour in the enterprise is in principle governed by the promotion of productive efficiency in relation to the maximisation of profit; but while responding to the same exigencies as the capitalist market in general, the influence of the division of labour must be analytically separated as a distinctive source of structuration (and, as will be discussed later, as a significant influence upon class consciousness). The division of labour, it is clear, may be a basis of the fragmentation as well as the consolidation of class relationships. It furthers the formation of classes to the degree to which it creates homogeneous groupings which cluster along the same lines as those which are fostered by mediate structuration. Within the modern industrial order, the most significant influence upon proximate structuration in the division of labour is undoubtedly that of technique. The effect of industrial technique (more recently, however, modified by the introduction of cybernetic systems of control) is to create a decisive separation between the conditions of labour of manual and non-manual workers. 'Machine-minding', in one form or another, regardless

of whether it involves a high level of manual skill, tends to create a working environment quite distinct from that of the administrative employee, and one which normally enforces a high degree of physical separation between the two groupings.¹⁸

This effect of the division of labour thus overlaps closely with the influence of the mediate structuration of class relationships through the differential apportionment of mobility chances; but it is, in turn, potentially heavily reinforced by the typical authority system in the enterprise. In so far as administrative workers participate in the framing, or merely in the enforcement, of authoritative commands, they tend to be separated from manual workers, who are subject to those commands. But the influence of differential authority is also basic as a reinforcing agent of the structuration of class relationships at the 'upper' levels. Ownership of property, in other words, confers certain fundamental capacities of command, maximised within the 'entrepreneurial' enterprise in its classical form. To the extent to which this serves to underlie a division at 'the top', in the control of the organisation (something which is manifestly influenced, but not at all destroyed, if certain of the suppositions advanced by the advocates of the theory of the separation of 'ownership and control' are correct) it supports the differentiation of the 'upper' from the 'middle' class.

The third source of the proximate structuration of class relationships is that originating in the sphere of consumption rather than production. Now according to the traditional interpretations of class structure, including those of Marx and Weber, 'class' is a phenomenon of production: relationships established in consumption are therefore quite distinct from, and secondary to, those formed in the context of productive activity. There is no reason to deviate from this general emphasis. But without dropping the conception that classes are founded ultimately in the economic structure of the

capitalist market, it is still possible to regard consumption patterns as a major influence upon class structuration. Weber's notions of 'status' and 'status group', as I previously pointed out, confuse two separable elements: the formation of groupings in consumption, on the one hand, and the formation of types of social differentiation based upon some sort of non-economic value providing a scale of 'honour' or 'prestige' on the other. While the two may often coincide, they do not necessarily do so, and it seems worthwhile to distinguish them terminologically. Thus I shall call *'distributive groupings'* those relationships involving common patterns of the consumption of economic goods, regardless of whether the individuals involved make any type of conscious evaluation of their honour or prestige relative to others; 'status' refers to the existence of such evaluations, and a 'status group' is, then, any set of social relationships which derives its coherence from their application.¹⁹

In terms of class structuration, distributive groupings are important in so far as they interrelate with the other sets of factors distinguished above in such a way as to reinforce the typical separations between forms of market capacity. The most significant distributive groupings in this respect are those formed through the tendency towards community or neighbourhood segregation. Such a tendency is not normally based only upon differentials in income, but also upon such factors as access to housing mortgages, etc. The creation of distinctive 'working-class neighbourhoods' and 'middle-class neighbourhoods', for example, is naturally promoted if those in manual labour are by and large denied mortgages for house buying, while those in non-manual occupations experience little difficulty in obtaining such loans. Where industry is located outside of the major urban areas, homogeneous 'working-class communities' frequently develop through

the dependence of workers upon housing provided by the company.

In summary, to the extent to which the various bases of mediate and proximate class structuration overlap, classes will exist as distinguishable formations. I wish to say that *the combination of the sources of mediate and proximate structuration distinguished here, creating a threefold class structure, is generic to capitalist society*. But the mode in which these elements are merged to form *a specific class system*, in any given society, differs significantly according to variations in economic and political development. It should be evident that structuration is never an all-or-nothing matter. The problem of the existence of distinct class 'boundaries', therefore, is not one which can be settled *in abstracto*: one of the specific aims of class analysis in relation to empirical societies must necessarily be that of determining how strongly, in any given case, the 'class principle' has become established as a mode of structuration. Moreover, the operation of the 'class principle' may also involve the creation of forms of structuration within the major class divisions. One case in point is that which Marx called the 'petty bourgeoisie'. In terms of the preceding analysis, it is quite easy to see why ownership of small property in the means of production might come to be differentiated both from the upper class and from the ('new') middle class. If it is the case that the chances of mobility, either inter- or intragenerationally, from small to large property ownership are slight, this is likely to isolate the small property-owner from membership of the upper class as such. But the fact that he enjoys directive control of an enterprise, however minute, acts to distinguish him from those who are part of a hierarchy of authority in a larger organisation. On the other hand, the income and other economic returns of the petty bourgeois are likely to be similar to the white-collar worker, and hence they may belong to similar distributive groupings. A sec-

and potentially important influence upon class formation is to be traced to the factor of skill differential within the general category of manual labour. The manual worker who has undergone apprenticeship, or a comparable period of training, possesses a market capacity which sets him apart from the unskilled or semi-skilled worker. This case will be discussed in more detail [elsewhere];²⁰ it is enough merely to indicate at this point that there are certain factors promoting structuration on the basis of this differentiation in market capacity (e.g., that the chances of intergenerational mobility from skilled manual to white-collar occupations are considerably higher than they are from unskilled and semi-skilled manual occupations).

So far I have spoken of structuration in a purely formal way, as though class could be defined in terms of relationships which have no 'content'. But this obviously will not do: if classes become social realities, this must be manifest in the formation of common patterns of behaviour and attitude. Since Weber's discussion of classes and status groups, the notion of 'style of life' has normally come to be identified as solely pertaining to the mode whereby a status group expresses its claim to distinctiveness. However, in so far as there is marked convergence of the sources of structuration mentioned above, classes will also tend to manifest common styles of life.

An initial distinction can be drawn here between 'class awareness' and 'class consciousness'.²¹ We may say that, in so far as class is a structured phenomenon, there will tend to exist a common awareness and acceptance of similar attitudes and beliefs, linked to a common style of life, among the members of the class. 'Class awareness', as I use the term here, does *not* involve a recognition that these attitudes and beliefs signify a particular class affiliation, or the recognition that there exist other classes, characterised by

different attitudes, beliefs, and styles of life; 'class consciousness', by contrast, as I shall use the notion, does imply both of these. The difference between class awareness and class consciousness is a fundamental one, because class awareness may take the form of a *denial of the existence or reality of classes*.²² Thus the class awareness of the middle class, in so far as it involves beliefs which place a premium upon individual responsibility and achievement, is of this order.

Within ethnically and culturally homogeneous societies, the degree of class structuration will be determined by the interrelationship between the sources of structuration identified previously. But many, if not the majority, of capitalist societies are not homogeneous in these respects. Traditionally, in class theory, racial or religious divisions have been regarded as just so many 'obstacles' to the formation of classes as coherent unities. This may often be so, where these foster types of structuration which deviate from that established by the 'class principle' (as typically was the case in the battles fought by the rearguard of feudalism against the forces promoting the emergence of capitalism). The idea that ethnic or cultural divisions serve to dilute or hinder the formation of classes is also very explicitly built into Weber's separation of (economic) 'class' and 'status group'. But this, in part at least, gains its cogency from the contrast between estate, as a legally constituted category, and class, as an economic category. While it may be agreed, however, that the *bases* of the formation of classes and status groups (in the sense in which I have employed these concepts) are different, nonetheless the tendency to class structuration may receive a considerable impetus *where class coincides with the criteria of status group membership*—in other words, where structuration deriving from economic organisation 'overlaps' with, or, in Dahrendorf's terms, is 'superimposed' upon, that deriving from evaluative categorisations based

upon ethnic or cultural differences.²³ Where this is so, status group membership itself becomes a form of market capacity. Such a situation frequently offers the strongest possible source of class structuration, whereby there develop clear-cut differences in attitudes, beliefs and style of life between the classes. Where ethnic differences serve as a 'disqualifying' market capacity, such that those in the category in question are heavily concentrated among the lowest-paid occupations, or are chronically unemployed or semi-employed, we may speak of the existence of an *underclass*.²⁴

NOTES

1. Theodor Geiger, *Die Klassengesellschaft im Schmelztiegel* (Cologne 1949); Karl Renner, *Wandlungen der modernen Gesellschaft* (Vienna 1953); Ralf Dahrendorf, *Class and Class Conflict in Industrial Society* (Stanford 1959).

2. For a cogent representation of this view, see W. G. Runciman, 'Class, status and power', in J. A. Jackson, *Social Stratification* (Cambridge 1968).

3. See my *Capitalism and Modern Social Theory* (Cambridge 1971), pp. 185ff. and *passim*.

4. *Economy and Society*, vol. 2 (New York 1968), pp. 926–40, and vol. 1, pp. 302–7.

5. *Ibid.*, vol. 2, p. 928.

6. *Ibid.*, p. 930.

7. *Ibid.*, p. 927.

8. *Ibid.*, p. 305.

9. Dahrendorf, *Class and Class Conflict*.

10. See, for example, David Lockwood, *The Blackcoated Worker* (London 1958), pp. 202–4; Frank Parkin, *Class Inequality and Political Order* (London 1971).

11. Raymond Aron, *La lutte des classes* (Paris 1964).

12. See Robert A. Nisbet, 'The decline and fall of social class', *Pacific Sociological Review* 2, 1959.

13. Anthony Giddens, *The Class Structure of the Advanced Societies* (New York 1973), ch. 4.

14. What I call class structuration, Gurvitch calls negatively 'résistance à la pénétration par la société globale'. Georges Gurvitch, *Le concept de*

classes sociales de Marx à nos jours (Paris 1954), p. 116 and *passim*.

15. Stanislaw Ossowski, *Class Structure in the Social Consciousness* (London 1963).

16. Vilfredo Pareto, *The Mind and Society* (New York 1935); Gaetano Mosca, *The Ruling Class* (New York 1939).

17. We may, however, agree with Schumpeter that 'The family, not the physical person, is the true unit of class and class theory' (Joseph Schumpeter, *Imperialism, Social Classes*, Cleveland 1961). This is actually completely consistent with the idea that mobility is fundamental to class formation.

18. Lockwood, *The Blackcoated Worker*, op. cit.

19. It might be pointed out that it would easily be possible to break down the notion of status group further: according, for example, to whether the status evaluations in question are made primarily by others outside the group, and rejected by those inside it, etc.

20. Giddens, *The Class Structure of the Advanced Societies*.

21. This is not, of course, the same as Lukács' 'class-conditioned unconsciousness'; but I believe that Lukács is correct in distinguishing qualitatively different 'levels' of class consciousness. Georg Lukács, *History and Class Consciousness* (London 1971), pp. 52ff.

22. cf. Nicos Poulantzas, *Pouvoir politique et classes sociales de l'état capitaliste* (Paris 1970). It is misleading, however, to speak of *classes sans conscience*, as Crozier does. See Michel Crozier, 'Classes sans conscience ou préfiguration de la société sans classes', *Archives européennes de sociologie* 1, 1960; also 'L'ambiguïté de la conscience de classe chez les employés et les petits fonctionnaires', *Cahiers internationaux de sociologie* 28, 1955.

23. Or, to use another terminology, where there is 'overdetermination' (Louis Althusser, *For Marx*, London 1969, pp. 89–128).

24. Marx's *Lumpenproletariat*, according to this usage, is only an underclass when the individuals in question tend to derive from distinctive ethnic backgrounds. Leggett has referred to the underclass as the 'marginal working class', defining this as 'a sub-community of workers who belong to a subordinate ethnic or racial group which is usually proletarianised and highly segregated' (John C. Leggett, *Class, Race, and Labor*, New York 1968, p. 14).

15 ■ Frank Parkin

Marxism and Class Theory**A Bourgeois Critique****The 'Boundary Problem' in Sociology**

The persistent attractions of Marxist class theory have almost certainly been boosted by the less than inspiring alternative offered by academic sociology. In so far as there is any sort of tacitly agreed upon model of class among western social theorists it takes the form of the familiar distinction between manual and non-manual labour. No other criterion for identifying the class boundary seems to enjoy such widespread acceptance among those who conduct investigations into family structure, political attitudes, social imagery, life-styles, educational attainment, and similar enquiries that keep the wheels of empirical sociology endlessly turning. Paradoxically, however, although the manual/non-manual model is felt to be highly serviceable for research purposes, it is not commonly represented as a model of class cleavage and conflict. That is to say, the two main social categories distinguished by sociology for purposes of class analysis are not invested with antagonistic properties comparable to those accorded to proletariat and bourgeoisie in Marxist theory. This would be less cause for comment if proponents of the manual/non-manual model normally construed the social order as a harmonious and integrated whole; but to construe it instead in terms of conflict, dichotomy, and cleavage, as most of these writers now appear

to do, seems to reveal an awkward contrast between the empirical model of class and the general conception of capitalist society.

The strongest case that could be made out for identifying the line between manual and non-manual labour as the focal point of class conflict would be one that treated capitalist society as the industrial firm writ large. It is only within the framework of 'factory despotism' that the blue-collar/white-collar divide closely corresponds to the line of social confrontation over the distribution of spoils and the prerogatives of command. And this is particularly the case in those industrial settings where even the lowest grades of white-collar staff are cast in the role of managerial subalterns physically and emotionally removed from the shop-floor workers. Within the microcosm of capitalism represented by the typical industrial firm, the sociological model of class has something to recommend it as an alternative to one constructed around the rights of property.

The drawback is, however, that social relations within the capitalist *firm* are a less accurate guide to class relations within capitalist *society* than they might once have been. The reason for this is that the post-war expansion of the public sector has given rise to an ever-increasing assortment of non-manual groups in local government and welfare services that cannot in any real sense be thought

of as the tail-end of a broad managerial stratum aligned against a manual workforce. Frequently, in fact there is no manual workforce to confront in the occupational settings within which these white-collar groups are employed.¹ And even where teachers, social workers, nurses, local government clerks, lower civil servants, and the like do form part of an organization that includes janitors, orderlies, cleaners, and other workers by hand, they do not usually stand in the same quasi-managerial relationship to them as does the staff employee to the industrial worker in the capitalist firm.

The usual rationale for treating intermediate and lower white-collar groups as a constituent element of a dominant class is that these groups traditionally have identified themselves with the interests of capital and management rather than with the interests of organized labour. But for various reasons this identification is easier to accomplish in the sphere of private industry and commerce than in the public sector. In the latter, as already pointed out, not only is there usually no subordinate manual group physically present to inspire a sense of white-collar status elevation, but also the charms of management are likely to seem less alluring when the chain of command stretches ever upwards and out of sight into the amorphous and unlovely body of the state. Moreover, public sector employees do not have the same opportunities as those in the commercial sector for transferring their special skills and services to different and competing employers; all improvements in pay and conditions must be negotiated with a monopoly employer, and one who is under close budgetary scrutiny. All this makes for a relationship of some tension between white-collar employees and the state *qua* employer, a condition more akin to that found between manual labour and management than between white-collar employees and management in the private sector. Thus, the validity of the

manual/non-manual model as a representation of class conflict relies more heavily upon a view of the commercial employee as the prototypical case of the white-collar worker than really is justified, given the enormous growth of public-sector employment.

What this suggests is that manual and non-manual groups can usefully be thought of as entities socially differentiated from each other in terms of life-chances and opportunities, but not as groups standing in a relationship of exploiter and exploited, of dominance and subordination, in the manner presumably required of a genuine conflict model. Expressed differently, the current sociological model does not fulfil even the minimal Weberian claim that the relations between classes are to be understood as ‘aspects of the distribution of power’. Instead of a theoretical framework organized around the central ideas of mutual antagonism and the incompatibility of interests we find one organized around the recorded facts of mere social differentiation. . . .

The ‘Boundary Problem’ in Marxism

The variety of [Marxist] interpretations on offer make it more than usually difficult to speak of ‘the’ Marxist theory of class. In some respects the range of differences within this camp has tended to blur the simple contrast between Marxist and bourgeois theories; and this is particularly so given the tendency for Marxists to adopt familiar sociological categories under substitute names. The most striking example of this is the tacit acknowledgment of the role of *authority* in the determination of bourgeois status. This arises from the need to find some theoretical principle by which the managerial stratum, in particular, can be assigned to the same class as the owners of capital. Although allusions may occasionally be made to the fact that managers are sometimes shareholders in the companies that employ them, it is clear that

this is a contingent feature of managerial status and could not be regarded as theoretically decisive. Managers with and without private company shares do not appear to be different political and ideological animals.

The exercise of discipline over the workforce, on the other hand, is a necessary feature of the managerial role, not a contingent one; and as such it recommends itself as a major criterion of bourgeois class membership. Indeed, for some Marxists managerial authority has in certain respects superseded property ownership as *the* defining attribute of a capitalist class. According to Carchedi, 'the manager, rather than the capitalist rentier, is the central figure, he, rather than the capitalist rentier, is the non-labourer, the non-producer, the exploiter. He, rather than the capitalist rentier, is capital personified.'²

Interestingly, by proclaiming that the supervision and control of subordinates is the new hallmark of bourgeois status, Marxist theorists have come surprisingly close to endorsing Dahrendorf's view of the determinate role of authority in establishing the class boundary.³ Their strict avoidance of this term in favour of some synonym or circumlocution ('mental labour', 'global function of capital', 'labour of superintendence') is perhaps a tacit admission of this embarrassing affinity with Dahrendorf's position. Although none of these writers would accept Dahrendorf's proposition that authority is a general phenomenon that encompasses property, it is nevertheless the case that their treatment of authority relations, however phrased, takes up far more of their analysis than the discussion of property relations.

To make property the centrepiece of class analysis would bring with it the duty of explaining precisely why the apparatus of managerial authority and control was thought to grow out of the institution of private ownership. Presumably it has come to the attention of western Marxists that so-

cieties that have done away with property in its private forms nevertheless have their own interesting little ways of seeing to the 'superintendence of labour'. The view that class and authority relations under capitalism are a unique product of private ownership must rest on a belief that these things are ordered in a very different way under the socialist mode of production. The fact that this mode of production figures not at all in any of the class analyses referred to suggests that Marxists are none too happy about drawing the very comparisons that are so essential to their case. After all, supposing it was discovered that factory despotism, the coercive uses of knowledge, and the privileges of mental labour were present not only in societies where the manager was 'capital personified', but also in societies where he was the party personified? Marxists would then be faced with the unwelcome choice of either having to expand the definition of capitalism to embrace socialist society, or of disowning the cherished concepts of private property and surplus extraction upon which their class theory is grounded. The obvious reluctance to engage in the comparative analysis of class under the two ostensibly different modes of production is therefore understandable enough. As for the credibility of Marxist class theory, it would seem that the advent of socialist society is about the worst thing that could have happened to it.

A further difficulty encountered by this theory is the attempt to arrive at some general principles by which to demarcate the established professions from routine white-collar employees, a distinction required by the evident self-identification of the former with the general interests of the bourgeoisie. In place of any general principles, however, resort is had to an eclectic assortment of descriptive indices demonstrating that 'higher' white-collar groups are in various ways simply better off than 'lower'

white-collar groups. Braverman, for example, lists advantages such as higher pay, security of employment, and the privileged market position of the professions.⁴ In similar vein, Westergaard and Resler suggest drawing a line of class demarcation beneath professional and managerial groups on the grounds that ‘they are not dependent on the markets in which they sell their labour in anything like the way that other earners are’.⁵ Their incomes ‘are determined by market rules and mechanisms over which, in effect, they themselves have considerable influence in their own corners of the market’.⁶

The one notable thing about this kind of analysis is that despite its avowedly Marxist provenance it is indistinguishable from the approach of modern bourgeois social theory. It is, after all, Weber rather than Marx who provides the intellectual framework for understanding class in terms of market opportunities, life-chances, and symbolic rewards. The focus upon income differences and other market factors is difficult to reconcile with the standard Marxist objection to bourgeois sociology that it mistakenly operates on the level of distribution instead of on the level of productive relations. It might also be said that it is from Weber rather than Marx that the postulated link between class position and bureaucratic authority most clearly derives. The fact that these normally alien concepts of authority relations, life-chances, and market rewards have now been comfortably absorbed by contemporary Marxist theory is a handsome, if unacknowledged, tribute to the virtues of bourgeois sociology. Inside every neo-Marxist there seems to be a Weberian struggling to get out. . . .

Social Closure

By social closure Weber means the process by which social collectivities seek to maximize rewards by restricting access to resources and opportunities to a limited circle of eligibles.

This entails the singling out of certain social or physical attributes as the justificatory basis of exclusion. Weber suggests that virtually any group attribute—race, language, social origin, religion—may be seized upon provided it can be used for ‘the monopolization of specific, usually economic opportunities’.⁷ This monopolization is directed against competitors who share some positive or negative characteristic; its purpose is always the ‘closure of social and economic opportunities to *outsiders*’.⁸ The nature of these exclusionary practices, and the completeness of social closure, determine the general character of the distributive system.

Surprisingly, Weber’s elaboration of the closure theme is not linked in any immediate way with his other main contributions to stratification theory, despite the fact that processes of exclusion can properly be conceived of as an aspect of the distribution of power, which for Weber is practically synonymous with stratification. As a result, the usefulness of the concept for the study of class and similar forms of structured inequality becomes conditional on the acceptance of certain refinements and enlargements upon the original usage.

An initial step in this direction is to extend the notion of closure to encompass other forms of collective social action designed to maximize claims to rewards and opportunities. Closure strategies would thus include not only those of an exclusionary kind, but also those adopted by the excluded themselves as a direct response to their status as outsiders. It is in any case hardly possible to consider the effectiveness of exclusion practices without due reference to the countervailing actions of socially defined ineligibles. As Weber acknowledges: ‘Such group action may provoke a corresponding reaction on the part of those against whom it is directed’.⁹ In other words, collective efforts to resist a pattern of dominance governed by exclusion principles can properly be regarded as the

other half of the social closure equation. This usage is in fact employed by Weber in his discussion of ‘community closure’ which, as Neuwirth has shown, bears directly upon those forms of collective action mounted by the excluded—that is, ‘negatively privileged status groups’.¹⁰

The distinguishing feature of exclusionary closure is the attempt by one group to secure for itself a privileged position at the expense of some other group through a process of subordination. That is to say, it is a form of collective social action which, intentionally or otherwise, gives rise to a social category of ineligibles or outsiders. Expressed metaphorically, exclusionary closure represents the use of power in a ‘downward’ direction because it necessarily entails the creation of a group, class, or stratum of legally defined inferiors. Countervailing action by the ‘negatively privileged’, on the other hand, represents the use of power in an upward direction in the sense that collective attempts by the excluded to win a greater share of resources always threaten to bite into the privileges of legally defined superiors. It is in other words a form of action having usurpation as its goal. *Exclusion* and *usurpation* may therefore be regarded as the two main generic types of social closure, the latter always being a consequence of, and collective response to, the former.¹¹

Strategies of exclusion are the predominant mode of closure in all stratified systems. Where the excluded in their turn also succeed in closing off access to remaining rewards and opportunities, so multiplying the number of substrata, the stratification order approaches the furthest point of contrast to the Marxist model of class polarization. The traditional caste system and the stratification of ethnic communities in the United States provide the clearest illustrations of this closure pattern, though similar processes are easily detectable in societies in which class formation is paramount. Strategies of usurpation vary in scale from those designed

to bring about marginal redistribution to those aimed at total expropriation. But whatever their intended scale they nearly always contain a potential challenge to the prevailing system of allocation and to the authorized version of distributive justice.

All this indicates the ease with which the language of closure can be translated into the language of power. Modes of closure can be thought of as different means of mobilizing power for the purpose of engaging in distributive struggle. To conceive of power as a built-in attribute of closure is at the very least to dispense with those fruitless searches for its ‘location’ inspired by Weber’s more familiar but completely unhelpful definition in terms of the ubiquitous struggle between contending wills. Moreover, to speak of power in the light of closure principles is quite consistent with the analysis of class relations. Thus, to anticipate the discussion, the familiar distinction between bourgeoisie and proletariat, in its classic as well as in its modern guise, may be conceived of as an expression of conflict between classes defined not specifically in relation to their place in the productive process but in relation to their prevalent modes of closure, exclusion and usurpation, respectively. . . .

In modern capitalist society the two main exclusionary devices by which the bourgeoisie constructs and maintains itself as a class are, first, those surrounding the institutions of property; and, second, academic or professional qualifications and credentials. Each represents a set of legal arrangements for restricting access to rewards and privileges: property ownership is a form of closure designed to prevent general access to the means of production and its fruits; credentialism is a form of closure designed to control and monitor entry to key positions in the division of labour. The two sets of beneficiaries of these state-enforced exclusionary practices may thus be thought of as the core components of the dominant class under

modern capitalism. Before taking up the discussion of common class interests fostered by private property and credentials it may be useful to consider each of the two principal closure strategies separately.

It has already been remarked upon how the concept of property has been devalued in the modern sociology of class as a result of the heavy weighting accorded to the division of labour. This has not always been true of bourgeois sociology. Weber was in full accord with Marx in asserting that “‘Property’ and ‘lack of property’ are . . . the basic characteristics of all class situations’.¹² The post-Weberian tendency to analyse social relations as if the propertyless condition had painlessly arrived is perhaps a natural extension of the use of ‘western’ or ‘industrial’ to denote societies formerly referred to as capitalist. The post-war impact of functionalist theory certainly contributed to this tendency, since the proclamation of belief in the ultimate victory of achievement values and the merit system of reward naturally cast doubt on the importance of property as an institution. The inheritance of wealth after all requires notably little expenditure of those talents and efforts that are said to be the only keys to the gates of fortune.

The extent to which property has come to be regarded as something of an embarrassing theoretical anomaly is hinted at in the fact that it receives only the most cursory acknowledgment in Davis and Moore’s functionalist manifesto, and even then in the shape of an assertion that ‘strictly legal and functionless ownership . . . is open to attack’ as capitalism develops.¹³ To propose that the imposition of death duties and estate taxes constitutes evidence for an assault upon property rights is somewhat like suggesting that the introduction of divorce laws is evidence of state support for the dissolution of the family. Property in this scheme of things can only be understood as a case of cultural lag—one of those quaint institutional rem-

nants from an earlier epoch which survives by the grace of social inertia.

Several generations earlier Durkheim had reasoned along similar lines in declaring that property inheritance was ‘bound up with archaic concepts and practices that have no part in our present day ethics’.¹⁴ And although he felt it was not bound to disappear on this account he was willing to predict that inherited wealth would ‘lose its importance more and more’, and if it survived at all it would only be ‘in a weakened form’.¹⁵ Durkheim was not of course opposed to private property as such, only its transmission through the family. ‘It is obvious that inheritance, by creating inequalities amongst men from birth, that are unrelated to merit or services, invalidates the whole contractual system at its very roots’.¹⁶ Durkheim wanted society made safe for property by removing those legal practices that could not be squared with conceptions of liberal individualism and which therefore threatened to cause as much moral and social disturbance as the ‘forced’ division of labour.

There was not much likelihood of property itself declining as an institution because it was part of the order of things invested with a sacred character, understood in that special Durkheimian sense of an awesome relationship rooted deeply in the *conscience collective*. Although the sacred character of property arose originally from its communal status, the source of all things holy, the marked evolutionary trend towards the individualization of property would not be accompanied by any decline in its divinity. Personal rights to property were therefore seen by Durkheim as part of that general line of social development by which the individual emerges as a distinct and separate entity from the shadow of the group. The individual affirms himself as such by claiming exclusive rights to things over and above the rights of the collectivity. There is more than an echo here of Hegel’s dictum that ‘In his property a

person exists for the first time as reason'.¹⁷ As Plamenatz comments:

'It makes sense to argue, as Hegel does, that it is partly in the process of coming to own things, and to be recognised as their owners, that human beings learn to behave rationally and responsibly, to lead an ordered life. It is partly in the process of learning to distinguish mine from thine that a child comes to recognise itself as a person, as a bearer of rights and duties, as a member of a community with a place of its own inside it'.¹⁸

As Plamenatz goes on to say, however plausible as a defence of personal property this may be, as a defence of capitalist property relations it is 'lamentably inadequate'.¹⁹

The reason for this is that Hegel, like Durkheim, and many contemporary sociologists, never clearly distinguishes between property as rights to personal *possessions* and property as capital. Parsons is only one of many who reduces all forms of property to the status of a possession; this is understood as 'a right or a bundle of rights. In other words it is a set of expectations relative to social behaviour and attitudes'.²⁰ If property is simply a specific form of possession, or a certain bundle of rights, then everyone in society is a proprietor to some degree. On this reckoning there can be no clear social division between owners and non-owners, only a gradual, descending scale from those with very much to those with very little. This is well in line with Parsons' usual theoretical strategy of asserting the benign quality of any resource by reference to its widespread distribution. The possession of a toothbrush or an oilfield confers similar rights and obligations upon their owners, so that property laws cannot be interpreted as class laws. As Rose and his colleagues have suggested:

'the ideological significance of such a universalistic and disinterested legal interpretation

of property in modern capitalist society is two-fold. First, as the law protects and recognises *all* private property, and as virtually all members of the society can claim title to *some* such property, it may be claimed that all members of society have some vested interest in the *status quo*. From such a perspective, therefore, it can be argued that, far from representing an irreconcilable conflict of interests, the distribution of property in modern capitalist society gives rise to a commensurability of interests, any differences being variations of degree rather than kind. The office developer, the shareholder, the factory-owner, the householder and even the second-hand car owner may thus be represented as sharing fundamentally common interests, if not identities'.²¹

What the sociological definition of property as possessions interestingly fails to ask is why only certain limited forms of possession are legally admissible. It is patently not the case, for example, that workers are permitted to claim legal possession of their jobs; nor can tenants claim rights of possession to their homes, nor welfare claimants enforceable rights to benefits. Possession in all these cases is pre-empted by the conflicting claims of employers, landlords, and the state respectively, which are accorded legal priority. Although the law may treat the rights of ownership in true universalistic fashion it is silent on the manner by which only some 'expectations' are successfully converted to the status of property rights and others not. . . .

The case for restoring the notion of property into the centre of class analysis is that it is the most important single form of social closure common to industrial societies. That is to say, rights of ownership can be understood not as a special case of authority so much as a specific form of exclusion. As Durkheim expresses it, 'the right of property is the right of a given individual to exclude other individual and collective entities from

the usage of a given thing'.²² Property is defined negatively by 'the exclusion it involves rather than the prerogatives it confers'.²³ Durkheim's reference to *individual* rights of exclusion clearly indicates that once again he has possessions in mind, and that, characteristically, he sees no important distinction between objects of personal ownership, and the control of resources resulting in the exercise of power.

It is clearly necessary to distinguish property as possessions from property as capital, since only the latter is germane to the analysis of class systems. Property as capital is, to paraphrase Macpherson, that which 'confers the right to deny men access to the means of life and labour'.²⁴ This exclusionary right can obviously be vested in a variety of institutional forms, including the capitalist firm, a nationalized industry, or a Soviet enterprise. All these are examples of property that confers legal powers upon a limited few to grant or deny general access to the means of production and the distribution of its fruits. Although personal possessions and capital both entail rights of exclusion, it is only the exclusionary rights embedded in the latter that have important consequences for the life-chances and social condition of the excluded. To speak of property in the context of class analysis is, then, to speak of capital only, and not possessions.

Once property is conceptualized as a form of exclusionary social closure there is no need to become entangled in semantic debates over whether or not workers in socialist states are 'really' exploited. The relevant question is not whether surplus extraction occurs, but whether the state confers rights upon a limited circle of eligibles to deny access to the 'means of life and labour' to the rest of the community. If such exclusionary powers are legally guaranteed and enforced, an exploitative relationship prevails as a matter of definition. It is not of overriding importance to know whether these exclusionary powers are

exercised by the formal owners of property or by their appointed agents, since the social consequences of exclusion are not demonstrably different in the two cases. Carchedi and other neo-Marxists may therefore be quite correct in suggesting that 'the manager is capital personified'; but all that needs to be added is first, that this dictum holds good not only for monopoly capitalism, but for *all*, including socialism, systems in which access to property and its benefices is in the legal gift of a select few; and, second, that it squares far more comfortably with the assumptions of bourgeois, or at least Weberian, sociology than with classical Marxist theory.

Of equal importance to the exclusionary rights of property is that set of closure practices sometimes referred to as 'credentialism'—that is, the inflated use of educational certificates as a means of monitoring entry to key positions in the division of labour. Well before the onset of mass higher education, Weber had pointed to the growing use of credentials as a means of effecting exclusionary closure.

'The development of the diploma from universities, and business and engineering colleges, and the universal clamour for the creation of educational certificates in all fields make for the formation of a privileged stratum in bureaus and offices. Such certificates support their holders' claims for intermarriages with notable families . . . , claims to be admitted into the circles that adhere to "codes of honour", claims for a "respectable" remuneration rather than remuneration for work well done, claims for assured advancement and old-age insurance, and, above all, claims to monopolize social and economically advantageous positions. When we hear from all sides the demand for an introduction of regular curricula and special examinations, the reason behind it is, of course, not a suddenly awakened "thirst for education" but the desire for restricting the supply of these positions

and their monopolization by the owners of educational certificates. Today the “examination” is the universal means of this monopolization, and therefore examinations irresistibly advance’.²⁵

The use of credentials for closure purposes, in the manner elaborated by Weber, has accompanied the attempt by an ever-increasing number of white collar occupations to attain the status of professions. Professionalization itself may be understood as a strategy designed, amongst other things, to limit and control the supply of entrants to an occupation in order to safeguard or enhance its market value. Much of the literature on the professions has tended to stress their differences from workaday occupations, usually accepting the professions’ own evaluation of their singularity in creating rigorous codes of technical competence and ethical standards. It is perfectly possible to accept that the monopolization of skills and services does enable the professions to exercise close control over the moral and technical standards of their members, whilst also endorsing Weber’s judgment that ‘normally this concern for efficient performance recedes behind the interest in limiting the supply of candidates for the benefices and honours of a given occupation’.²⁶

It would seem to be the professions’ anxiety to control the supply side of labour that accounts, in part at least, for the qualifications epidemic referred to by Dore as the ‘diploma disease’.²⁷ This is the universal tendency among professions to raise the minimum standards of entry as increasing numbers of potential candidates attain the formerly scarce qualifications. The growing reliance upon credentials as a precondition of professional candidature is commonly justified by reference to the greater complexity of the tasks to be performed and the consequent need for more stringent tests of individual capacity. Yet Berg’s careful analysis of

these claims was able to turn up no evidence to show that variations in the level of formal education were matched by variations in the quality of work performance.²⁸ Nor was there anything to suggest that professional tasks were in fact becoming more complex such as to justify a more rigorous intellectual screening of potential entrants. Berg’s conclusion, in line with Weber’s, is that credentials are accorded their present importance largely because they simplify and legitimate the exclusionary process. It is on these grounds, among others, that Jencks suggests that ‘the use of credentials or test scores to exclude “have not” groups from desirable jobs can be viewed in the same light as any other arbitrary form of discrimination’.²⁹

Formal qualifications and certificates would appear to be a handy device for ensuring that those who possess ‘cultural capital’ are given the best opportunity to transmit the benefits of professional status to their own children. Credentials are usually supplied on the basis of tests designed to measure certain class-related qualities and attributes rather than those practical skills and aptitudes that may not so easily be passed on through the family line. It is illuminating in this respect to contrast the white-collar professions with the sporting and entertaining professions. What is especially remarkable about the latter is how relatively few of the children of successful footballers, boxers, baseball and tennis stars, or the celebrities of stage and screen have succeeded in reproducing their parents’ elevated status. One reason for this would seem to be that the skills called for in these pursuits are of a kind that must be acquired and cultivated by the individual in the actual course of performance, and which are thus not easily transferred from parent to child. That is, there seems to be no equivalent to cultural capital that can be socially transmitted to the children of those gifted in the performing arts that could give them a head start in the fiercely

competitive world of professional sport and show business. Presumably, if the rewards of professional sport could be more or less guaranteed along conventional career or bureaucratic lines serious proposals would eventually be put forward to limit entry to those candidates able to pass qualifying examinations in the theory of sporting science. This would have the desired effect of giving a competitive edge to those endowed with examination abilities over those merely excelling in the activity itself.³⁰

The reason why professional sports, and the entertainment professions in general, are likely to be resistant to the ‘diploma disease’ offers a further instructive comment upon the nature of the white-collar professions. The supreme advantage of occupational closure based upon credentials is that all those in possession of a given qualification are deemed competent to provide the relevant skills and services for the rest of their professional lives. There is no question of retesting abilities at a later stage in the professional career. The professional bodies’ careful insistence that members of the lay public are not competent to sit in judgement on professional standards effectively means that a final certificate is a meal ticket for life. In the sporting and entertainment professions, by contrast, the skills and abilities of the performers are kept under continuous open review by the public; those who consume the services are themselves the ultimate arbiters of an individual’s competence and hence his market value, as expressed via their aggregate purchasing power. There can be no resort to the umbrella protection of a professional licence when sporting prowess and the ability to entertain are felt to be in decline in the eyes of those who pass collective judgement.

Against this exacting yardstick, then, credentialism stands out as a doubly effective device for protecting the learned professions from the hazards of the marketplace. Not

merely does it serve the convenient purpose of monitoring and restricting the supply of labour, but also effectively masks all but the most extreme variations in the level of ability of professional members, thereby shielding the least competent from ruinous economic punishment. The small irony is that credentialist strategies aimed at neutralizing the competitive effects of the market confer most benefit upon that class that is most prone to trumpet the virtues of a free market economy and the sins of collectivism.

The use of systematic restrictions upon occupational entry has not of course been wholly confined to the white-collar professions. Certain skilled manual trades have adopted similar techniques designed to regulate supply, as in the case of the apprenticeship system or certain forms of the closed shop. Some unskilled occupations such as dock work and market-portering have also sought to restrict entry to the kinsmen of those already employed, though this does not normally guarantee control over the actual volume of labour supply. The crucial difference between these attempts at occupational exclusion by manual trades and those adopted by the professions is that the latter generally seek to establish a *legal monopoly* over the provision of services through licensure by the state. Whereas the learned professions have been remarkably successful in winning for themselves the status of what Weber calls ‘legally privileged groups’, it has been far less common for the manual trades to secure the blessing of the state for their exclusionary tactics. Indeed, the resort to ‘restrictive practices’ on the part of organized labour is commonly condemned as a breach of industrial morality that should be curbed rather than sanctified by law. Presumably the fact that governments have usually been reluctant to legislate formally against such practices is not unrelated to the awkwardness that might arise in drawing legal distinctions between

these practices and the exclusionary devices of the professions, including the profession of law itself.

A further point of difference between professional closure and restrictive practices by trade unions is that the main purpose behind the latter activity has been the attempt to redress in some small part the disadvantages accruing to labour in its uneven contest with capital. Closure by skilled workers has been a strategy embarked upon in the course of struggle against a superior and highly organized opponent, and not primarily with the conscious intent of reducing the material opportunities of other members of the labour force. Credentialism, on the other hand, cannot be seen as a response to exploitation by powerful employers; the learned or free professions were never directly subordinate to an employing class during the period when they were effecting social closure. Their conflict, concealed beneath the rhetoric of professional ethics was, if anything, with the lay public. It was the struggle to establish a monopoly of certain forms of knowledge and practice and to win legal protection from lay interference. The aim was to ensure that the professional-client relationship was one in which the organized few confronted the disorganized many. Under modern conditions, where many professionals are indirectly in the service of the state and occasionally in conflict with the government of the day over pay and conditions, a somewhat better case could perhaps be made for likening the position of professions to that of craft unions, in so far as both could be said to employ closure for purposes of bargaining with a more powerful agency. But however acrimonious relations may become between professional bodies and the state, it is worth noting that the state rarely if ever threatens to take sanctions against professions in the way that would most seriously damage their interests—namely, by rescinding their legal monopoly.

On all these grounds it is necessary to regard credentialism as a form of exclusionary social closure comparable in its importance for class formation to the institution of property. Both entail the use of exclusionary rules that confer benefits and privileges on the few through denying access to the many, rules that are enshrined in law and upheld by the coercive authority of the state. It follows from this that the dominant class under modern capitalism can be thought of as comprising those who possess or control productive capital and those who possess a legal monopoly of professional services. These groups represent the core body of the dominant or exploiting class by virtue of their exclusionary powers which necessarily have the effect of creating a reciprocal class of social inferiors and subordinates. . . .

Class Reproduction

There is a definite tension between the commitment to closure by way of property and credentials on the part of one generation and the desire to pass on benefits to subsequent generations of kith and kin. It is not in the least necessary to deny that most members of the exclusionary class will strive to put their own advantages to the service of their children, while asserting at the same time that bourgeois forms of closure are not exactly tailor-made for self-recruiting purposes. In fact exclusionary institutions formed under capitalism do not seem to be designed first and foremost to solve the problem of class reproduction through the family line. The kinship link can only be preserved as a result of *adaptation* by the bourgeois family to the demands of institutions designed to serve a different purpose; it does not come about as a natural consequence of the closure rules themselves. In systems based on aristocratic, caste, or racial exclusion, families of the dominant group can expect to pass on their privileged status to their own descendants as

a direct result of the closure rules in operation, however socially lethargic those families might be. The bourgeois family, by contrast, cannot rest comfortably on the assumption of automatic class succession; it must make definite social exertions of its own or face the very real prospect of generational decline. In other words, although the typical bourgeois family will certainly be better equipped than most to cope with the closure system on its children's behalf, it must still approach the task more in the manner of a challenge with serious risks attached than as a foregone conclusion. Even when it is successful it must face the prospect of sharing bourgeois status with uncomfortably large numbers of parvenus. What kind of system is this to provoke such anxieties in the breasts of those supposedly in command?

The answer must be that it is a system designed to promote a class formation biased more in the direction of sponsorship and careful selection of successors than of hereditary transmission. Although *both* aims might be held desirable, the first takes ideological precedence over the second, so that succession along kinship lines must be accomplished in conformity with the application of criteria that are ostensibly indifferent to the claims of blood. There is nothing especially bizarre about an arrangement whereby a dominant class relinquishes its children's patrimony in order to ensure that the calibre of its replacements is of the highest possible order. It would only appear strange to those unable to conceive that the attachment to doctrine could ever take precedence over the claims of kinship. As Orwell noted in his discussion of communist party oligarchies:

‘The essence of oligarchical rule is not father-to-son inheritance, but the persistence of a certain world-view and a certain way of life, imposed by the dead upon the living. A ruling group is a ruling group so long as it can nominate its successors. The Party is not con-

cerned with perpetuating its blood but with perpetuating itself’.³¹

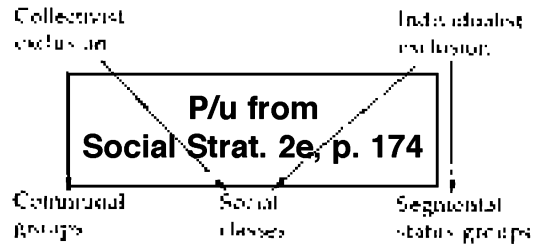
There are also powerful forces in capitalist society that are more dedicated to the perpetuation of bourgeois values than bourgeois blood. Ideological commitment to the rights of property and the value of credentials may be just as fierce as any faith in Leninist party principles. Each represents a set of ideals that can be held quite irrespective of the consequences upon the family fortunes of their advocates. The party militant's belief in a system of political selection and exclusion that could tell against his own ideologically wayward children has its counterpart in the liberal's belief in the validity of meritocratic criteria that would find against his not too clever offspring. It was perhaps examples of this kind that Weber had in mind when referring to patterns of closure distinguished by a ‘rational commitment to values’. The same idea is also more than hinted at in Marx's well-known assertion that the bourgeoisie always puts the interests of the whole class above the interests of any of its individual members. These priorities are not, presumably, reversed whenever the individual members in question happen to be someone's children.

To suggest that predominant forms of closure under modern capitalism are in some tension with the common desire to transmit privileges to one's own is to point up politically significant differences of interpretation of bourgeois ideology. The classical liberal doctrine of individualism contains a powerful rejection of those principles and practices that evaluate men on the basis of group or collectivist criteria. The political driving force of individualist doctrines arose in part from the opposition of the emergent middle classes to aristocratic pretensions and exclusiveness centred around the notion of descent. The emphasis upon lineage was an obvious hindrance to those who had raised

themselves into the ranks of property by way of industry and commerce, but who lacked the pedigree necessary to enter the charmed circles inhabited by those of political power and social honour. Although non-landed wealth could occasionally be cleansed through marriage into the nobility, the new rising class sought to make property respectable in its own right by divorcing it from its associations with particular status groups. Property in all its forms was to become the hallmark of moral worth without reference back, as it were, to the quality of proprietorial blood. In the individualist credo, property thus assumed the same characteristic as money in the marketplace, where the ability to pay overrides all questions as to the actual source of the buyer's cash. . . .

One reason for pressing the distinction between collectivist and individualist criteria underlying all forms of exclusion is to suggest that subordinate classes or strata are likely to differ in their political character according to which of the two sets of criteria is predominant. Looked at in ideal-typical terms, purely collectivist types of exclusion, such as those based on race, religion, ethnicity, and so on, would produce a subordinate group of a communal character—that is, one defined in terms of a total all-encompassing negative status. Blacks under *apartheid* or minority groups herded into religious and racial ghettos are the familiar modern examples. The polar archetypal case would be that of exclusion based solely on individualist criteria, giving rise to a subordinate group marked by intense social fragmentation and inchoateness. The example here is furnished by the model of a pure meritocracy in which class is virtually replaced by a condition of discrete segmental statuses never quite reaching the point of coalescence. In non-fictional societies, of course, individualist and collectivist criteria are usually applied in some combination or other, so producing stratified systems located at various points between

these two extremes. This can be depicted in simplified form as follows:



Thus, of the three major types of subordination, classes are presented as a combination of both types of exclusionary criteria. Schematically, a subordinate class could be located towards either of the opposite poles according to the relative weighting of the two sets of criteria. The proletariat of early and mid-nineteenth century Europe, for example, would approximate to the communal pole by virtue of its wholesale exclusion from civil society arising from the treatment of its members as a *de facto* collectivity. The badge of proletarian status carried with it the kinds of stigmata commonly associated with subordinate racial and ethnic groups. It was a total condition which permitted little leeway for the cultivation of those small part-time identities that bring temporary release from the humiliations of servile status. Correspondingly, of course, the proletarian condition under communal exclusion offered fertile ground for movements and ideologies which raised large questions about the nature of the political order and its legitimacy, and not merely about the fact of unequal shares.

It is the very hallmark of the communal condition that subordination is experienced through a myriad of direct personal degradations and affronts to human dignity, encouraged by the submersion of the individual into the stereotype of his 'membership' group. It is largely as a result of this that the politics of communal exclusion so frequently stresses the need for subordinate groups to create an alternative moral identity to that

fashioned for them by their oppressors. Although the condition of the early proletariat was never completely of a communal kind, it was not so different from that of a despised ethnic group, if only because the visible signs and trappings of status were as unmistakably clear as racial features. Certainly the mixture of horror, fear, and revulsion felt by the upper classes for the great unwashed was not a far remove from the sentiments usually held by dominant racial or ethnic groups towards those whom they simultaneously exploit and despise.

To speak of a gradual shift in the nature of exclusionary rules, from collectivism to individualism, is thus to point to those tendencies making for the progressive erosion of the communal components of proletarian status, otherwise referred to as working-class incorporation into civil society. Although under advanced capitalism labour remains an exploited commodity, the status of the worker does not derive to anything like the same extent from his immersion in a total collective identity and its accompanying rituals of personal degradation. Mills' portrayal of the pattern of 'status cycles' by which the modern urban worker is able to find escape in class anonymity during leisure periods and vacations may be somewhat overdrawn;³² but there is a real sense in which the absence of clearly visible and unambiguous marks of inferior status has made the enforcement of an all-pervasive deference system almost impossible to sustain outside the immediate work situation. It would now take an unusually sharp eye to detect the social class of Saturday morning shoppers in the High Street, whereas to any earlier generation it would have been the most elementary task. More to the point, even assuming that a lynx-eyed bourgeois could accurately spot a worker in mufti, what real hope could he now entertain of having any claim to deference actually honoured? A system of deference can only operate effectively when the status of strangers

can accurately be judged, and the information required for this is difficult to come by without the aid of a collectivist stereotype. In this respect the personal dignity of the modern worker has been enhanced by the evolution towards individualist exclusion, even though his subordination to capital remains a central fact of life.

As class subordination becomes increasingly less communal in character, the political ideals and programmes that flourish among its members tend to become less inspired by visions of a new moral order and the promise of emancipation, and rather more preoccupied with the issues of distributive justice. Those who deplore the apparent flickering of those energies and passions that produced nineteenth-century socialism might care to reflect on the possibility that this has less to do with the iniquities of working-class leadership than with the system of modern exploitation, in which the engines of political resentment are not so lavishly fuelled by the personal degradations arising from wholesale collectivist exclusion. . . .

Conclusion

By way of concluding this part of the discussion, it might be appropriate to offer some general remarks on the explanatory status of the closure model. This model, like any other, recommends the use of a particular sociological vocabulary and an attendant battery of concepts that contain barely disguised moral assumptions about the nature of class society. It is not strictly speaking a 'theory' of class but a way of conceptualizing it that differs from that proposed by other variants of bourgeois sociology or by Marxism. Most of what we conventionally call theories of class are in fact conceptual methods of this kind. They are, for the most part, take-it-or-leave-it moral classifications, not sets of propositions that stand or fall under the impact of

evidence. What conceivable social facts could destroy either the Marxist conception of class as an exploitative relationship, or the liberal conception of class as an exchange relationship? Since conceptual models are ways of presenting social reality, it follows that the preference for one presentation over another entails a personal judgement of some kind about the moral standing of class society.

On this score, the closure model is almost bound to appear defective by liberal and Marxist theorists alike. Liberal theory endorses a contractual view of class, in which the notion of mutual interest and harmony is the essential ingredient. Marxism, on the other hand, assumes not merely the absence of harmony and common class interests, but, more importantly, the presence of irresolvable antagonisms that drive the system to ultimate breakdown. The neo-Weberian position advanced here is that the relation between classes is neither one of harmony and mutual benefit, nor of irresolvable and fatal contradiction. Rather, the relationship is understood as one of mutual antagonism and permanent *tension*; that is, a condition of unrelieved distributive struggle that is not necessarily impossible to 'contain'. Class conflict may be without cease, but it is not inevitably fought to a conclusion. The competing notions of harmony, contradiction, and tension could thus be thought of as the three broad possible ways of conceptualizing the relation between classes, and on which all class models are grounded.

Since class models are not subject to direct empirical assault, the case for advancing the cause of one in preference to another rests partly on the claim that it draws attention to a set of problems and issues that are otherwise obscured. Thus, one of the attractions of the closure model is that it highlights the fact of communal cleavage and its relationship to class, and seeks to analyse both within the same conceptual framework. More generally, it proposes that intra-class

relations be treated as conflict phenomena of the same general order as inter-class relations, and not as mere disturbances or complications within a 'pure' class model. Hence the extension of the concept of exploitation to cover both sets of phenomena. There is, in addition, a recommendation that social classes be defined by reference to their mode of collective action rather than to their place in the productive process or the division of labour. The reason for this is that incumbency of position in a formally defined structure does not normally correspond to class alignment where it really counts—at the level of organized political sentiment and conduct. This serious lack of fit between all positional or systemic definitions of class and the actual behaviour of classes in the course of distributive struggle, is not due to any lack of the categories employed. It arises from the initial theoretical decision to discount the significance and effect of variations in the cultural and social make-up of the groups assigned to the categories in question. Models constructed upon such formal, systemic definitions require of their advocates much ingenuity in accounting for the continuous and wholesale discrepancies between class position and class behaviour. A good deal of the intellectual energy of western Marxism has been dissipated in wrestling with this very problem which is of its own conceptual making.

NOTES

1. The hospital setting provides, perhaps, the most important exception. Industrial conflicts between medical staff and the manual workers' unions over issues such as 'pay beds' are unusual in having clear-cut ideological, rather than bread-and-butter, causes.
2. Carchedi 1975:48. For Braverman, too, managers and executives are 'part of the class that personifies capital . . .' (1974:405).
3. Dahrendorf 1959.
4. Braverman 1974: Chapter 18.

5. Westergaard and Resler 1975:92.
6. Westergaard and Resler 1975:346.
7. Weber (eds Roth and Wittich) 1968:342.
8. Weber (eds Roth and Wittich) 1968:342.
9. Weber (eds Roth and Wittich) 1968:342.
10. Neuwirth 1969.
11. These arguments were first tentatively sketched out in my 'Strategies of Social Closure in Class Formation' (Parkin 1974). In that publication the two types of closure were referred to as *exclusion* and *solidarism*. This latter term does not, however, satisfactorily describe a mode of collective action standing in direct opposition to exclusion, since solidaristic behaviour can itself be used for blatantly exclusionary ends. That is to say, solidarism does not properly refer to the purposes for which power is employed. The term *usurpation* more adequately captures the notion of collective action designed to improve the lot of a subordinate group at the expense of a dominant group. Solidarism is simply one means among others to this end.
12. Weber (eds Gerth and Mills) 1946:182.
13. Davis and Moore 1945:247.
14. Durkheim 1957:174.
15. Durkheim 1957:175 and 217.
16. Durkheim 1957:213.
17. Plamenatz 1975:120.
18. Plamenatz 1975:121.
19. Plamenatz 1975:121.
20. Parsons 1951:119. The entry in the index under 'Property' invites the reader to 'see Possessions'.
21. Rose *et al.* 1976:703.
22. Durkheim 1957:142.
23. Durkheim 1957:142.
24. Macpherson 1973.
25. Weber (eds Gerth and Mills) 1946:241–42.
26. Weber (eds Roth and Wittich) 1968:344.
27. Dore 1976.
28. Berg 1973.
29. Jencks 1972:192.
30. It transpires that the idea is not so far-fetched after all. The Council for National Academic Awards has recently approved the syllabus for a BA Degree in Sports Studies. Undergraduates will be instructed in 'the variables influencing performance in sport; a science and its sports application; scientific methods, statistics and computing; and wide practical experience in a number of sports.' *Daily Telegraph*, Monday, 28 August 1978, p. 3.

31. Orwell 1949:215.
32. Mills 1956:257–58.

BIBLIOGRAPHY

- Berg, I. (1973) *Education and Jobs: The Great Training Robbery*. Harmondsworth: Penguin Books.
- Braverman, H. (1974) *Labor and Monopoly Capital*. New York: Monthly Review Press.
- Carchedi, G. (1975) On the Economic Identification of the New Middle Class. *Economy and Society* 4 (1).
- Dahrendorf, R. (1959) *Class and Class Conflict in Industrial Society*. London: Routledge.
- Davis, K. and Moore, W. E. (1945) Some Principles of Stratification. *American Sociological Review* X (2).
- Dore, R. (1976) *The Diploma Disease*. London: Allen and Unwin.
- Durkheim, E. (1957) *Professional Ethics and Civic Morals*. London: Routledge.
- Jencks, C. (1972) *Inequality*. New York: Basic Books.
- Macpherson, C. B. (1973) A Political Theory of Property. In *Democratic Theory: Essays in Retrieval*. Oxford University Press.
- Mills, C. W. (1956) *White Collar*. New York: Oxford University Press.
- Neuwirth, G. (1969) A Weberian Outline of a Theory of Community: Its Application to the 'Dark Ghetto'. *British Journal of Sociology* 20 (2).
- Orwell, G. (1949) *Nineteen Eighty-Four*. London: Secker and Warburg.
- Parkin, F. (ed.) (1974) *The Social Analysis of Class Structure*. London: Tavistock.
- Parsons, T. (1951) *The Social System*. London: Routledge.
- Plamenatz, J. (1975) *Karl Marx's Philosophy of Man*. Oxford: Clarendon Press.
- Rose, D., Saunders, P., Newby, H., and Bell, C. (1976) Ideologies of Property: A Case Study. *Sociological Review* 24 (4).
- Weber, M. (1946) *From Max Weber*. Gerth, H. H. and Mills, C. W. (eds) New York: Oxford University Press.
- . (1968) *Economy and Society*. Roth, G. and Wittich, C. (eds) New York: Bedminster Press.
- Westergaard, J. and Resler, H. (1975) *Class in a Capitalist Society*. London: Heinemann.

■ DURKHEIM AND POST-DURKHEIMIANS ■

16 ■ Emile Durkheim

The Division of Labor in Society

In *The Division of Labor in Society*, we emphasise the state of legal and moral anomie in which economic life exists at the present time. In fact, in this particular sphere of activity, professional ethics only exist in a very rudimentary state. There are professional ethics for the lawyer and magistrate, the soldier and professor, the doctor and priest, etc. Yet if we attempted to express in somewhat more precise terms contemporary ideas of what should be the relationship between employer and white-collar worker, between the industrial worker and the factory boss, between industrialists in competition with one another or between industrialists and the public, how imprecise would be the statements that we could formulate! Some vague generalities about the loyalty and commitment that employees of every kind owe to those who employ them, or about the moderation that employers should manifest in exercising their economic superiority, a certain condemnation of any competition that is too blatantly unfair, or of any too glaring exploitation of the consumer: this is almost the sum total of what the ethical consciousness of these professions comprises. Moreover, most of these precepts lack any juridical character. They are backed only by public opinion and not

by the law—and it is well known how indulgent that opinion shows itself to be about the way in which such vague obligations are fulfilled. Those actions most blameworthy are so often excused by success that the boundary between the permissible and the prohibited, between what is just and what is unjust, is no longer fixed in any way, but seems capable of being shifted by individuals in an almost arbitrary fashion. So vague a morality, one so inconsistent, cannot constitute any kind of discipline. The upshot is that this entire sphere of collective life is for the most part removed from the moderating action of any rules.

It is to this state of anomie that must be attributed the continually recurring conflicts and disorders of every kind of which the economic world affords so sorry a spectacle. For, since nothing restrains the forces present from reacting together, or prescribes limits for them that they are obliged to respect, they tend to grow beyond all bounds, each clashing with the other, each warding off and weakening the other. . . .

Political society as a whole, or the state, clearly cannot draw up the system of rules that is now lacking. Economic life, because it is very special and is daily becoming increasingly specialised, lies outside their authority

and sphere of action. Activity within a profession can only be effectively regulated through a group close enough to that profession to be thoroughly cognisant of how it functions, capable of perceiving all its needs and following every fluctuation in them. The sole group that meets these conditions is that constituted by all those working in the same industry, assembled together and organised in a single body. This is what is termed a corporation, or professional group.

Yet in the economic field the professional group no more exists than does a professional ethic. Since the last century when, *not without reason*, the ancient corporations were dissolved, hardly more than fragmentary and incomplete attempts have been made to reconstitute them on a different basis. Doubtless, individuals who are busy in the same trade are in contact with one another by the very fact that their activities are similar. Competition with one another engenders mutual relationships. But these are in no way regular; depending upon chance meetings, they are very often entirely of an individual nature. One industrialist finds himself in contact with another, but the body of industrialists in some particular speciality do not meet to act in concert. Exceptionally, we do see all members of the same profession come together at a conference to deal with some problem of common interest. But such conferences last only a short while: they do not survive the particular circumstances that gave rise to them. Consequently the collective life for which they provided an opportunity dies more or less entirely with them. . . .

Can we legitimately believe that corporative organisation is called upon to play in contemporary societies a more considerable part? If we deem it indispensable it is not because of the services it might render the economy, but on account of the moral influence it could exercise. What we particularly see in the professional grouping is a moral force capable of curbing individual egoism, nurturing among workers a more envigo-

rated feeling of their common solidarity, and preventing the law of the strongest from being applied too brutally in industrial and commercial relationships. Yet such a grouping is deemed unfit for such a role. Because it springs from temporal interests, it can seemingly only serve utilitarian ends, and the memories that survive of the corporations during the *ancien régime* only confirm this impression. We incline to visualise them in the future as they were towards the end of their former existence, intent above all on maintaining or increasing their privileges and monopolies. We fail to see how such narrow vocational concerns might have any beneficial effect upon the morality of the corporation or its members.

However, we should refrain from extending to the entire corporative system what may have been true of certain corporations during a very short period in their development. Far from the system having been, because of its very constitution, infected by a kind of moral sickness, during the greater part of its existence it played above all a moral role. This is especially evident with the Roman corporation. ‘Among the Romans,’ declares Walzing, ‘the corporations of artisans were far from having so pronounced a professional character as in the Middle Ages. We come across no regulations concerning methods, no obligatory apprenticeship, and no monopoly. Nor was their purpose to accumulate the capital necessary to exploit an industry.’¹ Doubtless their associating together gave them more power to safeguard the common interest, when the need arose. But this was only one of the useful by-products that the institution engendered. It was not the justification for its existence, nor its main function. Above all else, the corporation was a collegiate religious body. Each one possessed its own particular god, who, when the means were available, was worshipped in a special temple. Just as every family had its *Lar familiaris* and every city its *Genius publicus*, so every collegiate body had its protect-

ing divinity, the *Genius collegii*. Naturally this professional form of worship was not without its festivities, and sacrifices and banquets were celebrated in common together. Moreover, all kinds of circumstances would serve as the occasion for festive gatherings; distribution of food and money was often made at the expense of the community. . . .

The facts cited adequately demonstrate that a professional grouping is not at all incapable of exerting a moral effect. The very important place that religion held in its life highlights very particularly the true nature of its functions, for in such times every religious community constituted a moral environment, just as every kind of moral discipline necessarily tended to take on a religious form. Moreover, this characteristic of corporative organisation is due to the effect of very general causes which we can see at work in different circumstances. Within a political society, as soon as a certain number of individuals find they hold in common ideas, interests, sentiments and occupations which the rest of the population does not share in, it is inevitable that, under the influence of these similarities, they should be attracted to one another. They will seek one another out, enter into relationships and associate together. Thus a restricted group is gradually formed within society as a whole, with its own special features. Once such a group is formed, a moral life evolves within it which naturally bears the distinguishing mark of the special conditions in which it has developed. It is impossible for men to live together and be in regular contact with one another without their acquiring some feeling for the group which they constitute through having united together, without their becoming attached to it, concerning themselves with its interests and taking it into account in their behaviour. And this attachment to something that transcends the individual, this subordination of the particular to the general interest, is the very well-spring of all moral activity. Let this sentiment only cry-

tallise and grow more determinate, let it be translated into well-defined formulas by being applied to the most common circumstances of life, and we see gradually being constituted a corpus of moral rules.

Domestic morality did not arise any differently. Because of the prestige that the family retains in our eyes, if it appears to us to have been and continue to be a school of altruism and abnegation, the highest seat of morality, it is through the very special characteristics it is privileged to possess, ones that could not be found at any level elsewhere. We like to believe that in blood kinship there exists an extraordinarily powerful reason for moral identification with others. But, as we have often had occasion to show,² blood kinship has in no way the extraordinary effectiveness attributed to it. The proof of this is that in a large number of societies relations not linked by the blood tie are very numerous in a family. Thus so-called artificial kinship is entered into very readily and has all the effects of natural kinship. Conversely, very frequently those closely knit by ties of blood are morally and legally strangers to one another. For example, this is true of blood kin in the Roman family. Thus the family does not derive its whole strength from unity of descent. Quite simply, it is a group of individuals who have drawn close to one another within the body politic through a very specially close community of ideas, feelings and interests. Blood kinship was able to make such a concentration of individuals easier, for it naturally tends to have the effect of bringing different consciousnesses together. Yet many other factors have also intervened: physical proximity, solidarity of interest, the need to unite to fight a common danger, or simply to unite, have been causes of a different kind which have made people come together.

Such causes are not peculiar to the family but are to be found, although in different forms, within the corporation. Thus if the former group has played so important a role in the moral history of humanity, why

should not also the latter be capable of so doing? Undoubtedly one difference will always exist between them, inasmuch as family members share in common their entire existence, whereas the members of a corporation share only their professional concerns. The family is a kind of complete society whose influence extends to economic activity as well as to that of religion, politics, and science, etc. Everything of any importance that we do, even outside the home, has repercussions upon it and sparks off an appropriate reaction. In one sense the corporation's sphere of influence is more limited. Yet we must not forget the ever more important place that our profession assumes in our lives as work becomes increasingly segmented. . . .

What past experience demonstrates above all is that the organisational framework of the professional group should always be related to that of economic life. It is because this condition was not fulfilled that the system of corporations disappeared. Thus, since the market, from being municipal as it once was, has become national and international, the corporation should assume the same dimensions. Instead of being restricted exclusively to the artisans of one town, it must grow so as to include all the members of one profession scattered over the whole country,³ for in whatever region they may be, whether they live in town or countryside, they are all linked to one another and share a common life. Since this common life is in certain respects independent of any territorial boundaries, a suitable organism must be created to give expression to this life and to regulate its functions. Because of the dimensions that it assumes, such an organism should necessarily be closely in contact and directly linked with the central organism of the life of the collectivity. Events important enough to affect a whole category of industrial enterprises within a country necessarily have wide repercussions of which the state cannot fail to be aware. This impels it to intervene. Thus for good reason the royal power tended instinc-

tively not to leave large-scale industry outside its ambit as soon as it appeared. It could not fail to take an interest in a form of activity which by its very nature is always liable to affect society as a whole. Yet such regulatory action, although necessary, should not degenerate into utter subordination, as happened in the seventeenth and eighteenth centuries. The two organisms, although in contact with each other, should remain distinct and autonomous; each has functions that it alone can perform. If it falls to political assemblies to lay down the general principles for industrial legislation, they are not capable of diversifying them according to the various types of industry. It is this diversification that is the corporation's proper task.⁴ A unitary organisation over a whole country also in no way precludes the formation of secondary organisations which include similar workers in the same region or locality. Their role could be to spell out even more specifically, in accordance with local or regional needs, the regulations for a profession. Thus economic activity could be regulated and demarcated without losing any of its diversity.

Moreover, we must reject the belief that the corporation's sole role should consist in laying down and applying rules. It is undoubtedly true that wherever a group is formed, a moral discipline is also formed. But the institution of that discipline is only one of the numerous ways in which any collective activity manifests itself. A group is not only a moral authority regulating the life of its members, but also a source of life *sui generis*. From it there arises a warmth that quickens or gives fresh life to each individual, which makes him disposed to empathise, causing selfishness to melt away. Thus in the past the family has been responsible for legislating a code of law and morality whose severity has often been carried to an extreme of harshness. But it has also been the environment where, for the first time, men have learnt to appreciate the outpouring of feel-

ing. We have likewise seen how the corporation, both in Rome and during the Middle Ages, created these same needs and sought to satisfy them. The corporations of the future will be assigned even greater and more complex functions, because of their increased scope. Around their purely professional functions will be grouped others which at present are exercised by the communes and private associations. Among these are functions of mutual assistance which, in order to be entirely fulfilled, assume between helpers and helped feelings of solidarity as well as a certain homogeneity of intellect and morals, such as that readily engendered by the exercise of the same profession. Many educational activities (technical education, adult education, etc.) should also, it seems, find in the corporation their natural habitat. The same is also true for a certain type of artistic activity. It would seem in accordance with the nature of things that such a noble form of diversion and recreation should develop alongside the more serious aspects of life, acting as a balancing and restorative influence. In fact we now already see trade unions acting at the same time as friendly societies, and others are setting up communal centres where courses are organised, and concerts and dramatic performances held. Hence the activity of a corporation can take on the most varied forms.

We may even reasonably suppose that the corporation will be called upon to become the foundation, or one of the essential foundations, of our political organisation. We have seen that, although it first began outside the social system, it tended to become more and more closely involved in it as economic life developed. We have therefore every reason to anticipate that, if progress continues on the same lines, the corporation is destined to assume an ever more central and preponderant place in society. It was once the elementary division of communal organisation. Now that the commune, from being the autonomous unit that it once was, has been ab-

sorbed into the state just as the municipal market was absorbed into the national market, may we not legitimately think that the corporation should also undergo a corresponding transformation and become the elementary division of the state, the basic political unit? Society, instead of remaining what it is today—a conglomerate of land masses juxtaposed together—would become a vast system of national corporations. The demand is raised in various quarters for electoral colleges to be constituted by professions and not by territorial constituencies. Certainly in this way political assemblies would more accurately reflect the diversity of social interests and their interconnections. They would more exactly epitomise social life as a whole. Yet if we state that the country, in order to become conscious of itself, should be grouped by professions, is not this to acknowledge that the organised profession or the corporation should become the essential organ of public life?

In this way a serious gap in the structure of European societies, and in our own in particular, would be filled. We shall see how, as history unfolds, an organisation based on territorial groupings (village, town, district or province, etc.) becomes progressively weaker. There is no doubt that we each belong to a commune or a *département*, but the ties binding us to them become daily more loose and tenuous. These geographical divisions are in the main artificial, and no longer arouse deep emotions within us. The provincial spirit has vanished beyond recall. ‘Parish pump’ patriotism has become an anachronism that cannot be restored at will. Strictly local or *département* matters hardly affect or enthrall us either any longer, save in so far as they go hand in hand with matters relating to our profession. Our activity extends much beyond these groups, which are too narrow for it; moreover, much of what happens within them leaves us indifferent. Thus what might be described as the spontaneous collapse of the old social structure has occurred.

But this internal organisation cannot disappear without something taking its place. A society made up of an extremely large mass of unorganised individuals, which an overgrown state attempts to limit and restrain, constitutes a veritable sociological monstrosity. For collective activity is always too complex to be capable of finding expression in the one single organ of the state. Moreover, the state is too remote from individuals, its connections with them too superficial and irregular, to be able to penetrate the depths of their consciousness and socialise them from within. This is why, when the state constitutes the sole environment in which men can fit themselves for the business of living in common, they inevitably ‘contract out,’ detaching themselves from one another, and thus society disintegrates to a corresponding extent. A nation cannot be maintained unless, between the state and individuals, a whole range of secondary groups are interposed. These must be close enough to the individual to attract him strongly to their activities and, in so doing, to absorb him into the mainstream of social life. We have just demonstrated how professional groupings are fitted to perform this role, and how indeed everything marks them out for it. Hence we can comprehend how important it is, particularly in the economic sphere, that they should emerge from that inchoate and disorganised state in which they have lain for a century, since professions of this kind today absorb the greater part of the energies of society.⁵

NOTES

1. Waltzing, *Etude historique sur les corporations profession chez les Romains*, vol. I, p. 194.

2. Cf. especially *Année sociologique*, vol. I, pp. 313 ff.

3. We need not discuss the international organisation which, because of the international character of the market, would necessarily develop at a level above that of the national organisation. For at present the latter alone can constitute a legal entity. In the present state of European law the former can only result from arrangements freely concluded between national corporations.

4. This specialisation could not occur without the help of elected assemblies charged with representing the corporation. In the present state of industry, these assemblies, as well as those tribunals entrusted with the task of applying the regulations of an occupation, should clearly include representatives of employees and employers, as is already the case with the industrial arbitration tribunals. The proportion of each should correspond to the respective importance attributed by public opinion to these two factors of production. But if it is necessary for both sides to meet on the governing councils of the corporation it is no less indispensable for them to constitute distinct and independent groups at the lower level of corporative organisation, because too often their interests vie with one another and are opposing. To feel that they exist freely, they must be aware of their separate existence. The two bodies so constituted can then appoint their representatives to the common assemblies.

5. Moreover, we do not mean that territorial constituencies are destined to disappear completely, but only that they will fade into the background. Old institutions never vanish in the face of new ones to such an extent that they leave no trace of themselves. They persist not only by the mere fact of survival, but also because there persists some trace of the needs to which they corresponded. Material proximity will always constitute a link between men. Consequently the political and social organisation based on territory will certainly subsist. But it will no longer enjoy its present predominance, precisely because that link is losing some of its force. What is more, we have shown above that, even at the base of the corporation will still be found geographical divisions. Moreover, between the various corporations from a same locality or region there will necessarily be special relationships of solidarity which will, from time to time, demand an appropriate organisation.

17 ■ David B. Grusky and Jesper B. Sørensen

Are There Big Social Classes?

The study of social class has a volatile history in which waves of creative class analytic scholarship are interspersed with periods of cynicism about the class analytic enterprise. In the present cynical phase, criticisms of both Marxian and non-Marxian class analysis continue to escalate, with many commentators now feeling bold enough to argue that the concept of class is “ceasing to do any useful work for sociology” (Pahl 1989, p. 710; also, Pakulski and Waters 1996; Clark and Lipset 1991). By way of response, the most ardent defenders of class models have simply reaffirmed the class analytic status quo, albeit sometimes with the concession that class-based formulations now apply in rather weakened form (e.g., Wright 1996; Hout, Brooks, and Manza 1993; Goldthorpe and Marshall 1992). The debate between these two camps has proceeded along stylized lines. Indeed, although the literature is well stocked with all manner of defense and critique of conventional class analysis, there have been few, if any, truly constructive efforts to refashion class analysis.

Against this intellectual backdrop, we have recently suggested that critics of class analysis have too quickly dismissed the power of class analytic language, whereas defenders of class analysis have not appreciated that such language, for all its power, yields little insight when applied to conventional, highly aggregate social classes (see Grusky and Sørensen 1998; Grusky 1999; Grusky and Weeden

2001). This formulation leads to the prescription that class analysis should be ratcheted down to an analytic level where real social groupings (i.e., “occupations”) form around functional niches in the division of labor. The great virtue of disaggregating is that the nominal categories of conventional class analysis can be replaced by *Gemeinschaftlich* groupings that are embedded in the very fabric of society and are thereby meaningful not merely to sociologists but to the lay public as well.

The foregoing line of argument is not entirely without precedent. Indeed, whenever sociologists have turned their attention to the professions (e.g., Abbott 1988), the longstanding tendency has been to emphasize the great heterogeneity and sectional divisiveness within this (putative) new class. The recent commentary of Freidson (1986) is illustrative here: “The range of education, income, and prestige of the professional occupations in question . . . [makes] it hard to imagine them sharing a common culture of any significance, a common set of material interests, or a common inclination to act politically in the same fashion and direction” (p. 57). Although this critique is surely of interest, it falls short of our own position insofar as it pertains only to the professional sector and fails to engage more broadly with contemporary anticlass critiques. In similar fashion, stratification scholars are currently quite interested in “unpacking” conventional class categories (Marshall et al.

1988), yet the ultimate objective has invariably been to argue for some new and preferred form of *reaggregation*.

We would be hard-pressed, then, to locate a direct line of intellectual heritage. If forced to identify a partial and approximate one, the principal inspiration would have to be scholars such as Durkheim ([1893] 1933), Bourdieu (1984), and their intellectual descendants (e.g., Lamont 1992; also, Freidson 1994; Van Maanen and Barley 1984). Under the Durkheimian developmental model, occupational associations come to serve as important intermediaries between the modern state and individual, yet they play a largely integrative role and eschew the more partisan behavior of “maintaining or increasing privileges and monopolies” (Durkheim [1893] 1933, p. 10). Likewise, Bourdieu (1984) has argued that sociologists should “rethink Weber’s opposition between class and *Stand*” (p. xii), but his recent empirical work emphasizes the cultural rather than economic implications of occupational closure. This work is nonetheless distinguished by its relatively detailed analyses; that is, Bourdieu resorts to disaggregate data in characterizing the habitus and the lifestyles it generates, because the conditions of existence of conventional aggregate classes are assumed to be unacceptably heterogeneous. In our own analysis, we shall similarly insist on extreme disaggregation, yet we regard the resulting occupations as economic and cultural groupings that constitute precisely that unification of “class and *Stand*” that Bourdieu (1984) so ambitiously sought.

The Case for Disaggregation

The following discussion summarizes the main virtues of disaggregation in understanding patterns of class identification, social closure, collective action, and lifestyles and attitudes. For each of these topics, our summary of the conceptual rationale for disaggregation will be brief, as more comprehensive analyses can be found elsewhere (see

Grusky and Sørensen 1998; Grusky 1999; Grusky and Weedon 2001).

Class Identification

We can usefully begin by considering the subjective domain of stratification as revealed in patterns of class identification and awareness. Although some sociologists remain convinced that contemporary identities are strongly shaped by aggregate affiliations (e.g., Marshall et al. 1988), the prevailing post-Marxist position is that conventional classes now have only a weak hold over workers. For example, Emmison and Western (1990) report that only 7 percent of all Australians regard their social class as a “very important” identity, whereas other commentators (e.g., Saunders 1989) have stressed that open-ended queries about class identification tend to yield confused responses, refusals to answer, and even explicit denials that classes exist. This evidence has led many sociologists to conclude that class is now a “passive identity” (Bradley 1996, p. 72) and that the realm of production is no longer the principal locus of identity formation.

We regard such accounts as overreactive to concerns that, although legitimate, surely do not require abandoning class analysis altogether. The Emmison–Western results are again revealing on this point, because they indicate that detailed occupations continue to be one of the main social identities for contemporary workers (see Emmison and Western 1990, pp. 247–48). This result should come as no surprise; after all, occupational categories are deeply embedded in the institutions of advanced industrialism, whereas aggregate classes are highly abstract constructs that are evidently more appealing to academics than to workers, employers, or the state. As Treiman (1977) notes, workers invariably represent their career aspirations in occupational terms, while professional and vocational schools train workers for occupationally defined skills, and employers construct and advertise jobs in terms of corresponding occupational designations. The

class analytic fallacy thus amounts to insisting on aggregate categories even when disaggregate ones are more deeply institutionalized and hence subjectively more salient.

Social Closure

If subjectivist models of class were once dominant in sociology (e.g., Warner, Meeker, and Eells 1949), they have now been superseded by analytic approaches that focus on the social processes by which class membership is restricted to qualified eligibles (Freidson 1994, pp. 80–84; Murphy 1988; Collins 1979; Parkin 1979; Weber [1922] 1968). These models emphasize not only the institutionalized means by which closure is secured (e.g., private property, credentials, licenses) but also the efforts of excluded parties to challenge these institutions and the inequality that they maintain. Although closure theory provides, then, a new sociological language for understanding interclass relations, the actual class mappings posited by closure theorists have proven to be standard aggregate fare. The two-class solution proposed, for example, by Parkin (1979, p. 58) features an exclusionary class comprising those who control productive capital and professional services and a subordinate class comprising all those who are excluded from these positions of control.

We might usefully ask whether an aggregate formulation is fundamental to closure theory or merely superfluous adjunct. The latter interpretation strikes us as more plausible; that is, if closure theory could somehow be reinvented without the coloration of class analytic convention, its authors would likely emphasize that the real working institutions of closure (i.e., professional associations, craft unions) are largely local associations “representing the credential-holders themselves” (Murphy 1988, p. 174). These associations establish and enforce local jurisdictional settlements that prevent other occupations from providing competing services. In most cases, the associated closure devices (e.g., licensing,

credentialling, apprenticeships) do not govern entry to aggregate classes, but instead serve only to control entry (and exit) at the more detailed occupational level. By contrast, there are no analogous organizations that represent aggregate classes, nor are there jurisdictional settlements or closure devices that are truly aggregate in scope. This conclusion implies that conventional aggregate mappings of “exploitation classes” (e.g., Wright 1985) conceal the highly disaggregate level at which rent is extracted and interests are formed (see Sørensen 1996; 1994). Indeed, given that unions and associations establish local rather than class-wide restrictions on labor supply, the “rent” that is thereby generated should create interests principally at the disaggregate level.

Collective Action

For most neo-Marxists, social closure is of interest not because it provides a vehicle for pursuing purely local concerns (i.e., “trade union consciousness”), but rather because it allegedly facilitates the development of classwide interests and grander forms of interclass conflict. The aggregate classes identified by contemporary sociologists have so far shown a decided reluctance to act in accord with such theorizing. This quiescence at the aggregate level has led to considerable neo-Marxian handwringing as well as more radical claims that postmodern interests are increasingly defined and established outside the realm of production (e.g., Laraña, Johnston, and Gusfield 1994). The latter form of postmodernism, popular as it is, overlooks the simple fact that much collective action flows unproblematically out of structurally defined groupings, albeit only when those groupings are defined in less aggregate terms than is conventionally the case. The three principal types of collective action at the level of unit occupations are (a) downwardly directed closure strategies designed to restrict access to occupational positions, (b) lateral competitive struggles between occupational associations over functional niches in the division of labor, and (c) upwardly directed collective action oriented

toward securing occupation-specific benefits (e.g., monopoly protection) from the state and from employers. We thus concur with Krause (1971, p. 87) that “there has historically been more occupation-specific consciousness and action than cross-occupational combination” (also, see Freidson 1994, pp. 75–91).

This is not to suggest that local conflict at the unit occupational level drives the course of human history. To the contrary, local associations typically pursue sectional objectives, and the wider systemic effects of such micro-level conflict are neither obvious nor necessarily profound (cf. Durkheim [1893] 1933). We might conclude, then, that our disaggregate class analysis is an intellectually modest project, but it bears noting that aggregate class analysts have likewise scaled back their ambitions and effectively discarded comprehensive class-based theories of history (e.g., Goldthorpe and Marshall 1992, p. 385).

Class Outcomes

In this sense, the class analytic project is becoming gradually more limited in its objectives, with many contemporary scholars now satisfied to merely document that class membership conditions individual-level outcomes of all kinds (e.g., attitudes, voting behavior, lifestyles). The resulting analyses typically examine either the categorical effects of aggregate classes or the gradational effects of variables that represent the many dimensions (e.g., socioeconomic status, substantive complexity) underlying disaggregate occupations. Although these approaches have yielded new and important results, it is nonetheless troubling that they typically conceal or ignore the *Gemeinschaftlich* character of (some) disaggregate occupations. If modern closure is indeed secured principally at the detailed occupational level, then the resulting restriction of social interaction will generate and maintain occupational subcultures that are correspondingly disaggregate. These local cultures are initially forged through intensive secondary socialization of the kind provided in appren-

ticeships, police and military academies, and graduate and professional schools. As Caplow (1954) noted long ago, many occupations require prolonged training that serves to inculcate explicit codes of behavior, whereas aggregate classes have no comparable influence or authority over secondary patterns of socialization. The occupational habitus is further strengthened insofar as workers choose occupations receptive to their values and employers choose workers with values that are (putatively) compatible with occupational demands. The great failing of conventional analyses of lifestyles, dispositions, and attitudes is that *Gemeinschaftlich* occupations are regarded as nominal categories and are therefore blithely aggregated or dimensionalized.

The moral to our story, then, is that sociologists have searched for structuration at the wrong level of analysis. Ironically, class analysts have sought realist solutions at the aggregate level when only nominal ones were viable, whereas occupational analysts have settled on nominal solutions (e.g., socioeconomic scales) when in fact realist ones were feasible. Among Marxian and non-Marxian scholars alike, the division of labor is typically represented as purely “technical” in character (see, esp., Wright 1980; Abercrombie and Urry 1983, p. 109), even though nominal task-based groupings are often converted into real social collectivities with a shared culture and set of interests. We think that sociological research stands to benefit from taking such local organization more explicitly into account.

Issues of Trend

Although disaggregate structuration has been largely overlooked by contemporary class analysts, it is nonetheless possible that such structuration, strong though it may be, is growing gradually weaker in ways that are consistent with a standard poststructuralist vision. The prevailing view, especially among European commentators, is that the site of production is indeed of diminishing rele-

vance in understanding stratification systems. The virtues of poststructuralism may be “taken for granted among contemporary social and cultural analysts” (Casey 1995, p. 8), but the lack of substantiating evidence for this position is quite striking; and it is accordingly premature to foreclose on all further debate about the principal forces of change.

The available literature on such matters can be readily simplified by classifying theories in terms of the institutional domains that they reference. We shall thus proceed by distinguishing between (a) the types of technical tasks embodied in the division of labor, (b) the organizational settings in which these tasks are carried out, and (c) the associational forms that characteristically develop at the site of production (e.g., trade unions, professional associations). As shall be evident, the foregoing domains do not evolve in isolation from one another, but it is still analytically useful to distinguish between them.

Sociotechnical Change

The current fashion is to approach longstanding debates about sociotechnical change from a post-Fordist perspective (Piore and Sabel 1984). As Amin (1994) notes, post-Fordists suggest that early industrialism brought about much craft deskilling and homogenization, yet this process is alleged to be reversing itself as “Fordist” factories are gradually supplanted by small-scale production, flexible specialization, and a rejuvenated artisanal sector, all of which serve to reintroduce those distinctions of manual labor that Marx ([1894] 1964) promised would ultimately disappear. This account may therefore be seen as a freshened form of postindustrial theory in which the forces of upgrading and reskilling are presumed to play out not merely in the professional, technical, and service categories but in the craft sector as well. In this context, one might expect postmodernists to view post-Fordism with some antipathy, yet in fact these two accounts are often conflated in the literature. For exam-

ple, the “new times” post-Fordism of Hall (1988) and his colleagues (e.g., Hall and Jacques 1989) becomes virtually indistinguishable from conventional postmodernism, as it emphasizes that sociotechnical changes weaken aggregate solidarities and generate a new stratification order based on “lifestyle, taste, and culture rather than categories of social class” (Hall 1988, p. 24). This account rests on the characteristic postmodernist assumption that an “increasingly fragmented” productive realm (Hall 1988, p. 24) necessarily weakens *all* forms of solidarity within the division of labor.

The pathbreaking work of Piore and Sabel (1984) clearly has merit, but we would necessarily take issue with these more elaborated accounts that attempt to smuggle in poststructuralism under a post-Fordist banner. If one accepts the core post-Fordist claim that flexible specialization breathes new life into artisanal production (e.g., Piore and Sabel 1984), the appropriate implication is not that all production-based solidarities shall wither away but rather that such solidarities are increasingly localistic. In any standard post-Fordist account, the new and emerging forms of craft production are assumed to require worker “solidarity and communitarianism” (Piore and Sabel 1984, p. 278), and the rejuvenated artisanal sector therefore brings early-industrial “craft communities into the twenty-first century on the basis of a newly decentralized production process” (Aronowitz and DiFazio 1994, p. 98). The end result, then, is a manifestly prostructuralist account whereby modern craftworkers are increasingly “bound to an often familial community [that] promotes both greater control and a sense of belonging” (Aronowitz and DiFazio 1994, p. 97).

The same conclusion holds with respect to older sociotechnical models of differentiation (e.g., Parsons 1970; Dahrendorf 1959). When such models were initially formulated, there was little interest in elaborating a positive theory of local structuration, because the principal objective was merely to counter Marxian

approaches by calling attention to the class-decomposing effects of differentiation. If a positive theory of local solidarities *were* attempted, it would likely emphasize that (a) the process of differentiation generates local collective action as emergent occupational groupings vie with one another for jurisdiction over new functional niches, and (b) the resulting occupations become meaningful communities not only because of the “mechanical solidarity” spawned by functional similarities (Durkheim [1893] 1933, p. 16) but also because of the affiliative ties forged in the originating jurisdictional struggles (see Abbott 1988).

The above considerations suggest that differentiation creates solidarities that are increasingly localistic. At the same time, one must bear in mind that the newly differentiated occupations are, by virtue of their newness, hampered in developing stereotypical behavioral expectations that can then be enforced by the outside public. The subcultures of these occupations may therefore be less binding; for instance, the occupations of “systems analyst,” “day trader,” or “Web site designer” may not evoke stereotypical expectations that are as well formed as those characterizing more established occupations, such as professor (absentminded), cook (excitable), or reporter (cynical). In his seminal work, Caplow (1954, pp. 134–35) makes much of this liability of newness, as he regarded the “public stereotype as itself the most important agent for the conditioning of roles.” The latter argument fails, however, to appreciate that newness can itself be an asset; indeed, just as religious cults generate solidarity by capitalizing on missionary zeal, so too one suspects that new occupations can impose behavioral expectations on incumbents without these expectations being well known or appreciated by outsiders. Under this formulation, rapid differentiation prevents the public from understanding the increasingly complex mosaic of occupational subcultures and communities, but it may not greatly weaken the hold of these local communities over their members.

Organizational Change

As sociologists so frequently point out, the division of labor is not intrinsic to the structure of tasks, but rather is a social construction that reflects organizational constraints as well as the interests of relevant parties (e.g., Edwards 1979). The rise of industrialism in the eighteenth century can be attributed, for example, to the spread of vertical strategies of coordination that fragmented tasks into increasingly simple jobs and thus rendered them amenable to purely administrative or bureaucratic oversight (Weber [1922] 1968). By contrast, preindustrial craft workers defined and organized the production process themselves, and the division of labor was accordingly controlled by self-regulating occupational experts rather than organizationally empowered administrators (Zabusky and Barley 1996, pp. 188–92). The obvious question that then arises is whether vertical methods of control will continue to encroach on occupationally defined labor as postindustrialism evolves.

This question cannot be as easily answered as some postmodernists seemingly suggest. In fact, one can identify incipient organizational theories on either side of this debate, with the contemporary literature thus encompassing both (a) postoccupational theories describing the gradual withering away of functionally defined positions, and (b) revisionist theories suggesting that a new occupationally oriented logic of production is on the rise. The former literature, which is clearly dominant, rests on the claim that contemporary organizations are relying increasingly on teamwork, cross-training, and multiactivity jobs that break down conventional skill-based distinctions (e.g., Casey 1995). These new polyvalent jobs are created either by combining formerly distinct skills or by appending managerial and coordinative functions to production positions. The resulting story thus privileges the forces of integration over those of differentiation; that is, whereas many early industrial craft occu-

pations (e.g., shoemaker) were dissolved through task simplification, the postindustrial organizations described by Casey (1995) putatively eliminate occupations through task fusion, elaboration, and complication.

The preceding account, popular though it may be, is not without its critics, some of whom have argued that “pressures for an occupational logic of organizing may in fact be rising” (Barley 1995, p. 40). This revisionist argument rests on the twofold claim that (a) the occupationally organized sectors of the labor force (e.g., professionals) are rapidly expanding in size, and (b) the remaining vertically organized sectors of the labor force (e.g., management) are increasingly differentiating into functional areas and therefore becoming “occupationalized” (Freidson 1994). In developing these claims, Barley (1995) suggests that the seeds of the future have been sown in the burgeoning technical sector, where the work process is dominated by experts who have so far rigorously defended their occupational jurisdictions and have accordingly resisted cross-training, job mergers, and all forms of hierarchy. The resulting “technicist archetype” (Barley 1995) thus rests on the collaboration of experts who control knowledge through extended training within a community of practice. Under the latter formulation, teams and work groups figure no less prominently than in the postoccupationalist archetype (e.g., Casey 1995), but of course the constituent experts now control mutually exclusive bodies of knowledge. The resulting team solidarity may be seen, then, as organic rather than mechanical.

Although most expert teams are presently formed within the confines of firms, one might anticipate that production will increasingly be contracted out to independent workers who are brought together by managers or brokers. The construction industry serves as the conventional exemplar here both because of its extreme occupationalism and characteristic reliance on outsourcing.

In fact, the emerging fashion among organizational theorists is to represent the construction industry not as a historical remnant that “God forgot and the industrial revolution overlooked” (Lawrence and Dyer 1983, p. 599), but rather as a heroic survivor that will in the end supersede mass production, thereby shaping the future of work more generally. For our purposes, it suffices to stress that these revisionist theories are inconsistent with those of postoccupationalism, and not merely because they rest explicitly on a well-developed (and intensifying) division of labor. We would further emphasize that the concomitant growth of outsourcing and externalization increases pressures to identify and affiliate with occupations rather than organizations.

Associational Change

The final institutions of interest to us are the various intermediary associations (e.g., trade unions, professional associations) that characteristically develop at the site of production. Within the Marxian framework, the longstanding concern has been that “trade union consciousness” is intrinsically sectional, thus requiring intellectuals and party functionaries to carry out supplementary ideological work that presumably cultivates more encompassing class-based interests (esp. Lenin 1927). This Marxian concern appears now to have been well founded. If the history of guilds, unions, and related production-based associations is reevaluated from the long view, it is evident that true classwide organization emerged for only a brief historical moment and that postmodern forms are reverting back to localism and sectionalism. The widely documented difficulties facing contemporary unions should be interpreted accordingly; namely, despite an evident weakening in the “encompassiveness of union movements” (Visser 1988, p. 167), there is much evidence suggesting that purely local unions and associations have by no means lost their hold over workers (e.g., the

American Federation of Teachers). In many countries, centralized bargaining between national unions and employers is indeed on the decline, yet decentralized negotiations have taken their place as “instrumental collectivism, based on sectional self-interest, becomes the order of the day” (Marshall et al. 1988, p. 7). This interpretation, if borne out, does not speak to destructurement per se but rather to increasing disaggregation and differentiation of associational forms.

The professional sector has given rise to organizational forms that are yet more localistic. As Parkin (1979) points out, professionals eschew all types of interoccupational confederation, whereas they typically seek out sectional associations that can defend jurisdictional claims and thereby protect against incursions by neighboring occupations. In assessing the future of professionalization, one must consider not only the ongoing growth of traditional professional occupations (e.g., lawyer) and the consequent increase in the number of workers who find themselves in classlike groupings, but also the emergence of new high-skill sectors that may allow further occupations to undertake professionalization projects. To be sure, oppositional movements may possibly emerge and stall these closure projects, yet there is relatively little in the contemporary political arena that might now be interpreted as incipient antiprofessionalism. This conclusion serves to emphasize our larger point that the future of local solidarities is more ambiguous than standard poststructuralist formulas allow.

Conclusions

In his celebrated preface to *The Division of Labor*, Durkheim ([1893] 1933, p. 28) predicted that occupational associations would gradually become “intercalated between the state and the individual,” thereby providing an organizational counterbalance to the threat of class formation on one hand and

state tyranny on the other. This account is ritually rehearsed by Durkheimian scholars but has never been treated as a credible developmental model. As the Marxian project falls out of favor, scholars have therefore settled into some version of Weberianism or postmodernism, neither of which pays much attention to occupation-level structuration. We have outlined above a quasi-Durkheimian third road that refocuses attention on local forms of structuration within the division of labor.

In laying out this case for disaggregation, we have largely ignored cross-national variability in local structuration, but not because we believe such variability to be either trivial or inconsequential. To the contrary, we suspect that convergence theories (e.g., Erikson and Goldthorpe 1992) may be rather less appealing when disaggregate analyses are attempted, because national idiosyncrasies are necessarily concealed through the abstracting and aggregating operations of class analysis. The case of Germany, for example, provides a revealing example of the extent to which local institutional forms can support and sustain disaggregate structuration. As class analysts have long stressed, Germany has a well-developed system of vocational training and apprenticeship, both of which serve to encourage occupation-specific investments and promote professional commitment and craftsmanship (e.g., Blossfeld 1992). In systems of this sort, workers must invest in a single trade early in their careers, and the correspondingly high costs of retraining produce relatively closed occupational groupings.

If the German system reveals, then, the limits of disaggregate structuration, the case of Japan conversely reveals the extent to which such structuration can be institutionally suppressed. The standard characterization of Japan emphasizes such distinguishing features as (a) an educational curriculum that is generalist in orientation rather than functionally differentiated, (b) a vocational training

system that cultivates firm-specific “*nenko* skills” (Dore 1973) through teamwork and continuous job rotation, (c) an organizational commitment to “lifetime employment” that further strengthens firm-specific ties at the expense of more purely occupational ones, and (d) a weakly developed system of enterprise unions that cuts across functional specializations and thereby eliminates any residual craft-based loyalties. This conjunction of forces produces a postoccupational system that some commentators might well regard as prototypically postmodern.

Although further cross-national comparisons of the preceding sort would surely be instructive, we think that comparative analysis becomes especially powerful when local and aggregate forms of structuration are considered in tandem. In the past, structuration has been treated as a unidimensional concept, and scholars accordingly sought to characterize countries on a simple continuum representing the extent to which their stratification systems were well formed (cf. Giddens 1973). The two cases discussed above suggest that such practice may not be altogether misleading; after all, Japan is well known for its attenuated class structure as well as its postoccupationalism (Nakane 1970), while Germany likewise combines strong vocationalism with a deeply class-based labor market and political system. We would nonetheless caution against assuming that such cross-level consistency is the norm. In fact, low-level structuration is often assumed to *undermine* the development of class-based organization, with the United States serving as the typical case in point. The scholarly literature on American exceptionalism is obviously wide ranging, but one of the continuing themes is that class formation was inhibited in the American case not so much by simple individualism as by low-level structuration in the form of craft unions and professionalism (e.g., Dahrendorf 1959). The zero-sum imagery underlying such analyses suggests that

aggregate and disaggregate structuration may sometimes work at cross-purposes.

It is also worth considering the obverse case in which class-based organization flourishes in the absence of competing local structuration. This is clearly the stuff of textbook Marxism, yet ironically it comes closest to empirical realization within countries, such as Sweden, that opted for the social democratic road quite early. In standard analyses of Swedish exceptionalism (e.g., Therborn 1988), the well-known solidarism of labor is attributed not merely to the historic weakness of guild organization and craft unionism, but also to party negotiating tactics that privileged classwide collective bargaining over purely sectional wage demands. At the same time, the “active labor market” programs embodied in the Rehn-Meidner model (Esping-Andersen 1988, pp. 47–53) provide extensive state assistance for worker retraining and relocation, thereby blurring interoccupational boundaries and further undermining local sectionalism and closure. In this context, unit-level occupations are still defined by functional positions in the Swedish division of labor, but the social trappings (e.g., associations, closure) that usually emerge around such technical distinctions have been partly repressed. Although Sweden appears, then, to be properly characterized by the neo-Marxian formula that “technical features do not entail social features” (Abercrombie and Urry 1983, p. 109), it is unclear whether this form of structuration extends much beyond Sweden and Scandinavia more generally. If it is more widespread than we suspect, then our preferred line of argumentation is admittedly weakened.

The larger conclusion to be drawn is that sociologists in all countries have typically been too quick to fall back on purely nominal categories and the descriptive models that they imply. The longstanding Marxian distinction between *klasse an sich* and *klasse für sich* only reinforces such nominalist tendencies, as it legitimates the claim that conventional aggregate

categories, although presently latent or quiescent, may someday become meaningful and activated. This approach is of course peculiarly modern. In characterizing stratification systems of the past, sociologists have typically relied on categories that were embedded in the fabric of society (e.g., estates, castes), thereby rendering them sensible and meaningful to intellectuals and the lay public alike.

The modern analogues to such realist categories are the unit occupational groups that emerge around functional positions in the division of labor. If analyses are ratcheted down to this level, we can construct models that rely on real institutional forces and assume more nearly structural form. The proof of our approach rests, then, on the additional explanatory power and understanding that accrues from referencing the real institutional processes that create classes, constrain mobility chances, generate earnings, and define lifestyles. The task of mapping disaggregate stratification is hardly trivial, but the intellectual payoff to so proceeding is likely to be greater than that secured by carrying out yet another study at the aggregate level.

REFERENCES

- Abbott, Andrew. 1988. *The System of Professions: An Essay on the Division of Expert Labor*. Chicago: University of Chicago Press.
- Abercrombie, Nicholas, and John Urry. 1983. *Capital, Labour, and the Middle Classes*. London: Allen & Unwin.
- Amin, Ash. 1994. "Post-Fordism: Models, Fantasies, and Phantoms of Transition." Pages 1–39 in *Post-Fordism: A Reader*, edited by Ash Amin. Oxford: Blackwell.
- Aronowitz, Stanley, and William DiFazio. 1994. *The Jobless Future*. Minneapolis: University of Minnesota Press.
- Barley, Stephen R. 1995. "The Technician as an Occupational Archetype: A Case for Bringing Work into Organizational Studies." Working Paper, Stanford University.
- Blossfeld, Hans-Peter. 1992. "Is the German Dual System a Model for a Modern Vocational Training System?" *International Journal of Comparative Sociology* 33:168–81.
- Bourdieu, Pierre. 1984. *Distinction: A Social Critique of the Judgement of Taste*, translated by Richard Nice. New York: Cambridge University Press.
- Bradley, Harriet. 1996. *Fractured Identities: Changing Patterns of Inequality*. Cambridge: Polity Press.
- Caplow, Theodore. 1954. *The Sociology of Work*. Minneapolis: University of Minnesota Press.
- Casey, Catherine. 1995. *Work, Self, and Society*. London: Routledge.
- Clark, Terry N., and Seymour M. Lipset. 1991. "Are Social Classes Dying?" *International Sociology* 6:397–410.
- Collins, Randall. 1979. *The Credential Society: An Historical Sociology of Education and Stratification*. New York: Academic Press.
- Dahrendorf, Ralf. 1959. *Class and Class Conflict in Industrial Society*. Stanford: Stanford University Press.
- Dore, Ronald P. 1973. *British Factory—Japanese Factory*. London: Allen & Unwin.
- Durkheim, Emile. [1893] 1933. *The Division of Labor in Society*. New York: Macmillan.
- Edwards, Richard. 1979. *Contested Terrain*. New York: Basic Books.
- Emmison, Michael, and Mark Western. 1990. "Social Class and Social Identity: A Comment on Marshall et al." *Sociology* 24:241–53.
- Erikson, Robert, and John F. Goldthorpe. 1992. *The Constant Flux: A Study of Class Mobility in Industrial Societies*. Oxford: Clarendon Press.
- Esping-Andersen, Gösta. 1988. "The Making of a Social Democratic Welfare State." Pages 35–66 in *Creating Social Democracy: A Century of the Social Democratic Labor Party in Sweden*, edited by Klaus Misgeld, Karl Molin, and Klas Amark. University Park, PA: The Pennsylvania State University Press.
- Freidson, Eliot. 1994. *Professionalism Reborn: Theory, Prophecy, and Policy*. Chicago: University of Chicago Press.
- Freidson, Eliot. 1986. *Professional Powers: A Study of the Institutionalization of Formal Knowledge*. Chicago: University of Chicago Press.
- Giddens, Anthony. 1973. *The Class Structure of the Advanced Societies*. London: Hutchinson.
- Goldthorpe, John H., and Gordon Marshall. 1992. "The Promising Future of Class Analysis: A Response to Recent Critiques." *Sociology* 26:381–400.
- Grusky, David B. 1999. "Foundations of Class Analysis: A Durkheimian Perspective." Working paper, Department of Sociology, Cornell University.
- Grusky, David B., and Jesper B. Sørensen. 1998. "Can Class Analysis Be Salvaged?" *American Journal of Sociology* 103:1187–1234.

- Grusky, David B., and Kim A. Weeden. 2001. "Decomposition without Death: A Research Agenda for the New Class Analysis." *Acta Sociologica* 44: 203–18.
- Hall, Stuart. 1988. "Brave New World." *Marxism Today*, October: 24–9.
- Hall, Stuart, and Martin Jacques. 1989. *New Times: The Changing Face of Politics in the 1990s*. London: Lawrence and Wishart.
- Hout, Michael, Clem Brooks, and Jeff Manza. 1993. "The Persistence of Classes in Postindustrial Societies." *International Sociology* 8: 259–77.
- Krause, Elliot A. 1971. *The Sociology of Occupations*. Boston: Little, Brown.
- Lamont, Michele. 1992. *Money, Morals, and Manners: The Culture of the French and American Upper-Middle Class*. Chicago: University of Chicago Press.
- Laraña, Enrique, Hank Johnston, and Joseph R. Gusfield. 1994. *New Social Movements: From Ideology to Identity*. Philadelphia: Temple University Press.
- Lawrence, Paul R., and Davis Dyer. 1983. *Renewing American Industry*. New York: Free Press.
- Lenin, Vladimir I. 1927. *Collected Works of V. I. Lenin*. New York: International Publishers.
- Marshall, Gordon, David Rose, Howard Newby, and Carolyn Vogler. 1988. *Social Class in Modern Britain*. London: Unwin Hyman.
- Marx, Karl. [1894] 1964. *Selected Works: Volume 1*. Moscow: Progress Publishers.
- Murphy, Raymond. 1988. *Social Closure: The Theory of Monopolization and Exclusion*. Oxford: Clarendon.
- Nakane, Chie. 1970. *Japanese Society*. London: Weidenfeld and Nicolson.
- Pahl, R. E. 1989. "Is the Emperor Naked? Some Questions on the Adequacy of Sociological Theory in Urban and Regional Research." *International Journal of Urban and Regional Research* 13:709–20.
- Pakulski, Jan, and Malcolm Waters. 1996. *The Death of Class*. London: Sage Publications.
- Parkin, Frank. 1979. *Marxism and Class Theory: A Bourgeois Critique*. New York: Columbia University Press.
- Parsons, Talcott. 1970. "Equality and Inequality in Modern Society, or Social Stratification Revisited." Pages 13–72 in *Social Stratification: Research and Theory for the 1970s*, edited by Edward O. Laumann. Indianapolis: Bobbs-Merrill Company.
- Piore, Michael J., and Charles F. Sabel. 1984. *The Second Industrial Divide: Possibilities for Prosperity*. New York: Basic Books.
- Saunders, Peter. 1989. "Left Write in Sociology." *Network* 44:3–4.
- Sørensen, Aage B. 1996. "The Structural Basis of Social Inequality." *American Journal of Sociology* 101:1333–65.
- Sørensen, Aage B. 1994. "The Basic Concepts of Stratification Research: Class, Status, and Power." Pp. 229–41 in *Social Stratification: Class, Race, and Gender in Sociological Perspective*, edited by David B. Grusky. Boulder, Colo.: Westview.
- Therborn, Göran. 1988. "A Unique Chapter in the History of Democracy: The Social Democrats in Sweden." Pages 1–34 in *Creating Social Democracy: A Century of the Social Democratic Labor Party in Sweden*, edited by Klaus Misgeld, Karl Molin, and Klas Amark. University Park, PA: The Pennsylvania State University Press.
- Treiman, Donald J. 1977. *Occupational Prestige in Comparative Perspective*. New York: Academic Press.
- Van Maanen, John, and Stephen R. Barley. 1984. "Occupational Communities: Culture and Control in Organizations." *Research in Organizational Behavior* 6:287–365.
- Visser, Jelle. 1988. "Trade Unionism in Western Europe: Present Situation and Prospects." *Labour and Society* 13:125–82.
- Warner, W. Lloyd, Marchia Meeker, and Kenneth Eells. 1949. *Social Class in America*. Chicago: Science Research Associates.
- Weber, Max. [1922] 1968. *Economy and Society*. Berkeley: University of California Press.
- Wright, Erik O. 1996. "The Continuing Relevance of Class Analysis." *Theory and Society* 25:697–716.
- Wright, Erik O. 1985. *Classes*. London: Verso.
- Wright, Erik O. 1980. "Class and Occupation." *Theory and Society* 9:177–214.
- Zabusky, Stacia E., and Stephen R. Barley. 1996. "Redefining Success: Ethnographic Observations on the Careers of Technicians." Pages 185–214 in *Broken Ladders: Managerial Careers in the New Economy*, edited by Paul Osterman. Oxford: Oxford University Press.

18 ■ Kim A. Weeden

Occupational Closure and Earnings Inequality

In contemporary industrialized societies, inequality stems from three logically, although not temporally, distinct processes: positions in the division of labor are differentiated from each other, reward packages of greater or lesser value are attached to these positions, and people are allocated to these differentially rewarded positions. The first and third components continue to captivate students of occupations and of stratification, respectively, whereas the second component, the link between positions and their rewards, has received less concerted theoretical attention. This paper develops and evaluates a group-level approach to inequality that focuses attention on the second component, concentrating on the question, “Why do some occupations pay more than others?” Its basic argument derives from Weber’s (1978 [1922]) notion of social closure and its subsequent elaboration by inequality theorists (e.g., Collins 1979; Murphy 1988; Manza 1992; Parkin 1971, 1974, 1979; Tilly 1998): social groups formed around positions in the technical division of labor (i.e., occupations) create social and legal barriers that restrict “access to resources and opportunities to a limited circle of eligibles” (Parkin 1979, p. 3). These practices, in turn, shape the contours of the hierarchy of occupational rewards.

Despite its pedigree, closure theory has languished in the twin shadows cast by Marxian class analysis and status attainment research. Prior efforts to launch closure the-

ory into the spotlight of stratification research have been hampered by three main failings. First, Weber never offered a unified and well-developed statement of closure (Manza 1992, p. 227), and subsequent scholars borrowed and modified bits and pieces of his fragments to suit their own purposes. Perhaps as a result, closure theory has lacked a coherent account of earnings inequality and, in particular, of the mechanisms that link various closure strategies to earnings.

Second, closure theorists typically search for closure at the level of aggregate classes and all but ignore the deeply institutionalized practices that secure privilege for detailed occupations: educational credentialing, licensure, voluntary certification, representation by occupational associations, and unionization. To be sure, not all occupations are characterized by closure. Some are collections of workers who share little more than an administrative label, while others are still in the process of organizing and yet to evince the institutionalized markers of a successful collective mobility project. Nevertheless, to the extent that visible means of closure can be identified in the contemporary United States, the labor market of concern here, they are found at the occupation level (Grusky and Sørensen 1998). For all its Weberian heritage, the closure model outlined here is thus Durkheimian in its assumption that occupations are the salient units of contemporary social organization.

Third, closure theory has been shackled by the dearth of closure data and, consequently, convincing tests of its core predictions. This paper exploits an innovative data set that codes 488 detailed occupations in the United States by their closure practices. When linked to individual-level earnings data, this data set offers an opportunity to evaluate (a) whether occupational closure affects earnings; and (b) the closure strategies and mechanisms that are most effective in securing an economic payoff to the occupation.

Closure Mechanisms and Strategies

The central prediction of closure theory is straightforward: the greater the extent of closure characterizing an occupation, the higher the occupation's rewards. Closure theorists are typically content to end the story here, or perhaps with the qualification that some closure strategies (e.g., unionization) contradict modern individualist ideologies, generate public disapproval and reactionary backlash, and hence are less effective than others (e.g., Parkin 1979, pp. 74–86). In its current state, closure theory lacks a clear articulation of the mechanisms through which closure affects rewards and therefore can offer little real insight into how or why various closure strategies might differ in their economic and social payoff.

The broader stratification and professions literatures suggest four mechanisms linking occupational closure projects and the reward structure. First, closure may place artificial restrictions on the supply of labor in an occupation, thereby generating rents—wage premiums above that which would be expected under full market competition—for occupation members (e.g., Sørensen 2000). Second, closure projects may enhance overall levels of demand for a product or service, as when occupations lobby the state to expand the provision of particular services (MacDonald 1995) or market their services directly to the con-

sumer. Third, closure projects can channel this newfound demand to a particular occupation and protect an occupation's jurisdictional claim over the service or task (e.g., Abbott 1988). Fourth, closure projects send a quality-of-service signal (e.g., a “profession” or “craft”), thereby triggering normative expectations about the amount of pay practitioners deserve (Brint 1994).

Each of the five institutionalized closure practices affects occupational rewards through a unique combination of these mechanisms. Take, for example, educational credentials, the familiar markers of achievement conferred by the formal education system. By tying an occupation to academic knowledge, educational credentialing signals that the occupation is highly selective and triggers the “public's mistaken belief that abstract professional knowledge . . . implies effective professional work” (Abbott 1988, p. 54). Educational credentialing also restricts the labor supply where, as in the contemporary United States, two conditions obtain: (1) access to the credential is itself restricted, as when the primary educational system fails to prepare all students for occupation-specific training, the postsecondary educational system fails to expand rapidly enough to meet demand, occupational agents restrict the number of slots available in accredited institutions, or genetics conspire to create a dearth of native talent to fill those slots (Sørensen 2000; Weeden and Grusky 2007); and (2) employers value credentials as the symbols or markers of knowledge, whether because these markers indicate the acquisition of real skills, select on desirable traits (e.g., ambition, intelligence), or “merely” signify membership in a favored cultural caste (Collins 1979).

Licenses, which are state-mandated permissions for an individual to identify herself by an occupational title or practice a particular set of skills, operate through supply-side, demand-side, and normative mechanisms. Although licenses are formally granted by the

state, occupations may exert considerable control over the criteria for licensure, as when occupational representatives control licensing boards whose decisions are honored by the state. In this way, occupations can impose restrictions on the number of licensed practitioners. Even where occupational self-regulation is minimal, licensure cannot fail to restrict supply unless criteria are so lax, fees so low, and the process so painless that no potential licensees are discouraged or denied (Kleiner 2006). Licenses can also affect occupational rewards by convincing the public that a particular set of tasks is most efficiently, effectively, or safely performed by the named occupation alone, thereby channeling demand to that occupation. And, by virtue of the selectivity they imply and the embedded assumption that special expertise is required to judge a practitioner's competence, licenses help an occupation lay claim to professional or craft status.

Voluntary certifications trigger much the same symbolic and demand-side mechanisms as licenses, but, because they are legitimated by neither the state nor the formal educational system, they are largely ineffective at restricting the supply of labor in the certified occupation. Such programs do, however, allow occupations to define and claim ownership over their core tasks, presumably affecting a consumer's choice of a practitioner when in need of a service. Moreover, they can help solidify an occupation's claims that its practitioners meet a minimum level of technical competence and moral superiority and provide a service worthy of professional-level pay.

Occupational associations likewise operate purely in the symbolic realm. To be sure, occupational associations may be the identifiable agents behind efforts to develop occupation-specific educational credentials, licenses, or voluntary certification programs, and some of their impact on the reward structure may by these other strategies be mediated. In terms of direct effects on rewards, an occupa-

tional association operates by (1) increasing the overall level of demand for services, as when the association representing plastic surgeons takes out a full page ad in national magazines touting the benefits of altering one's facial features; (2) channeling new or existing demand to that occupation through efforts to protect or expand its task niche, as when the emergent occupation of financial planner competes with insurance agents, lawyers, accountants, and even advice columnists to capture a newfound demand for personal financial planning (Cohen 1996); and (3) influencing consumers' perception of the quality of the occupation's services through the association's involvement in the political economy, self-proclaimed expertise on social issues, organizational rhetoric, or even choice of a name (e.g., Freidson 1986; Caplow 1954).

The final closure strategy considered here, unionization, primarily affects occupational rewards through unions' state-sanctioned right to bargain collectively and to use the threat of withdrawing labor. Unions have a well-documented impact on the wages of both union members and nonunion members in the same firm, industry, geographic labor market, or broad skill group (e.g., Freeman and Medoff 1984). The latter "wage spillover" effect, wherein nonunion employers pay union wages to reduce the outflow of personnel or to undermine unionizing efforts in their own firms, has not been demonstrated at the detailed occupation level, but it is reasonable to assume that such a spillover effect might be present. The complexity, of course, is that although some unions are organized around occupations (e.g., pilots, flight attendants), others are organized around industries or seemingly disparate combinations of occupations (e.g., graduate students and auto workers). Any "occupational" effect of unionization may merely reflect the uneven distribution of occupations across industries as much as an occupational closure effect per se.

Figure 1: Mechanisms of Earnings Enhancement Characterizing Each Closure Strategy

| Closure Strategy | Restrict Supply | Increase Diffuse Demand | Channel Demand to the Occupation | Signal Quality of Service |
|----------------------------------|-----------------|-------------------------|----------------------------------|---------------------------|
| Licensing | + | - | + | + |
| Formal educational credentialing | + | - | - | + |
| Voluntary certification | - | - | + | + |
| Representation by associations | - | + | + | + |
| Unionization | + | - | - | - |

Note: Symbols indicate the presence (+) or absence (-) of the mechanism.

The preceding discussion is summarized in Figure 1, which lists the four mechanisms of closure and their associated closure strategies. Because the observable strategies operate through multiple unobservable mechanisms, disentangling the mechanisms empirically is difficult. However, it is possible to generate testable hypotheses about the relative effects of each closure strategy based on different assumptions about the importance of the underlying mechanism. Specifically, if stratification scholars are correct in assuming that supply-side monopolization has the strongest impact on the reward structure, we would expect voluntary certification and representation by associations to have relatively weak effects on earnings, because neither operates through this mechanism. If, on the other hand, demand-side protections have the strongest impact, we would expect formal educational credentialing to have relatively weak effects.

Methods and Models

The few extant efforts to link licensing, credentialing, unionization, or association representation to occupational earnings have failed to distinguish occupation-level effects on mean earnings from the wage-generating attributes of the individuals nested within

occupations (e.g., Cullen 1978; Stigler 1971; Friedman and Kuznets 1945). Without this distinction, positional accounts become hopelessly entangled with far more popular individual-level explanations. To avoid this problem, I exploit a simple hierarchical model, specified as follows:

$$Y_{ij} = c + \mathbf{a}'\mathbf{W}_j + \mathbf{b}'(\mathbf{X}_{ij} - \overline{\mathbf{X}}_{j..}) + \mu_j + \epsilon_{ij}$$

where Y_{ij} is the (logged) weekly earnings for person i in occupation j ; \mathbf{b}' is a vector of individual-level fixed effects on earnings; \mathbf{X}_{ij} is a vector of individual-level characteristics centered around their grand means; and ϵ_{ij} and μ_j are mean-zero individual- and occupation-level error terms, respectively. Of primary interest is \mathbf{a}' , a vector of coefficients that estimates the association between occupational characteristics, \mathbf{W}_j , and mean (logged) earnings after adjusting the latter for cross-occupational differences in incumbents' attributes.

Data and Variables

Individual-Level Data

The individual-level data are extracted from the earner study sample of the March Current Population Survey (CPS) (Bureau of Labor Statistics [BLS]). The 1992–1999

surveys, chosen to bracket the closure data, are pooled to maximize the number of 1990 Standard Occupation Classification (SOC) categories represented in the analysis. The final weighted sample contains 97,084 civilian wage and salary workers aged 21–64, each with positive CPS weights and usual weekly earnings, nested in 488 nonfarm occupations.

The March CPS data provide the dependent variable and all individual-level covariates of earnings. The dependent variable is the natural log of usual weekly earnings in 1999 dollars, including wages, tips, and bonuses. Weekly earnings that the BLS top-codes to protect confidentiality are replaced with the gender- and year-specific means of a Pareto distribution (see Hirsch and McPherson 2001, p. 6).

The individual-level covariates mimic those found in standard earnings functions. Education, an ordinal variable in the CPS, is converted into a continuous measure of years of education using gender- and race-specific means of each category calculated from the 1993–1998 General Social Surveys (Davis and Smith 2000). Experience is approximated by age minus years of education minus six, with negative results set to zero. Labor market commitment is approximated with a dummy variable coded “0” if the respondent worked part-time, part-year, or both, and “1” otherwise. Firm-specific capital is approximated by a binary variable coded “0” if a respondent had more than one employer in the past year, and “1” otherwise. Other individual-level correlates of earnings, indexed in the model with dummy variables, are gender, race, marital status, parenting status, Southern region, urban or rural location, industry, and establishment size. Union membership is also included to capture any direct effect of collective bargaining on wages; unfortunately, no suitable earnings data offer individual-level measures of licensing, voluntary certification, or association membership (see Weeden 2002 for further discussion).

Occupation-Level Data

The occupation-level variables, their means, and their bivariate correlations are presented in Table 1. The first set of variables measure occupational skills and working conditions and can be interpreted through either a functionalist or a human capital lens (as indicators of otherwise unobserved occupation-specific capital). Cognitive skills, nurturing, physical demands, and hazardous conditions are measured with unweighted means of the z scores of the component items from the *Dictionary of Occupational Titles* ([DOT] U.S. Department of Labor et al. 1991; England and Kilbourne 1988). Authority is measured with a binary variable indicating whether an occupation requires supervisory or managerial skills (England 1992, p. 137). A sixth occupation-level control, the proportion of women in an occupation (U.S. Department of Commerce 1993), is included because of its demonstrated association with occupational earnings (e.g., Kilbourne et al. 1994; cf. Tam 1997).

The final set of variables in Table 1 index the extent to which an occupation is characterized by each of the five closure strategies. Measures of educational credentialing and unionization are constructed from the 1994–1996 Outgoing Rotation Group surveys of the CPS (BLS, 1997), while measures of licensing, voluntary certification, and association representation were collected from published directories (Columbia Books 1995, 1997; Gale 1996a, 1996b, 1996c, 1997, 1999) and supplementary information obtained from state licensing departments, association web sites, and contacts with certification program coordinators. The latter were constructed with a “bottom-up” approach that involved (1) taking a census of all occupational licenses, voluntary certifications, and associations existing between 1995 and 1997 (see Weeden 2002); (2) identifying the occupation covered by the license, certification, or association (e.g., certified public accountant, registered nurse, paralegal); (3) mapping these groups onto the 1990 SOC

Table 1: Means, Standard Deviations, and Bivariate Correlations of Occupation-Level Skill, Gender Composition, and Closure Variables

| Variable | Mean | Std. Dev. | Bivariate Correlations | | | | | | | | | | | |
|--|--------|-----------|------------------------|-------|-------|-------|-------|-------|------|-------|-------|-------|------|--|
| | | | 1. | 2. | 3. | 4. | 5. | 6. | 7. | 8. | 9. | 10. | 11. | |
| Skills and working conditions ^a | | | | | | | | | | | | | | |
| 1. Cognitive skills | 0.014 | 0.848 | 1.00 | | | | | | | | | | | |
| 2. Authority | 0.160 | 0.366 | 0.36 | 1.00 | | | | | | | | | | |
| 3. Nurturing | 0.312 | 0.880 | 0.35 | 0.14 | 1.00 | | | | | | | | | |
| 4. Physical demands | -0.163 | 0.803 | -0.03 | -0.32 | -0.30 | 1.00 | | | | | | | | |
| 5. Hazardous conditions | -0.175 | 0.737 | -0.43 | -0.14 | -0.46 | 0.25 | 1.00 | | | | | | | |
| 6. Percent female ^b | 0.466 | 0.302 | -0.03 | -0.16 | 0.46 | -0.08 | -0.49 | 1.00 | | | | | | |
| Closure strategies (proportion of occupation members covered) ^c | | | | | | | | | | | | | | |
| 7. Licensing | 0.330 | 0.404 | 0.39 | 0.24 | 0.43 | -0.19 | -0.17 | 0.06 | 1.00 | | | | | |
| 8. Educational credentialing | 0.574 | 0.257 | 0.83 | 0.18 | 0.57 | -0.20 | -0.56 | 0.23 | 0.48 | 1.00 | | | | |
| 9. Voluntary certification | 0.078 | 0.220 | 0.23 | -0.09 | 0.12 | 0.10 | -0.10 | 0.07 | 0.26 | 0.29 | 1.00 | | | |
| 10. Association representation | 0.153 | 0.344 | 0.48 | 0.07 | 0.19 | 0.10 | -0.18 | -0.01 | 0.26 | 0.49 | 0.35 | 1.00 | | |
| 11. Unionization | 0.161 | 0.154 | -0.22 | -0.25 | 0.10 | 0.08 | 0.33 | -0.18 | 0.01 | -0.11 | -0.12 | -0.06 | 1.00 | |

Note: $N=97,084$ individuals in 488 occupations. Values are weighted by occupation size. All correlations except the value in italics are significant at $p < .01$.

^a Scale items are as follows: (a) cognitive skills—general educational development, data complexity, numerical aptitude, intelligence, training time; (b) nurturing—face-to-face service (England 1992), talking/hearing, dealing with people; (c) physical demands—motor coordination, finger dexterity, form perception, spatial perception, visual requirements; (d) hazardous conditions—exposure to hazards, exposure to fumes, need to stoop, need to climb, environmental disamenities. Authority is a binary variable coded “1” if the occupation requires supervisory or managerial skills, “0” otherwise (see England 1992). Unless otherwise noted, items are from the DOT (see England and Kilbourne 1988).

^b Calculated from the 1990 EEO Census file (U.S. Dept. of Commerce 1993).

^c See text for variable definitions and data sources.

Table 2: Estimated Occupation-Level Effects, from Regression Analysis of (Logged) Weekly Earnings, of Individual Attributes, Occupational Skills, and Closure Strategies

| Independent Variable | Model 1 | |
|-------------------------------|---------|------------|
| | Coeff. | Std. Err. |
| FIXED EFFECTS ^a | | |
| Constant | 5.80 | (.030) *** |
| Skills and working conditions | | |
| Cognitive skills | 0.072 | (.015) *** |
| Authority | 0.127 | (.024) *** |
| Nurturing | -0.047 | (.011) *** |
| Physical demands | 0.047 | (.010) *** |
| Hazardous conditions | 0.008 | (.012) |
| Proportion female | -0.132 | (.030) *** |
| Closure strategies | | |
| Licensing | 0.085 | (.023) *** |
| Educational credentialing | 0.317 | (.054) *** |
| Voluntary certification | 0.116 | (.028) *** |
| Association representation | -0.033 | (.014) * |
| Unionization | 0.047 | (.044) |
| RANDOM EFFECTS | | |
| Occupation-level variance | 0.014 | (.001) |
| Individual-level variance | 0.220 | (.001) |
| -2 log-likelihood | 129,517 | |

Note: $N=97,084$ individuals in 488 occupations. See Table 1 for data sources.

^a The model fits all individual-level covariates (see text) and their interactions with sex as grand-mean centered fixed effects.

* $p < .05$

** $p < .01$

*** $p < .001$ (two-tailed tests)

using information in the DOT, the *Alphabetical Index of Occupations* (U.S. Department of Commerce 1992), and association literature; and (4) constructing continuous measures of the extent of closure in an occupation. The extent of licensure is estimated by the proportion of individuals who work in a state that licenses the occupation;¹ voluntary certification by the proportion of occupation members who hold a certification sponsored by an occupational association or independent certification board; association representation by the proportion of incumbents who belong to an occupational association (membership is imputed for 4 percent of associations); educational credentialing by the proportion of occupation members who have attended some college (see Hauser and Warren 1997); and union representation by the

proportion of incumbents who are union members.

Closure and Earnings

Table 2 presents the occupation-level coefficients from the model specified in the Equation. This model also fits the aforementioned individual-level covariates of earnings and their interaction with sex, but because these coefficients follow predictable patterns, they are not presented here. The estimated effects of the occupational skill and gender composition variables are likewise consistent with investment models of earnings and gender devaluation arguments, respectively.

The coefficients of the closure strategies generated under this full model offer substantial support for closure theory. Three of the closure

strategies—licensing, voluntary certification, and educational credentialing—have positive, significant effects on mean earnings. Moreover, the wage premiums associated with these strategies are hardly trivial. At mean levels of educational credentialing (see Table 1), this wage premium is 20 percent ($\exp[0.317 \cdot 0.574] = 1.200$), and if all incumbents have attended college, the estimated wage premium is 37 percent ($\exp[.317 \cdot 1] = 1.373$), relative to an occupation in which none of the incumbents have attended college. An occupation that is licensed in every state receives an estimated wage premium of nearly 9 percent ($\exp[0.085 \cdot 1] = 1.089$), and an occupation that is fully certified enjoys a wage premium of 12 percent ($\exp[0.116 \cdot 1] = 1.123$). Subsequent analyses (not shown) reveal that the payoff to these closure strategies is greater in the professions, but it is not limited to these occupations nor is it tightly linked to occupations' task complexity (Weeden 2002).

Support for closure theory in these results is not, however, entirely unqualified. Neither unionization nor association representation have significant, direct correlations with occupational rewards. It is possible that unionization generates real wage returns to both union and nonunion workers, but the relevant boundaries are formed around industries or local labor markets instead of occupations. Employers who face a labor shortage or fear union campaigns may look to other industry players, but deem occupational references irrelevant (Bridges 1995). Consistent with this interpretation, the coefficient of occupation-level unionization is positive and significant when the industry dummies are excluded from the model (not shown). Of course, these dummies may be picking up other industry attributes related to earnings, so the test is hardly conclusive.

The apparent failure of association representation to generate earnings returns is best understood by revisiting the mechanisms on which associations rely—enhancing diffuse demand, channeling demand, and sending a

quality-of-service signal (see Figure 1). It is possible that these mechanisms indeed generate earnings returns, but capturing consumer attention depends more on name recognition, legislative power, organizational capital, or the financial assets necessary to mount an effective marketing campaign than on membership density. Nevertheless, if we assume that these resources are correlated with membership density, we would still expect a positive coefficient of association representation. An alternative interpretation is that one or more of the underlying mechanisms does not generate earnings returns. The obvious suspect is the only mechanism unique to occupational associations: generating diffuse demand. This mechanism will only raise mean earnings if it is combined with effective ways to channel demand to the occupation and also to restrict an inflow of new recruits to that occupation. Apart from their licensing and certification activities, occupational associations have virtually no control over the labor supply and hence must rely on largely symbolic claims to protect task jurisdictions. Association rhetoric resounds with such symbolic claims, but these results suggest no one is listening.

Conclusion

This analysis reveals the promise of closure theory for understanding the contemporary sources of earnings inequality. The results support the basic claim that closure practices affect collective earnings. However, they also confirm the need to (1) distinguish between strategies of closure rather than treating them as substitutable means to the same ends, and (2) recognize that the particular combination of mechanisms underlying the closure strategies influence the rewards they generate. Specifically, three institutionalized strategies of closure—licensing, formal educational credentials, and voluntary certification—have statistically and substantively important effects on the hierarchy of material rewards.

The relatively strong impact of licensing and credentialing suggest that tangible restrictions on the supply of labor are particularly effective, at least as measured by the criterion of occupational earnings. Conversely, the relatively weak effect of voluntary certification and the lack of a positive effect of association representation suggest that largely symbolic jurisdictional claims and efforts to enhance diffuse demand have a more modest impact on occupational rewards. The emphasis of stratification scholars on supply-side monopolization appears partially justified.

Although the results of the analyses are consistent with the predictions of closure theory, they do not recommend theoretical imperialism. The measures of occupational skills have significant net associations with occupational earnings, even after explicit measures of formal closure are incorporated into the analyses. One could, of course, claim that skills are socially constructed as part of the closure project, and that the DOT's job analysts are no more immune to occupational rhetoric, meaning-making, or stereotypes than the lay public. However, if the occupations most able to influence DOT skill ratings are also those that make extensive use of formal closure strategies, the existence of a net skill effect suggests that skill is more than an indirect measure of market power. The data support only a more cautious interpretation, namely that both technical complexity and closure affect the occupational reward structure.

Moreover, closure theory does not account for the substantial portion of the variation in earnings—in these data, approximately 66 percent—that occurs within occupations. Some of this within-occupation heterogeneity may reflect the aggregate nature of the SOC. If the occupational categories available in earnings data mapped perfectly onto the actual groups around which closure takes place, some of this heterogeneity might rightfully be attributed to occupational closure. Neverthe-

less, it is unreasonable to claim that group-level processes are the only ones at work. In the United States, closure practices shape the supply of and demand for certain occupations, thereby affecting the reward structure, but do not supplant market-based processes that differentially reward individuals based on their human capital or other personal attributes.

These caveats notwithstanding, the analyses confirm the longstanding sociological assumption that individualistic models are insufficient for understanding earnings inequality. Although much work remains to be done to make the case airtight, closure theory provides a viable complement to dominant investment accounts of inequality. In so doing, it refocuses sociological attention on the collective mobility projects (Sorokin 1959) that link positions in the division of labor and their reward packages.

NOTES

1. The exception is engineers, for whom licensing laws are notoriously riddled with exemptions. I estimate the proportion of licensed civil engineers at 0.44, chemical engineers at 0.08, aerospace and marine engineering/naval architecture at zero, and the remaining specialties at 0.15 for an average proportion of 0.18 (CBRPELS 1998).

REFERENCES

- Abbott, Andrew. 1988. *The System of Professions: An Essay on the Division of Expert Labor*. Chicago: University of Chicago Press.
- Bridges, William P. 1995. "Where Do Markets Go To? An Analysis of Change in Internal and External Labor Markets." *Research in Social Stratification and Mobility* 14: 71-98.
- Brint, Steven. 1994. *In an Age of Experts: The Changing Role of Professionals in Politics and Public Life*. Princeton, NJ: Princeton University Press.
- Caplow, Theodore. 1954. *The Sociology of Work*. Minneapolis: University of Minnesota Press.
- CBRPELS (California Board of Registration for Professional Engineers and Land Surveyors). 1998. *Sunset Review Report, 1998* [online]. Accessed November 5, 2001. Available at <http://www.dca.ca.gov/pels/sunset/98sunset.pdf>.

- Cohen, Randi-Charlene. 1996. *Who's Planning for Your Future? Jurisdictional Competition among Organizations and Occupations in the Personal Financial Planning Industry*. Ph.D. dissertation, Stanford University.
- Collins, Randall. 1979. *The Credential Society: An Historical Sociology of Education and Stratification*. New York: Academic Press.
- Columbia Books. 1995. *National Trade and Professional Associations of the United States, 30th Annual Edition*. Washington, DC: Columbia Books.
- _____. 1997. *National Trade and Professional Associations of the United States, 32nd Annual Edition*. Washington, DC: Columbia Books.
- Cullen, John B. 1978. *The Structure of Professionalism: A Quantitative Examination*. New York: Petrocelli Books.
- Davis, James Allen, and Tom W. Smith. 2000. *General Social Survey Cumulative Data File, 1972–1998*. Chicago: NORC.
- England, Paula. 1992. *Comparable Worth: Theories and Evidence*. New York: Aldine de Gruyter.
- England, Paula, and Barbara Kilbourne. 1988. *Occupational Measures from the Dictionary of Occupational Titles for 1980 Census Detailed Occupations*. Richardson, TX: Paula England and Barbara Kilbourne. Distributed by Inter-University Consortium for Political and Social Research, Ann Arbor, MI.
- Freeman, Robert B., and James L. Medoff. 1984. *What Do Unions Do?* New York: Basic Books.
- Freidson, Eliot. 1986. *Professional Powers: A Study of the Institutionalization of Formal Knowledge*. Chicago: University of Chicago Press.
- Friedman, Milton, and Simon Kuznets. 1945. *Income from Independent Professional Practice*. New York: National Bureau of Economic Research.
- Gale. 1996a. *Professional and Occupational Licensing Directory*. Detroit: Gale Research Company.
- _____. 1996b. *Certification and Accreditation Programs Directory*. Detroit: Gale Research Company.
- _____. 1996c. *Encyclopedia of Associations, 31st Edition*. Vol. 1 (Part 3). Detroit: Gale Research Company.
- _____. 1997. *Encyclopedia of Associations, 32nd Edition*. Vol. 1 (Part 3). Detroit: Gale Research Company.
- _____. 1999. *Certification and Accreditation Programs Directory, 2nd Edition*. Detroit: Gale Research Company.
- Grusky, David B., and Jesper B. Sørensen. 1998. "Can Class Analysis Be Salvaged?" *American Journal of Sociology* 103(5): 1187–234.
- Hauser, Robert, and John Robert Warren. 1997. "Socioeconomic Indices for Occupations: A Review, Update, and Critique." *Sociological Methodology* 27: 177–298.
- Hirsch, Barry T., and David A. McPherson. 2001. *Union Membership and Earnings Data Book: Compilations from the Current Population Survey, 2001 edition*. Washington, DC: Bureau of National Affairs.
- Kilbourne, Barbara Stanek, George Farkas, Kurt Beron, Dorothea Weir, and Paula England. 1994. "Returns to Skill, Compensating Differentials, and Gender Bias: Effects of Occupational Characteristics on the Wages of White Women and Men." *American Journal of Sociology* 100(3): 689–719.
- Kleiner, Morris M. 2006. *Occupational Licensing: Ensuring Quality or Restricting Competition?* Kalamazoo, MI: W.E. Upjohn Institute.
- MacDonald, Keith M. 1995. *The Sociology of the Professions*. London: Sage.
- Manza, Jeff. 1992. "Classes, Status Groups, and Social Closure: A Critique of Neo-Weberian Closure Theory." Pp. 275–302 in *Current Perspectives in Social Theory*, edited by Ben Agger. Greenwich, CT: JAI Press.
- Murphy, Raymond. 1988. *Social Closure: The Theory of Monopolization and Exclusion*. Oxford: Clarendon Press.
- Parkin, Frank. 1971. *Class Inequality and Political Order: Social Stratification in Capitalist and Communist Societies*. New York: Praeger.
- _____, ed. 1974. *The Social Analysis of Class Structure*. New York: Harper and Row.
- _____. 1979. *Marxism and Class Theory: A Bourgeois Critique*. New York: Columbia University Press.
- Sørensen, Aage B. 2000. "Toward a Sounder Basis for Class Analysis." *American Journal of Sociology* 105(6): 1559–71.
- Sorokin, Pitirim. 1995 [1959]. "Social and Cultural Mobility." Pp. 245–50 in *Social Stratification: Class, Race, and Gender in Sociological Perspective*, 1st ed., edited by David B. Grusky. Boulder: Westview Press.
- Stigler, George J. 1971. "The Theory of Economic Regulation." *Bell Journal of Economics and Management Science* 2(1): 3–17.
- Tam, Tony. 1997. "Sex Segregation and Occupational Gender Inequality in the United States: Devaluation or Specialized Training?" *American Journal of Sociology* 102(6): 1652–92.
- Tilly, Charles. 1998. *Durable Inequality*. Berkeley: University of California Press.

- U.S. Bureau of Labor Statistics. Various years. *Current Population Surveys, March Supplement*. Washington, DC: U.S. Government Printing Office.
- _____. Various years. *Current Population Surveys, Outgoing Rotation Groups*. Washington, DC: U.S. Government Printing Office.
- U.S. Department of Commerce. 1992. *Alphabetical Index of Industries and Occupations*. Washington, DC: U.S. Government Printing Office.
- _____. 1993. *Census of Population and Housing, 1990: Equal Employment Opportunity (EEO) Supplemental Tabulations File*. Washington, DC: The Bureau of the Census.
- U.S. Department of Labor, U.S. Employment Service, and the North Carolina Occupational Analysis Field Center. 1991. *Dictionary of Occupational Titles: Revised Fourth Edition*. Washington, DC: U.S. Department of Labor, U.S. Employment Service, and Raleigh, NC: North Carolina Occupational Analysis Field Center. Distributed by Inter-University Consortium for Political and Social Research, Ann Arbor, MI.
- Weber, Max. 1978 [1922]. *Economy and Society: An Outline of Interpretive Sociology*, edited by Guenther Roth and Claus Wittich. Berkeley: University of California Press.
- Weeden, Kim A. 2002. "Why Do Some Occupations Pay More than Others? Social Closure and Earnings Inequality in the United States." *American Journal of Sociology* 108(1): 55–101.
- Weeden, Kim A., and David B. Grusky. 2007. "A New Guilded Age? Class-Biased Institutional Change and Rising Wage Inequality in the United States." Working paper, Cornell University.

Gradational Status Groupings

■ CLASSIC STATEMENTS ■

19 ■ Peter M. Blau and Otis Dudley Duncan,
with the collaboration of Andrea Tyree

Measuring the Status of Occupations

Two approaches have dominated the investigations of occupational hierarchy carried out by students of social stratification. One is the effort to develop a socioeconomic classification scheme for occupations. Perhaps the most influential work here was that of the census statistician Alba M. Edwards.¹ His “social-economic grouping” of occupations has been widely used in studies of occupational stratification and mobility. With certain modifications it led to the “major occupation groups” used by the Bureau of the Census since 1940. . . . To suggest that his grouping supplied a “scale,” Edwards contented himself with showing differences in average or typical levels of education and income of the workers included in the several categories: “Education is a very large factor in the social status of workers, and wage or salary income is a very large factor in their economic status.”²

A more recent development is the derivation of scores for *detailed* census occupation titles representing a composite index of education and income levels of workers in each such occupation. Priority for this specific technique probably belongs to social scientists in Canada,³ with a similar approach

being taken in this country by both a private researcher worker⁴ and, lately, in official publications of the U. S. Bureau of the Census.⁵

The second approach to occupational stratification is to secure, from samples more or less representative of the general public, ratings of the “general standing” or “prestige” of selected occupations. Such ratings have been shown to be remarkably close to invariant with respect to (a) the composition and size of the sample of raters; (b) the specific instructions or form of the rating scale; (c) the interpretation given by respondents to the notion of “general standing”; and (d) the passage of time.⁶ The high order of reliability and stability evidenced by prestige ratings would commend their use in problems requiring social distance scaling of the occupations pursued by a general sample of the working force, but for one fact: ratings have hitherto been available only for relatively small numbers of occupation titles. Many research workers have resorted to ingenious schemes for splicing *ad hoc* judgments into the series of rated occupations, but no general solution to the problem has been widely accepted.

Peter M. Blau and Otis Dudley Duncan, with the collaboration of Andrea Tyree. “Measuring the Status of Occupations” in *The American Occupational Structure*, pp. 118–124. Copyright © 1967 by Peter M. Blau and Otis Dudley Duncan. Reprinted with the permission of The Free Press, A Division of the Simon & Schuster Adult Publishing Group. All rights reserved.

Work currently in progress at the National Opinion Research Center promises to overcome this difficulty by supplying prestige ratings for a comprehensive list of occupations. In the absence of such ratings at the time of the Occupational Changes in a Generation (OCG) survey we fell back on the idea of a socioeconomic index of occupational status. The particular index we used, however, was one designed to give near-optimal reproduction of a set of prestige ratings. A full account of the construction of this index is given elsewhere,⁷ and only a few general points need to be made before presenting some illustrations of the scale values assigned to occupations.

In the derivation of the socioeconomic index of occupational status, prestige ratings obtained from a sizable sample of the U.S. population in 1947 were taken as the criterion. These were available for 45 occupations whose titles closely matched those in the census detailed list. Data in the 1950 Census of Population were converted to two summary measures: per cent of male workers with four years of high school or a higher level of educational attainment, and per cent with incomes of \$3,500 or more in 1949 (both variables being age-standardized). The multiple regression of per cent “excellent” or “good” prestige ratings on the education and income measures was calculated. The multiple correlation, with the 45 occupations as units of observation, came out as .91, implying that five-sixths of the variation in aggregate prestige ratings was taken into account by the combination of the two socioeconomic variables. Using the regression weights obtained in this calculation, all census occupations were assigned scores on the basis of their education and income distributions. Such scores may be interpreted either as estimates of (unknown) prestige ratings or simply as values on a scale of occupational socioeconomic status (“occupational status” for short). The scale is represented by two-digit numbers ranging from 0 to 96. It closely resembles the scales of Blishen, Bogue, and the

U.S. Bureau of the Census mentioned earlier, although there are various differences in detail among the four sets of scores.

One of the most serious issues in using any index of occupational status in the study of mobility has to do with the problem of temporal stability. . . . Fortunately, we now have a detailed study of temporal stability in occupational prestige ratings. The results are astonishing to most sociologists who have given the matter only casual thought. A set of ratings obtained as long ago as 1925 is correlated to the extent of .93 with the latest set available, obtained in 1963. The analysts conclude, “There have been no substantial changes in occupational prestige in the United States since 1925.”⁸ Less complete evidence is available for the socioeconomic components of our index, but information available in the Censuses of 1940, 1950, and 1960 points to a comparably high order of temporal stability,⁹ despite major changes in the value of the dollar and the generally rising levels of educational attainment. . . .

Two-digit status scores are available for 446 detailed occupation titles. Of these, 270 are specific occupation categories; the remainder are subgroupings, based on industry or class of worker, of 13 general occupation categories. The reader may consult the source publication for the scores of particular occupations of interest.¹⁰ Here we shall only illustrate the variation of the scores by citing illustrative occupations, not always those of the greatest numerical importance (see Table 1). . . .

Table 1 makes it clear that occupations of very different character may have similar status scores. In particular, there is considerable overlap of scores of occupations in distinct major occupation groups. Indeed, only five points separate the lowest occupation in the “professional, technical, and kindred workers” group from the highest among “laborers, except farm and mine.” Nevertheless, the major occupation group classification accounts for three-fourths of the variation in scores among detailed occupations. The sta-

Table 1. Occupations Illustrating Various Scores on the Index of Occupational Status*

| <i>Score Interval</i> | <i>Title of Occupation (Frequency per 10,000 Males in 1960 Experienced Civilian Labor Force in Parentheses)</i> |
|-----------------------|---|
| 90 to 96 | Architects (7); dentists (18); chemical engineers (9); lawyers and judges (45); physicians and surgeons (47) |
| 85 to 89 | Aeronautical engineers (11); industrial engineers (21); salaried managers, banking and finance (30); self-employed proprietors, banking and finance (5) |
| 80 to 84 | College presidents, professors and instructors (31); editors and reporters (14); electrical engineers (40); pharmacists (19); officials, federal public administration and postal service (13); salaried managers, business services (11) |
| 75 to 79 | Accountants and auditors (87); chemists (17); veterinarians (3); salaried managers, manufacturing (133); self-employed proprietors, insurance and real estate (9) |
| 70 to 74 | Designers (12); teachers (105); store buyers and department heads (40); credit men (8); salaried managers, wholesale trade (41); self-employed proprietors, motor vehicles and accessories retailing (12); stock and bond salesmen (6) |
| 65 to 69 | Artists and art teachers (15); draftsmen (45); salaried managers, motor vehicles and accessories retailing (18); self-employed proprietors, apparel and accessories retail stores (8); agents, n.e.c. (29); advertising agents and salesmen (7); salesmen, manufacturing (93); foremen, transportation equipment manufacturing (18) |
| 60 to 64 | Librarians (3); sports instructors and officials (12); postmasters (5); salaried managers, construction (31); self-employed proprietors, manufacturing (35); stenographers, typists, and secretaries (18); ticket, station, and express agents (12); real estate agents and brokers (33); salesmen, wholesale trade (106); foremen, machinery manufacturing (28); photoengravers and lithographers (5) |
| 55 to 59 | Funeral directors and embalmers (8); railroad conductors (10); self-employed proprietors, wholesale trade (28); electrotypers and stereotypers (2); foremen, communications, utilities, and sanitary services (12); locomotive engineers (13) |
| 50 to 54 | Clergymen (43); musicians and music teachers (19); officials and administrators, local public administration (15); salaried managers, food and dairy products stores (21); self-employed proprietors, construction (50); bookkeepers (33); mail carriers (43); foremen, metal industries (28); toolmakers, and die-makers and setters (41) |
| 45 to 49 | Surveyors (10); salaried managers, automobile repair services and garages (4); office machine operators (18); linemen and servicemen, telephone, telegraph and power (60); locomotive firemen (9); airplane mechanics and repairmen (26); stationary engineers (60) |
| 40 to 44 | Self-employed proprietors, transportation (8); self-employed proprietors, personal services (19); cashiers (23); clerical and kindred workers, n.e.c. (269); electricians (77); construction foremen (22); motion picture projectionists (4); photographic process workers (5); railroad switchmen (13); policemen and detectives, government (51) |
| 35 to 39 | Salaried and self-employed managers and proprietors, eating and drinking places (43); salesmen and sales clerks, retail trade (274); book-binders (3); radio and television repairmen (23); firemen, fire protection (30); policemen and detectives, private (3) |
| 30 to 34 | Building managers and superintendents (7); self-employed proprietors, gasoline service stations (32); boilermakers (6); machinists (111); millwrights (15); plumbers and pipe fitters (72); structural metal workers (14); tinsmiths, coppersmiths, and sheet metal workers (31); deliverymen and routemen (93); operatives, printing, publishing and allied industries (13); sheriffs and bailiffs (5) |
| 25 to 29 | Messengers and office boys (11); newsboys (41); brickmasons, stonemasons, and tile setters (45); mechanics and repairmen, n.e.c. (266); plasterers (12); operatives, drugs and medicine manufacturing (2); ushers, recreation and amusement (2); laborers, petroleum refining (3) |
| 20 to 24 | Telegraph messengers (1); shipping and receiving clerks (59); bakers (21); cabinetmakers (15) excavating, grading, and road machine operators (49); railroad and car shop mechanics and repairmen (9); tailors (7); upholsterers (12); bus drivers (36); filers, grinders, and polishers, metal (33); welders and flame-cutters (81) |
| 15 to 19 | Blacksmiths (5); carpenters (202); automobile mechanics and repairmen (153); painters (118); attendants, auto service and parking (81); laundry and dry cleaning operatives (25); truck and tractor drivers (362); stationary firemen (20); operatives, metal industries (103); operatives, wholesale and retail trade (35); barbers (38); bartenders (36); cooks, except private household (47) |
| 10 to 14 | Farmers (owners and tenants) (521); shoemakers and repairers, except factory (8); dyers (4); taxicab drivers and chauffeurs (36); attendants, hospital and other institution (24); elevator operators (11); fishermen and oystermen (9); gardeners, except farm, and groundskeepers (46); longshoremen and stevedores (13); laborers, machinery manufacturing (10) |
| 5 to 9 | Hucksters and peddlers (5); sawyers (20); weavers, textile (8); operatives, footwear, except rubber, manufacturing (16); janitors and sextons (118); farm laborers, wage workers (241); laborers, blast furnaces, steel works, and rolling mills (26); construction laborers (163) |
| 0 to 4 | Coal mine operatives and laborers (31); operatives, yarn, thread and fabric mills (30); porters (33); laborers, saw mills, planning mills, and millwork (21) |

*n.e.c. means "not elsewhere classified"

Sources: Reiss, *op. cit.*, Table B-1; and U.S. Bureau of the Census, *1960 Census of Population*, Final Report, PC(1)-1D, Table 201.

tus scores offer a useful refinement of the coarser classification but not a radically different pattern of grading.

Table 1 probably does not illustrate adequately the variation by industry subclass of such occupation categories as “operatives, not elsewhere classified” and “laborers, not elsewhere classified.” Such variation is fairly substantial. It must be understood, however, that particularly at these levels of the census classification scheme the occupation-industry categories represent groups of jobs with quite heterogeneous specifications, although the groups are thought to be somewhat homogeneous as to the degree of skill and experience required for their performance. No one has yet faced the question of what a study of occupational mobility would look like if all the 20,000 or more detailed titles in the *Dictionary of Occupational Titles* were coded without prior grouping.

The use of occupational status scores carries a theoretical implication. We are assuming, in effect, that the occupation structure is more or less continuously graded in regard to status rather than being a set of discrete status classes. The justification of such an assumption is not difficult. One needs only to look at any tabulation of social and economic characteristics of persons engaged in each specific occupation (whatever the level of refinement in the system of occupational nomenclature). We discover that the occupations overlap—to a greater or lesser degree, to be sure—in their distributions of income, educational attainment, consumer expenditures, measured intelligence, political orientations, and residential locations (to mention but a few items). One may sometimes find evidence supporting the interpretation that there are “natural breaks” in such distributions. Interpretations of this kind were advanced [elsewhere]¹¹ in respect to the dividing line between farm and nonfarm and between white-collar and manual occupations. The evidence did not permit the conclusion that such occupation categories are

entirely disjunct. The analysis . . . suggests that boundaries may be discerned between the three broad groups, [but] also shows that these are by no means sharp lines without any overlap.

If we choose to think of occupational status as exhibiting continuous variation, the appropriate analytical model is one that treats status as a quantitative variable. This point of view has far-reaching implications for the conceptualization of the process of mobility as well as for the analysis and manipulation of data purporting to describe the process.

NOTES

1. Alba M. Edwards, *Comparative Occupation Statistics for the United States, 1870 to 1940*, Washington: Government Printing Office, 1943.

2. *Ibid.*, p. 180.

3. Enid Charles, *The Changing Size of the Family in Canada*, Census Monograph No. One, Eighth Census of Canada, 1941, Ottawa: The Kings Printer and Controller of Stationery, 1948; Bernard R. Blishen, “The Construction and Use of an Occupational Class Scale,” *Canadian Journal of Economics and Political Science*, 24 (1958), 519–531.

4. Donald J. Bogue, *Skid Row in American Cities*, Chicago: Community and Family Study Center, University of Chicago, 1963, Chapter 14 and Appendix B.

5. U. S. Bureau of the Census, *Methodology and Scores of Socioeconomic Status*, Working Paper, No. 15 (1963); U. S. Bureau of the Census, “Socioeconomic Characteristics of the Population: 1960,” *Current Population Reports*, Series P-23, No. 12 (July 31, 1964).

6. Albert J. Reiss, Jr., et al., *Occupations and Social Status*, New York: Free Press of Glencoe, 1961; Robert W. Hodge, Paul M. Siegel, and Peter H. Rossi, “Occupational Prestige in the United States, 1925–63,” *American Journal of Sociology*, 70 (1964), 286–302.

7. Otis Dudley Duncan, “A Socioeconomic Index for All Occupations,” in Reiss, *op. cit.*, pp. 109–138.

8. Hodge, Siegel, and Rossi, *op. cit.*, p. 296.

9. Reiss, *op. cit.*, p. 152. (Work in progress by Hodge and Treiman further supports this point.)

10. Duncan, *op. cit.*, Table B-1, pp. 263–275.

11. Peter M. Blau and Otis Dudley Duncan, *The American Occupational Structure*, New York: The Free Press, 1967, Chapter 2.

20 ■ Donald J. Treiman

Occupational Prestige in Comparative Perspective

In the three decades since World War II there have been some eighty-five studies of occupational prestige conducted in more than sixty countries throughout the world, ranging from highly industrialized places such as the U.S. to traditional societies such as India, Thailand, Nigeria, and New Guinea. Although these studies vary somewhat in their specific details, they all utilize the same basic procedure: a sample of the population is asked to rate or rank a set of occupational titles with respect to their prestige or social standing. These ratings are then aggregated into mean scores (or other measures of central tendency) and the scores are treated as indicators of the relative prestige of the evaluated occupations.

A remarkable feature of these studies is that they yield the same results regardless of the exact wording of the questionnaire. It does not matter whether respondents are asked about the “prestige” or “social standing” or “respect” accorded certain occupations, or whether they are asked to rate occupations on a scale or to rank them in any other way. The results are the same. A second striking feature is that the educated and uneducated, the rich and poor, the urban and rural, the old and young, all on the average have the same perceptions of the prestige hierarchy. There is no systematic subgroup

variation in the relative ratings of jobs. This is of considerable importance since it allows us to make use of data drawn from rather poor samples of the population—for example, students, members of voluntary organizations, representatives of special subcultures—without fear that if we had a different sample we would get different answers. The third noteworthy feature is that although the distribution of the labor force in various occupations varies substantially from place to place, the same sorts of occupations tend to exist everywhere. Even if there are not many airplane pilots or professors in a given country, there tend to be at least a few of them and the population at large knows what these jobs are. In general, the organization of work into specific jobs is amazingly uniform across societies. Pretty much everywhere there are distinctions between weavers and tailors, and between carpenters, painters, and plumbers. And the uniformity in occupations across societies is reflected in the uniformity of occupational titles appearing in prestige studies. As a result, matching occupational titles across countries is a less onerous task than might be expected.

These three features, uniform results regardless of measurement procedures, minimal subgroup variations, and similarity of occupational titles, make possible a systematic

Donald J. Treiman. “A Standard Occupational Prestige Scale for Use with Historical Data,” from *The Journal for Interdisciplinary History* VII (1976), pp. 285–290. Copyright © 1976 by the editors of *The Journal for Interdisciplinary History* and the Massachusetts Institute of Technology. Reprinted with permission of the editors of *The Journal for Interdisciplinary History* and the MIT Press, Cambridge, Massachusetts.

comparison of occupational prestige hierarchies among countries. The basic procedure is to match titles across countries, e.g., “physician” in the U.S. with “doctor” in Australia, “medecin” in Mauritania, “medico” in Argentina, “laege” in Denmark, and then to compute a product moment correlation between the prestige scores for all matching titles for each pair of countries. The correlation coefficients thus generated can be taken as measures of the similarity of prestige evaluations between each pair of countries. The fundamental conclusion from such computations is that there is substantial uniformity in occupational evaluations throughout the world: the average intercorrelation between pairs of countries is .81.¹ As such numbers go, it is extremely high and fully justifies treatment of the prestige hierarchy in any given country as reflecting, in large part, a single worldwide occupational prestige hierarchy. On the basis of this result, and in view of the need for a standard occupational scaling procedure, it seemed desirable to attempt to construct a standard occupational prestige scale which could be applied to any country.

In order to match occupational titles across countries, it was necessary to devise a comprehensive occupational classification scheme. To do this, I took advantage of an already existing scheme: the *International Standard Classification of Occupations*, Revised Edition (ISCO).² This classification is a “nested” scheme which clusters occupations into nine major groups, eighty-three minor groups, 284 unit groups, and 1,506 specific occupations. It was developed by the International Labour Office as a guide for national census offices to encourage the comparability of occupational statistics. Many foreign census bureaus do, in fact, utilize the ISCO scheme and do publish occupational statistics according to its guidelines.

However, since the 1,506 ISCO occupational categories did not correspond very well to the occupational titles for which I had prestige ratings, I followed the ISCO

scheme (with minor variations) down to the unit group level and then, within this level, made distinctions among specific occupations when they appeared warranted by my prestige data. The resulting classification contains 509 distinct occupational titles.

To derive generic prestige scores for each of these occupations, I converted all the data to a standard metric and then simply averaged scores across all countries in which a given title appeared. Scores for higher levels of aggregation (unit groups, minor groups, and major groups) were derived by various averaging procedures.³

How good is the prestige scale created by this procedure? The answer is—very good. Evidence for contemporary societies is extremely convincing. The average correlation of the new Standard Scale with the reported prestige hierarchies of fifty-five countries is .91 and only seven of the correlations are less than .87. Thus the Standard Scale is, on the average, the best available predictor of the prestige of occupations in any contemporary society. . . .⁴

A Theory of the Determinants of Prestige

Analysis of the universally shared occupational prestige hierarchy suggests that high prestige is allocated to those occupations which require a high degree of skill or which entail authority over other individuals or control over capital. Moreover, the nature of occupational specialization is such that specific occupations are relatively invariant in these characteristics across time and space. As a result, the prestige of specific occupations is relatively invariant as well. In fact, it is so uniform that a single occupational prestige scale will capture the basic features of the occupational hierarchy of any society.

Specialization of functions into distinct occupational roles necessarily results in inequalities among occupations with respect to skill, authority, and control over capital.

Some occupations, by their very definitions, require specific skills. For example, literacy is required of clerks because one cannot be a clerk if one is illiterate. Similarly, some occupations require control over capital or authority over other individuals as inherent definitions of their functions. For instance, a managerial job is one which involves “planning, direction, control, and co-ordination”;⁵ otherwise, it is not a managerial job but something else. Examples of these inherent inequalities can be located throughout the prestige hierarchy.

Skill, authority, and economic control are singled out as the basic resources which differentiate occupations because these are the fundamental aspects of power—they provide the crucial means to the achievement of desired goals. But the more powerful an occupation, the more important it is that it be performed well, since the consequences of competent or incompetent performances are more telling for such occupations. For example, if a garbage collector does his job poorly, little is lost; but if a surgeon is incompetent, a life can be lost. Or, similarly, if a chain store manager makes a poor business decision it may cost a firm a few hundred or at the most a few thousand dollars; but a poor decision on the part of a major executive can run into millions. Consequently, the more powerful an occupation, the greater the incentive to attract competent personnel to it. And since the basic mechanism for inducing people to perform tasks is to reward them, it follows that the most powerful positions will also be the most highly rewarded.

Other factors do enter into the determination of rewards, so that the relationship between power and privilege is not perfect. Some functions are in greater demand than others, depending upon the needs of society at any particular time. For example, in a hunting and gathering society, hunting is in greater demand than farming and thus hunting is more highly rewarded. And in a commercial economy, law is of great importance

and therefore highly rewarded. But these differences are relatively minor compared to the differences in occupational requisites and prerequisites which are inherent in the definitions of jobs and therefore stable across time and space.

Not only is the relative power and privilege of occupations essentially similar across societies, but so is the prestige accorded them, for prestige is granted in recognition of power and privilege. Prestige is the metric of “moral worth,” and the moral worth of positions reflects their control over socially valued resources and rewards, that is, their power and privilege.⁶ Since occupations are differentiated with respect to power, they will in turn be differentiated with respect to privilege and prestige. Thus, if this theory of prestige determinants is correct, these attributes of occupations will be highly correlated across societies.⁷ In particular, skill level, authority, economic power, wealth, income, and prestige will be highly intercorrelated with one another and will be highly correlated across countries.

In my work on occupational prestige in contemporary societies, I amassed data on the education levels of occupations for fifteen nations (as a surrogate measure for “skill”) and on income levels of occupations for eleven countries; no comparable measures of authority or control of capital were available. These data indicate a striking uniformity in occupational hierarchies. Like occupational prestige evaluations, occupational variations in education and in income proved to be highly similar from society to society. When measures of the average level of schooling of incumbents of each occupation were computed for the U.S., Argentina, Canada, West Germany, Ghana, Great Britain, India, Israel, Japan, the Netherlands, Norway, Taiwan, the U.S.S.R., Yugoslavia, and Zambia and these measures were intercorrelated, the average intercorrelation was .76, which is almost as high as the average prestige intercorrelation reported above (.81). And when

measures of the average income of incumbents of each occupation were computed for the U.S., Canada, Ceylon, Costa Rica, India, New Zealand, Pakistan, Surinam, Sweden, Taiwan, and Yugoslavia and these measures were intercorrelated, the average intercorrelation was .65, which is still a substantial correlation. Moreover, education, income, and prestige levels of occupations were highly correlated within each country: the average correlations were, respectively, .77 between education and income, .72 between education and prestige, and .69 between income and prestige. The average correlations with the Standard Scale were .79 for education and .70 for income.⁸

In short, the available data indicate that in the contemporary world occupational hierarchies are substantially invariant from place to place, even among countries varying widely in level of industrialization. This finding lends considerable empirical support to the theoretical argument outlined above.

NOTES

This is a revised version of a paper presented at the Conference on International Comparisons of Social Mobility in Past Societies held in 1972. Preparation of the paper was supported by a grant from the National Science Foundation to Columbia University (NSF #28050). I am grateful to the following for making unpublished material available or for giving leads to the work of others: Peter Decker, Sigmund Diamond, Clyde Griffen, James Henretta, David Herlihy, Theodore Hershberg, Richard Hopkins, Michael Katz, William Sewell, Jr., James Smith, and Stephan Thernstrom. Thanks are also extended to Vincent Covello, Theodore Riccardi, Jr., Rose M. Cascio, Michael Freeman, John Hammond, Jr., Herbert Klein, and Jane Ferrar. The comments of the participants at the MSSB Conference, and especially those of Griffen, were extremely helpful in preparing the revision.

1. This average correlation was computed over all pairs of countries with at least 10 occupational titles rated in common.

2. International Labour Office (Geneva, 1969).

3. These are described more fully in Donald J. Treiman, *Occupational Prestige in Comparative Perspective* (New York, 1976), ch. 8.

4. I have shown that the Standard Scale does a uniformly better job of predicting occupational prestige hierarchies in individual countries than do occupational status scales developed specifically for use in occupational mobility studies in those countries. See Donald J. Treiman, "Problems of Concept and Measurement in the Comparative Study of Occupational Mobility," *Social Science Research*, IV (1975), 183–230.

5. I.L.O., *International Standard Classification*, 95.

6. Cf. Edward Shils, "Deference," in John A. Jackson (ed.), *Social Stratification* (Cambridge, 1968), 104–132.

7. Some readers will recognize the similarity between this theory and that of Kingsley Davis and Wilbert E. Moore, "Some Principles of Stratification," *American Sociological Review*, X (1945), 242–249. The principal difference between the two lies in the claim by Davis and Moore that prestige is granted by society as an inducement to competent people to fill important jobs. My claim is that occupational income may be seen as such an inducement but that prestige must be viewed as a measure of moral worth, that is, of the extent to which an occupation embodies that which is valued by members of society. Since power and privilege are universally valued and since hierarchies of power and privilege are relatively invariant, prestige will also be relatively invariant.

8. The education data typically derive from the population censuses of each of the countries in question. The income data also typically derive from population censuses, but in some cases they are from enterprise censuses. Ordinarily annual income was utilized but in some instances weekly or monthly wage rates were available rather than annual income. In practice, alternative measures of the relative income of occupational groups tend to be highly correlated, despite differences among occupations in part-time or seasonal employment rates. See Treiman, *Occupational Prestige*, ch. 5, Tables 5.1 and 5.2. Data on both income and education levels were available for only five countries: the U.S., Canada, India, Taiwan, and Yugoslavia.

21 ■ John H. Goldthorpe and Keith Hope

Occupational Grading and Occupational Prestige

Introduction

Over the last forty years or so, there has accumulated in the literature of sociology and social psychology a relatively large number (probably several score) of studies in which respondents have been required to grade a selection of occupations in some hierarchical fashion. It has become customary to refer to such studies as being ones of ‘occupational prestige’. Indeed, when the matter of occupational prestige is now considered, it is almost invariably in terms of studies of the kind in question. Furthermore, the data provided by these enquiries have come to play an important part both in theoretical discussion and in the conduct of empirical investigation in the general problem-area of social stratification and mobility. Yet, oddly enough, ‘occupational prestige’ studies have rarely been subjected to critical examination other than from a technical point of view. . . .

The Meaning of Prestige

An appropriate starting point for a more radical appraisal of ‘occupational prestige’ studies than seems hitherto to have been made is with the concept of ‘prestige’ itself. In a sociological context, we would suggest, prestige can be most usefully understood as referring to a particular form of social advantage and power, associated with the incumbency of a role or membership of a collectivity: specifically, to advantage and power which are of a *symbolic*,

rather than of an economic or political nature. That is to say, such advantage and power imply the ability of an actor to exploit—in the pursuit of his goals—*meanings* and *values*, rather than superior material resources or positions of authority or of *force majeure*.

From this conception it follows that a hierarchy of prestige is constituted by intersubjective communication among actors, and must therefore be characterized in attitudinal and relational terms. It cannot be characterized—other than misleadingly—as a distribution in which units have differing amounts of some particular substance or quality. As a provisional statement, a prestige hierarchy might be one in which actors

(i) *defer* to their superiors—that is, acknowledge by speech or other action their own social inferiority—and seek, or at least appreciate, association with superiors;

(ii) *accept* their equals as partners, associates etc. in intimate social interaction—entertainment, friendship, courtship, marriage, etc.;

(iii) *derogate*¹ their inferiors, if only by accepting their deference and avoiding association with them other than where their own superiority is confirmed.

The attributes of roles or collectivities which differentiate actors in respect of their prestige are various. What they have in common is some symbolic significance—some generally recognised meaning—which, in conjunction with prevailing values, constitutes a claim to

John H. Goldthorpe and Keith Hope. “Occupational Grading and Occupational Prestige,” in *The Analysis of Social Mobility: Methods and Approaches*, edited by Keith Hope, pp. 23–24, 26–27, 30–37, 78–79. Copyright © 1972 by Oxford University Press. Used by permission of Oxford University Press.

social superiority or, conversely, some stigma of inferiority. For example, having the role of doctor and working in a hospital or clinic implies having knowledge of, control over and close involvement with matters which are generally regarded as ones of ultimate concern—matters of life and death. Belonging to an aristocratic family and owning a landed estate signifies descent from illustrious forebears and participation in an historically-rooted, distinctive and exclusive way of life. Working as a clerk in a bank evokes such generally valued characteristics as honesty, trustworthiness, discretion and dependability, and again in relation to ‘important’—in this case, financial—matters. In all of these cases, then, ‘deference-entitlements’ (Shils, 1968) exist, and are likely to be honoured at least by some actors in some contexts. In contrast, being, say, a gypsy scrap-metal dealer or a West Indian refuse-collector is likely to mean relatively frequent exposure to derogation, on account both of the symbolic significance of the ethnic memberships in question and of the implied occupational contact with what is spoiled, discarded and dirty.² In other words, prestige positions do not derive directly from the attributes of a role or collectivity ‘objectively’ considered, but rather from the way in which certain of these attributes are perceived and evaluated in some culturally determined fashion. . . .

Occupational Prestige

Assuming that a conception of prestige consistent with classical analyses is adopted, then the reference of ‘occupational prestige’ follows from it directly: it is to the chances of deference, acceptance and derogation associated with the incumbency of occupational roles and membership in occupational collectivities. Such prestige will be related to the ‘objective’ attributes of occupations—their rewards, requisite qualifications, work-tasks, work environments etc.—but only indirectly: only, that is, in so far as these attributes carry symbolic significance of a kind

that is likely to be interpreted as indicative of social superiority or inferiority, with corresponding interactional consequences.

We may, therefore, now go on to ask such questions as: (i) whether such a conception of occupational prestige has been that generally held by the authors of conventional occupational prestige studies; (ii) whether the results of such studies provide valid indicators of prestige in the sense in question; (iii) whether the uses to which results have been put have been appropriate ones. . . .

The Interpretation of Occupational Prestige Ratings

It has been regularly remarked that in occupational prestige ratings, as conventionally carried out, both cognitive and evaluative processes are involved. However, precisely what are supposed to be the objects of these processes has rarely been made clear. For example, if it really were occupational prestige in the sense we would favour which was being assessed, then what would have to be cognized (or, rather, *recognized*) and evaluated would be the symbolic significance of certain features of an occupation with regard to the chances of those engaged in the occupation meeting with deference, acceptance or derogation in their relations with others. If, for instance, the occupational ‘stimulus’ given were that of ‘coal miner’, a possible response might be on the lines of

‘dirty, degrading work’ →
 ‘rough, uncultivated men’ →
 ‘likely to be looked down on by most groups’

or, alternatively perhaps

‘difficult, dangerous work’ →
 ‘able, courageous men’ →
 ‘likely to be respected by many groups’.

But is this in fact the kind of thing that usually happens? There is little reason to

believe so, at least if we are guided by respondents' own accounts of what chiefly influenced their ratings.³ Rather, we would suggest, the operation that most respondents have tended to perform (perhaps in accordance with the principle of least effort) is a far more obvious and simple one: namely, that of rating the occupations on the basis of what they know, or think they know, about a number of objective characteristics, evaluated in terms of what they contribute to the general 'goodness' of a job. In other words—and consistently with their own accounts—respondents in occupational prestige studies have not typically been acting within a distinctively 'prestige' frame of reference at all. The sensitivity to symbolic indications of social superiority and inferiority which this would imply has not usually been evoked by the task of grading set them. Rather, this task has led them to assess occupations only in some far less specific fashion, according to a composite judgment on an assortment of their attributes which might be thought of as more or less desirable.⁴

Such an interpretation of what 'occupational prestige' ratings are actually about would seem, moreover, to fit far better with what is known of the pattern of variation in such ratings than would the idea that they relate to prestige *stricto sensu*. The basic feature of this pattern is that while some considerable amount of disagreement in rating occurs as between *individuals*, differences between the mean ratings of age, sex, regional, occupational and other collectivities are never very great. If one assumes that in making their judgments, respondents more or less consciously (i) consider a number of different occupational attributes which they take as determining how 'good' a job is; (ii) attach some subjective 'weight' to each of these; (iii) for each occupation presented apply their rating 'formula' to what they know about the occupation, and thus (iv) come to some overall assessment of it—then one might well anticipate some appreciable degree of variation

in ratings at the individual level. Individuals are likely to differ in their familiarity with particular jobs and in their priorities as regards what makes a job 'good'. However, one would not expect—other than in somewhat special and limited cases⁵—that such differences would be socially structured in any very striking way. Knowledge about the more general characteristics of other than rather esoteric occupations is relatively 'open'; and, again in general terms, the kinds of thing thought of as 'good' in a job are unlikely to give rise to systematic differences in ratings, especially since there is, in any case, a clear tendency for such advantages to go together. To take a particular example—from the NORC data—it is not surprising, given an interpretation of the kind we have proposed, that individuals should quite often disagree about the ratings of 'building contractor' *vis-à-vis* 'welfare-worker'—*nor* that, at the same time, in the case of age, sex, regional, occupational or other categories, the former job should invariably have the higher *mean* rating. (See Reiss 1961, pp. 55–6, 225–8).⁶

On the other hand, if we were to suppose that 'occupational prestige' scores did give a valid indication of a structure of prestige relations, then the degree of consensus that is shown among different social groups would indeed be remarkable, at least in those societies where other research has indicated some notable diversity in value systems and in particular between members of different social strata. For in this case it would not be a matter of evaluative consensus simply on what attributes make a job 'good', but rather on certain symbolic criteria of generalized superiority and inferiority, with all their attitudinal and behavioural implications. As Shils has observed, the conditions necessary for an entirely, or even a largely, 'integrated' prestige order to exist are in fact demanding ones. It would seem, therefore, the safest assumption to make that, within modern industrial societies, such conditions will prevail only locally, transiently or imperfectly, and

thus that social relations expressive of a prestige order will occur only in an intermittent or discontinuous fashion. On the basis of available empirical data, one might suggest that while derogation is still quite widely manifest—as, for example, in the form of differential association or status-group exclusivity—the claim to superiority thus made by one group is not necessarily, or even usually, acknowledged by those regarded as inferior; that is to say, the latter are often not inclined to display deference.⁷ This refusal may be revealed passively—by disregard for the claim to superiority, in that no particular ‘respect’ is shown, and little concern to reduce social distance from the ‘superior’ group; or, perhaps, some direct challenge to the claim may be made where real interests are felt to be threatened by it—as, say, by ‘exclusivity’ in housing areas, use of amenities, etc. . . .

The Uses of ‘Occupational Prestige’ Ratings

One notable use of the data in question results from the fact that over the last two decades occupational prestige studies have been carried out in a steadily increasing number of countries at different levels of economic development. The opportunity has therefore arisen of making cross-national comparisons which, it has been supposed, can throw light on the relationship between value systems and social structural characteristics and are thus relevant to the thesis of the ‘convergent’ development of societies as industrialism advances. For example, Inkeles and Rossi (1956), comparing occupational prestige ratings in studies from six industrial societies, showed that a high degree of similarity prevailed. On this basis, they concluded that common structural features of these societies were of greater influence on the evaluation of occupations than were differences in cultural traditions. Subsequently, however, occupational prestige ratings from several countries as yet little industrialized

have *also* been shown to be broadly in line with the hierarchy found in economically advanced societies—in so far, that is, as comparisons can be made. This result has then led to the modified argument (Hodge, Treiman and Rossi, 1966) that what is chiefly reflected in prestige ratings is the set of structural features shared by national societies of *any* degree of complexity—‘specialized institutions to carry out political, religious, and economic functions, and to provide for the health, education and welfare of the population . . .’. Occupations at the top of these institutional structures, it is suggested, are highly regarded because of their functional importance and also because they are those which require the most training and ability and those to which the highest rewards accrue. Thus, ‘any major prestige inversion would produce a great deal of inconsistency in the stratification system’ (p. 310).

In this way, therefore, it is clearly indicated how occupational prestige data may further be employed in support of a general theory of social stratification of a structural-functional type. Such an application has in fact been made quite explicitly in the work of Barber (1957). Following a Parsonian approach, Barber takes the results of the Inkeles-Rossi study as the main empirical foundation for the view that the factual order of stratification in modern societies tends in the main to be consistent with the dominant normative order. Inequality in social rewards and relationships, it is held, is structured in accordance with functional ‘needs’, and this arrangement is then seen as receiving general moral support: ‘functionally important roles are congruent with or partly determine a system of values’ (p. 6).

Clearly, for occupational prestige data to be used in the ways in question, it is necessary to assume that such data reflect prevailing values and norms *of a particular kind*: ones pertaining to the ‘goodness’—in the sense of the ‘fairness’ or ‘justice’—of the existing distribution of social power and advan-

tage. However, in view of our previous discussion, it is difficult to regard such an assumption as a valid one or indeed to understand why it ever should have been made. Even if it were to be supposed that data on publicly recognized occupational hierarchies do indicate a prestige order in something approximating the classical conception, it still then would not follow that they can provide evidence that the objective reality of stratification is morally legitimated. For while prestige relations do depend upon a certain range of shared understandings, consensus on principles of distributive justice is not necessarily involved.⁸ Moreover, as we have argued, by far the most plausible interpretation is that occupational prestige ratings reflect prevailing ideas at a much lower level of abstraction: that is, ideas of what is 'good' in the sense simply of what is generally found desirable in an occupation. And if *this* is the case, then the consensus that exists is obviously of no very great moral or legitimatory significance at all. Apart from quite unsurprising agreement on such matters as, for example, that high pay is preferable to low pay, more security to less, qualifications to lack of qualifications, etc., the consensus that is implied is of a cognitive and perceptual kind, not an evaluative one. The fact that, on average, all groups and strata agree that certain occupations should be rated higher than others tells one nothing at all about whether the occupational hierarchy that is thus represented is regarded as that which *ought* to exist. And in so far as the publicly recognized hierarchy corresponds to that proposed by structural-functional theorists, this would seem to indicate no more than that broadly similar sets of rating criteria are being applied: i.e. occupational rewards and occupational requirements.⁹

Thus, as regards the utilization of occupational prestige data in the advancement of stratification theory, our view must be that this has been fundamentally misguided. What, now, of their application in research? Primarily, of course, occupational prestige

ratings have been used in studies of social mobility, in which they have constituted the hierarchy—scalar or categorical—in the context of which mobility has been assessed. Assumptions about what prestige ratings rate are thus necessarily involved in the interpretation of mobility patterns, and the crucial issues that arise are once more ones of 'validity'.

Concerning the question: What, in mobility studies, may occupational prestige ratings be taken to indicate?—three main positions can be distinguished. These can be usefully considered in turn, together with their implications and problems.

(i) Ratings may be taken—as, for example, by Svalastoga (1959)—as indicative of the position of an occupation within a prestige order; that is, as indicative of the chances of those holding that occupation encountering deference, acceptance or derogation in their social lives. In this case, therefore, mobility between different occupational levels, other than of a marginal kind, may be interpreted as involving the probability of subcultural and relational discontinuity. While such a perspective does not necessarily mean that society is seen as divided up into more or less discrete strata, it does imply that social mobility, as measured, is not just a matter of individuals gaining more qualifications, more income, more interesting work etc., but further of their experiencing changes in their life-styles and patterns of association. The difficulty is, however, as already remarked, that the validity of occupational prestige ratings construed in this way has never been established, and that there are indeed strong grounds for doubting their validity. In other words, we are simply not in a position to infer, with any acceptable degree of precision and certitude, what are the typical consequences of mobility, as measured via occupational prestige ratings, for the actual social experience of those deemed to be mobile.

(ii) Prestige ratings may be taken as indicative of the status of occupations in the generic sense—that is, as being in effect comparable with composite measures of ‘socio-economic’ status, derived from data on income, education, housing, possessions etc. Justification for this position is twofold: first, to [quote] the observation of Reiss (1961), respondents in prestige-rating studies appear ‘to emphasize the relevance of indicators sociologists use to measure socio-economic status . . .’; secondly, as shown by Duncan (1961), it is possible, at least in the American case, to predict prestige ratings fairly accurately from census data on occupational income and education. If then, ‘occupational prestige’ is understood in the way in question, some reasonable basis may be claimed for interpreting occupationally-measured mobility in terms of movement between grades of occupation differentiated chiefly by their levels of rewards and requirements. At the same time, though, it must be emphasized that in this case no good grounds exist for any interpretation in terms of prestige *stricto sensu*, and, of course, no basis at all for any consideration of how far mobility may be incongruent from one form of stratification to another. Precisely because of the inevitably ‘synthetic’ nature (Ossowski, 1963) of socio-economic status, as indicated by prestige scores, the analysis of mobility must be strictly unidimensional. These limitations would lead one to suggest, therefore, that if it is accepted that occupational prestige ratings are not valid indicators of a prestige order but are being used simply to stand proxy for socio-economic status, then it would be preferable, where possible, to seek to measure the latter more directly—and without any concern to combine components so that a good ‘fit’ with prestige scores may be obtained. To discard the notion of prestige altogether would, in this case, mean losing nothing but the possibility of terminological confusion; and developing separate indices of occupational income, education etc., as well

as some composite measure, would permit the analysis of mobility in a multi-dimensional manner. In short, there seems no good argument for basing mobility research on occupational prestige ratings, interpreted as socio-economic status scores, other than where a lack of data on the socio-economic attributes of occupations makes this procedure an unavoidable *pis aller*.

(iii) Prestige ratings may be taken as indicating popular evaluations of the relative ‘goodness’ of occupations in terms of the entire range of prevailing criteria. In this case, related mobility data are open to interpretation as showing, basically, the chances of individuals entering more or less desirable grades of occupation, given certain grades of origin. While an interpretation of the data on these lines has rarely, if ever, been pursued consistently throughout a mobility study, it is that which, on grounds of validity, could best be defended. First, as we have already argued, grading occupations according to notions of their general ‘goodness’ is what respondents in occupational prestige studies appear, in the main, to be doing. Secondly, it is in regard to *this* understanding of prestige scores that it would seem most relevant to claim, following Duncan and Artis (1951) and Reiss (1961), that their validity lies in the degree of consensus which emerges, despite the use of quite various criteria of evaluation. The argument that this consensus points to ‘the existence of an underlying and agreed upon structure of occupational prestige’ is difficult to sustain once it is recognized just what consensus on a prestige order entails. But the idea of a broadly agreed upon ordering of occupations in terms of ‘goodness’ does, on the evidence in question, receive some clear—and not very surprising—support. Furthermore, if prestige ratings are taken as indicative of an occupational hierarchy of this kind, then the fact that they represent synthetic judgments and cannot be ‘disaggregated’ is no longer a problem in the analysis of mobility patterns. For if

mobility is being interpreted as being simply between grades of occupation of differing desirability in some overall sense, a unidimensional approach would appear the appropriate one. However, it must be added that what would then be a dubious and potentially dangerous step would be to shift from such an interpretation of specifically occupational mobility to one in which conclusions were drawn regarding the stability of status groups, income classes, or social strata in any sense whatsoever; that is, conclusions regarding *social* mobility as generally understood. In effect, of course, a shift of this nature has been made in most large-scale mobility studies carried out in the recent past. But while it might reasonably be held that such a manoeuvre is unlikely to be very misleading so far as the 'gross' patterns of social mobility are concerned, the difficulty is (apart from the limitation of unidimensionality) that we have no way of knowing at just *what* point and in *what* ways it might turn out to be quite deceptive. Yet again, the problem of validity recurs.

The general—and rather pessimistic—conclusion to which one is led is, therefore, the following: that to the extent that the meaning of occupational prestige ratings is correctly construed, the less useful they appear to be as a basis for mobility studies which pursue the 'classical' sociological interests of mobility research.

NOTES

1. We use 'derogate' in this context following Shils (1968). Were it not that its usual connotations go beyond its strict meaning, 'disparage'—literally 'to make unequal'—might be a preferable term.

2. On 'stigma symbols' as the obverse of 'prestige symbols', see Goffman (1963).

3. See Reiss (1961).

4. As regards the NORC [National Opinion Research Center] study, it is worth recalling what is usually forgotten: that this enquiry, at least in the view of those who devised it, was in fact specifically

aimed at finding out what people thought were the best jobs, in the sense of the most desirable. Where 'prestige' and 'standing' are referred to in the initial report on the study, they are obviously equated with desirability. See NORC (1947).

5. E.g. where respondents are rating occupations within their own status or situs areas, c.f. Gerstl and Cohen (1964).

6. Our interpretation of the meaning of 'occupational prestige' ratings is also consistent with the fact that certain variations in the task set to respondents appear to make little difference to the results achieved: e.g. whether respondents are asked to rate occupations according to their 'social prestige', 'social standing', 'social status', 'general desirability' etc: or whether they are asked for their own opinions or what they believe are generally prevailing opinions. It seems reasonable to suppose that if respondents are required to grade occupations according to any one criterion which, while rather imprecise, implies a 'better-worse' dimension, they will produce results of the kind in question; and further, that the level of consensus in this respect is such that the distinction between personal and general opinion is of little consequence—provided that there is no suggestion of a normative judgment being required; that is, one in terms of which jobs *ought* to be the best.

7. Cf. for example, Goldthorpe, Lockwood, Bechhofer and Platt (1969), chapters 4 and 5.

8. In fact, one might suggest the hypothesis that societies of the kind in which an integrated and stable prestige order is to be found will tend to be ones in which the factual order of stratification is not commonly appraised in terms of distributive justice, or indeed envisaged as capable of being in any way substantially different from what it is.

The distinction between the recognition of prestige and the attribution of justice is foreshadowed—as are several other points in the above paragraph—by Gusfield and Schwartz (1963) in a paper that has been curiously neglected by subsequent American writers on occupational grading.

9. It is a well-known problem of the structural-functional theory of stratification that other usable criteria of the functional importance of occupational roles are hard to find: employing the two criteria in question does, of course, introduce a serious degree of circularity into the argument.

REFERENCES

Barber, B. (1957). *Social stratification*. Harcourt, Brace & Co., New York.

- Duncan, O.D. (1961). A socioeconomic index for all occupations. In A.J. Reiss, *Occupations and social status*. Free Press of Glencoe, New York.
- Duncan, O.D. and Artis, J.W. (1951). *Social stratification in a Pennsylvania rural community*. Pennsylvania State College: Agricultural Experiment Station Bulletin 543.
- Gerstl, J. and Cohen, L.K. (1964). Dissensus, situs and egocentrism in occupational ranking. *Brit. J. Sociol.*, 15, 254–61.
- Goffman, E. (1963). *Stigma*. Prentice Hall, Englewood Cliffs.
- Goldthorpe, J.H., Lockwood, D., Bechhofer, F. and Platt, J. (1969). *The affluent worker in the class structure*. Cambridge University Press.
- Gusfield, J.R. and Schwartz, M. (1963). The meaning of occupational prestige: reconsideration of the NORC scale. *Amer. Sociol. Rev.*, 28, 265–71.
- Hodge, R.W., Treiman, D.J. and Rossi, P. (1966). A comparative study of occupational prestige. In *Class, status and power* (2nd edn.). (ed. R. Bendix and S.M. Lipset). Free Press of Glencoe, New York.
- Inkeles, A. and Rossi, P. (1956). National comparisons of occupational prestige. *Amer. J. Sociol.*, 61, 329–39.
- N.O.R.C. (1947). Jobs and occupations: a popular evaluation. *Opinion News*, 9th September. Reprinted in *Class, status and power* (1st edn.). (ed. R. Bendix and S.M. Lipset). Free Press of Glencoe, New York.
- Ossowski, S. (1963). *Class structure in the social consciousness*. Routledge, London.
- Reiss, A.J. (1961). *Occupations and social status*. Free Press of Glencoe, New York.
- Shils, E.A. (1968). Deference. In *Social stratification* (ed. J.A. Jackson). Cambridge University Press.
- Svalastoga, K. (1959). *Prestige, class and mobility*. Gyldendal, Copenhagen.

22 ■ David L. Featherman and Robert M. Hauser

Prestige or Socioeconomic Scales in the Study of Occupational Achievement?

At least in the United States and Australia, the processes of allocation to educational and occupational statuses from social origins (i.e., the process of stratification or of status attainment) seem largely socioeconomic in character (Featherman, Jones, and Hauser, 1975). Put another way, inter- and intragenerational movements of men among categories of their own and their parents' educations and occupations more closely follow the dimensions of social

space defined by the "socioeconomic" distances among occupation groups than by the "prestige" distances among occupations. Evidence for this interpretation is drawn from parallel results for the United States and Australia in which estimates for the structural equations of "status attainment" models with occupations scaled in units of Duncan's (1961) socioeconomic index (SEI) yield higher coefficients of multiple determination (R^2) than do esti-

mates based on occupations scaled in units of NORC prestige (Siegel, 1971) or of Treiman's (1977) international prestige index. In addition, the canonical structure of generational and career occupational mobility in both societies more nearly approximates a socioeconomic "space," as the canonical weights for occupation categories correlated higher with mean SEI scores for these occupations than with mean Siegel or Treiman scores.

In interpreting these data we suggest that prestige scores for occupations are less valid indicators of the dimensions of occupations pertinent to occupational mobility in industrial societies and of the status attainment processes operating therein than are socioeconomic scores. We reason from evidence for the United States (Reiss, 1961; Siegel, 1971) and Great Britain (Goldthorpe and Hope, 1974) that occupational prestige scores represent a congeries of salient dimensions or occupational characteristics. For example, the British ratings of the "social standing" of occupations are a linear combination (to the extent of 97% of their variance) of four oblique dimensions: standard of living, power and influence over other people, level of qualifications, and value to society (Goldthorpe and Hope, 1974: 14). Any two pairs of raters produce rankings which are modestly correlated at best ($r = .4$), consistent with the notion that unique variance in prestige gradings is quite high. Conversely, the mean ranks for the same occupations over socially and demographically defined groups correlate in the range of 0.8 and 0.9. This common variance appears to be socioeconomic; that is, over three-quarters of the linear variance in prestige scores is a reflection of the educational and economic properties of the ranked occupations. Thus, while raters in the United States and Britain used many and idiosyncratic features of occupations in assessing their relative social standing, apparently they all were aware of

and utilized the socioeconomic "desirability" of titles, to some extent, in reaching their decisions.

The salience of the socioeconomic properties of occupations across persons, groups, and perhaps societies may follow from the rather similar social organization of occupations in functionally similar economic systems (e.g., industrial capitalism). But more to the point of the relative centrality of "prestige" or socioeconomic dimensions to the process of status attainment, we speculate that commonalities in prestige grades and in the responsiveness of these rankings to socioeconomic attributes of occupations may reflect popular awareness of (what further comparative research may show to be) similar processes of status allocation across societies. In at least the cases of Australia and the United States, the socioeconomic model, patterned after the work of Blau and Duncan (1967), yields estimates of effect parameters which are substantially the same. Moreover, log-linear adjustments of mobility matrices for the effects of differential occupation structures (to wit, as provided in the table margins) uncovers largely similar interactions within the tables (to wit, constant patterns of inflow and outflow both between and within generations for both societies).

Our provisional conclusion is that prestige scores are "error-prone" estimates of the socioeconomic attributes of occupations. Whatever it is that prestige scores scale—and this does not appear to be prestige in the classical sense of deference/derogation (see Goldthorpe and Hope, 1972)—it is substantively different from socioeconomic status. Yet one is best advised to use a scale for occupations which most accurately captures the features of occupations having force for the social process one is studying. In instances of occupational mobility and related processes of status allocation, socioeconomic dimensions and socioeconomic scores for occupations

are the more central, and therefore are preferable over prestige scores.

NOTES

This research was supported by NSF Grant #44336, NICHD Grant #HD-05876 and institutional support from the College of Agricultural and Life Sciences, University of Wisconsin-Madison. Any opinions, findings, conclusions, or recommendations are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

Blau, P. M. and O. D. Duncan (1967) *The American Occupational Structure*. New York: Wiley.
 Duncan, O.D. (1961) "A socioeconomic index for all occupations," pp. 139–161 in A. Reiss, Oc-

cupations and Social Status. New York: Free Press.
 Featherman, D. L., F. L. Jones, and R. M. Hauser (1975) "Assumptions of social mobility research in the U.S.: the case of occupational status." *Social Science Research* 4 (December): 329–360.
 Goldthorpe, J. and K. Hope (1974) *The Social Grading of Occupations: A New Approach and Scale*. Oxford: Oxford Univ. Press.
 ———. (1972) "Occupational grading and occupational prestige," pp. 19–79 in K. Hope (ed.) *The Analysis of Social Mobility: Methods and Approaches*. Oxford: Clarendon.
 Reiss, A. J. (1961) *Occupations and Social Status*. New York: Free Press.
 Siegel, P. M. (1971) "Prestige in the American occupational structure." Unpublished doctoral dissertation, University of Chicago.
 Treiman, D. J. (1977) *Occupational Prestige in Comparative Perspective*. New York: Academic Press.

■ ARE SOCIOECONOMIC SCALES OBSOLETE? ■

23 ■ Robert W. Hodge

The Measurement of Occupational Status

Since the appearance of Blau and Duncan's monumental inquiry into "The American Occupational Structure" (1967), we have learned a great deal about processes of inter- and intra-generation occupational mobility. Indeed, there has been a virtual explosion of research on processes of status attainment. Much of this research rests upon the reduction of information about a person's detailed occupational pursuit to a single continuous variable, a transformation typically accomplished by utilizing Duncan's Socioeconomic Index for All Occupations (Duncan, 1961a) to assign status scores to the occupations held by fathers and sons at various points in their careers (see, for example, Blau and Duncan, 1967; Hauser and Featherman, 1977). Despite the reliance of most inquiries into processes of status attainment on Duncan's SEI scale, there has been little discussion of the properties and characteristics of this index by its users (see, however, Duncan, 1961b, and Featherman and Hauser, 1976). The purpose of this essay is to discuss the characteristics of Duncan's SEI scale, as well as several difficulties encountered in its use in studies of occupational mobility. . . .

On the Interpretation of Duncan's Index

The conceptual meaning of Duncan's SEI scale is by no means clear; at least three al-

ternative interpretations are available and none of these is entirely satisfactory. All of these interpretations rest upon features of the construction of Duncan's index and/or characteristics of the estimated weights of its components.

The most obvious interpretation of Duncan's SEI scale follows from the technique by which the weights of its components were derived. The reader will recall that they were established by regressing the percentage of excellent plus good ratings received by a few titles in the North-Hatt study which matched census lines on census-derived indicators of the age-standardized educational and income levels of these occupations. Scale values for all occupations were then obtained by substituting the education and income measures, available for all occupations from census data, into the resulting equation. Consequently, Duncan's SEI scale may be interpreted as *the expected percentage of excellent plus good ratings an occupation would receive in a prestige inquiry of the North-Hatt type*.

There are two defects with this interpretation of Duncan's SEI scale. First, the prediction equation for the prestige indicator is less than satisfactory. It accounts for a bit more than four-fifths of the variance in prestige ratings. One could, of course, regard the error variance as random, on the view that prestige ratings are just "error-prone" proxies

for the education and income levels of occupations. However, such a view cannot be sustained in the light of the substantial consensus which exists between subgroups of raters. The education and income levels of occupations fail to account for the consensus observed between subgroups of raters differing in their own occupations, their sex, race, and so forth. Not only is there consensus between subgroups of raters about overall prestige ratings, there is also consensus about that part of the prestige of an occupation which is not accounted for by the income and educational levels of its incumbents. This fact enables one to discount the view that prestige scores are just “error-prone” indicators of the “socioeconomic” level of an occupation, a point which appears to have escaped Featherman and Hauser who state (1976, p. 405), “Our provisional conclusion is that prestige scores are ‘error-prone’ estimates of the socioeconomic attributes of occupations.” This claim is quite possibly true *with respect to the intergenerational transmission of occupational status*. However, in view of the consensus over them from one subgroup of raters to the next, the “errors” themselves appear to be social facts in Durkheim’s sense, rather than random disturbances which have no life of their own. For this reason, the interpretation of Duncan’s SEI scale as a predicted prestige score flies in the face of what is known about occupational prestige.

Since the publication of Duncan’s SEI scale, pure prestige scales have become available for all occupations (Siegel, 1971; Treiman, 1977). Comparisons of the performance of these scales with Duncan’s index in studies of status attainment leave no doubt that the association between the detailed occupations of fathers and their sons is captured more completely by their values on Duncan’s index than on either Siegel’s or Treiman’s prestige scale (see, for example, Duncan *et al.*, 1972; Featherman and Hauser, 1976; Stevens and Featherman, 1981). This is yet another reason why the in-

terpretation of Duncan’s scale as expected prestige scores is dubious.

A second interpretation of Duncan’s SEI scale pays no attention to the method of its construction and makes reference only to its components. Without specifying the precise meaning of either, sociologists commonly make a distinction between “social status” and “economic status.” (These concepts, whatever they are, should not be confused with Weber’s concepts of “status honor” and “class,” which have quite specific meanings that are analytically, if not statistically, independent of the usual measures of “social status” and “economic status.”) Education is frequently utilized as an indicator of “social status,” while current income is a common measure of “economic status.” Since aggregate measures of the educational and income levels of an occupation’s incumbents enter into the computation of Duncan’s index, it is natural to refer to the combination of them as a *socioeconomic index of occupational status*.

This interpretation of Duncan’s SEI scale is, obviously, the one most frequently made in the literature. Duncan indicated his own preference for it by his decision to name his index as he did. The socioeconomic interpretation of Duncan’s SEI scale is clearly embedded in its use in Blau and Duncan’s study of occupational mobility (1967), Hauser and Featherman’s replication of it (1977), and Featherman and Hauser’s important discussion (1976) of the properties of socioeconomic and prestige indicators of occupational standing. Most users of the Duncan index have accepted this interpretation without serious consideration of alternatives. Despite the overwhelming consensus in the published literature about the proper interpretation of Duncan’s scale, it is interesting to note that Duncan himself, in his original presentation of the scale, is more than slightly ambiguous about its proper interpretation. At one point, he remarks (1961a, p. 115),

Our problem, then, is defined as that of obtaining a socioeconomic index for each of the occupations in the detailed classification of the 1950 Census of Population. This index is to have both face validity, in terms of its constituent variables, and *sufficient predictive efficiency with respect to the NORC occupational prestige ratings that it can serve as an acceptable substitute for them* [emphasis is added] in any research where it is necessary to grade or rank occupations in the way that the NORC score does but where some of the occupations are not on the NORC list.

This quotation seems to make clear that Duncan wanted to cut the cake both ways: the index was a socioeconomic one, but it was also a substitute for prestige ratings. Subsequently, Duncan made clear that he did not regard his index as the equivalent of a pure prestige scale, noting (1961a, p. 129) that, “It should be made perfectly clear that the socioeconomic index does *not* [Duncan’s italics] purport to be a prediction of the prestige ratings that occupations excluded from the NORC list would receive in a similarly conducted study of prestige ratings.” Subsequent research has clearly demonstrated that Duncan was absolutely correct in this judgment: his scale and prestige scales are very definitely not the same thing. But that still leaves open how his scale should be interpreted.

The interpretation of Duncan’s SEI scale as a socioeconomic index is seemingly agreeable and consistent with what its constituent indicators are thought to measure at the individual level. Nonetheless, this interpretation of Duncan’s and similar scales is not without its problems. In our view, there are two primary difficulties with socioeconomic indicators of an occupation’s location in the social structure of work. First, the combination of indicators of social and economic status such as education and income into a composite index of socioeconomic level begs the question of whether or not the effects of these factors are proportional to their weights in the index—a crucial and

necessary assumption whenever the index is subsequently employed in empirical research. There is ample evidence *at the individual level* that education and income can even have *effects of opposite sign* on some dependent variables such as fertility; combining them together with a person’s occupational level into an overall index of an individual’s socioeconomic level presumes unidimensionality where there is none and should be avoided. (For a further discussion of this point and additional examples, see Hodge, 1970.) But if it is sound practice to keep such variables as income and education separated at the individual level, one can at least question the wisdom of combining them *at the aggregate level* of occupations. It may well be that alternative, aggregate characteristics of occupations are not just alternative indicators of an occupation’s location in a single hierarchy, but reflect somewhat different forces at work on an occupation’s incumbents and their behavior.

What we regard, however, as an even greater difficulty with the socioeconomic interpretation of Duncan’s SEI scale stems from the analytical status of the concept of “socioeconomic” level. As far as we can see, it has none: its relationship to such *well-defined*, though *poorly measured*, concepts of stratification theory as “class,” “status,” and “power” is at best vague and imprecise. The concept of socioeconomic status has no independent analytical status in stratification theory: at the individual level it is no more or no less than whatever is measured by a person’s education, occupational pursuit, and income and, at the aggregate level of occupations, it is just some combination of whatever skills it takes to enter the occupation and whatever rewards are obtained from pursuing it at a given point in time within a given market structure. Socioeconomic status is what socioeconomic status scales measure; there is no underlying analytical concept to which we can refer a proposed indicator of socioeconomic status to decide whether it is well or ill conceived or to assess how it might be improved. For example,

referring to “the choice of summary statistics to represent the education and income distributions of the occupations, and the adjustment of these statistics for age differences among occupations,” Duncan observed (1961a, p. 119), “Reasonable procedures for accomplishing these two steps, different from the ones followed [in the construction of Duncan’s index] are easily proposed.” In a world like this, of course, there is no theoretical justification for the choice of one, as opposed to another plausible means of summarizing the constituent indicators and the choice of any one of the competing alternatives rests on the assumption that they all measure the same thing in approximately the same way. As Duncan put it (1961a, p. 119), “. . . it seems doubtful that the final result would be greatly altered by switching to one of the alternatives.” Whether this is literally the case, we cannot say, for while students of status attainment have been especially diligent at exposing the weaknesses of pure prestige scales in the study of occupational mobility, they have devoted but limited energy to examining the properties of alternative socioeconomic scales, to exposing the behavior of the component variables in these scales, and to making these scales temporally relevant to their research, a fault remedied in considerable measure by the work of Stevens and Featherman (1981).

We offer now a third and final interpretation of Duncan’s SEI scale. In Duncan’s SEI index, the weights of the education and income variables are nearly equal and the intercept is close to zero. We can find the constant k which will center the coefficients of the income and education variables in Duncan’s SEI scale around .5 by solving $.50 - .55k = .59k - .50$ for $K = (1)/(1.14) = .8772$.¹ Multiplying the values of Duncan’s SEI scale ($= D$) by this value and adding $6(k) = 6(.8772) = 5.26$ to eliminate the constant term leaves us with

$$\begin{aligned} D' &= k(D) + 6(k) = 8.772(D) + 5.26 \\ &= .4825(E) + 5.175(I), \end{aligned}$$

a transformed index ($= D'$) in which the education ($= E$) and income ($= I$) indicators are for all practical purposes simply averaged together. This transformation of Duncan’s SEI scale is, of course, made possible only because the summary measures of the education and income distributions of the occupations receive nearly equal weights when prestige is regressed on them. Our ability to effect this transformation suggests another interpretation of the Duncan SEI scale scores, viz., *a linear transformation of the best guess we could make of the age-standardized percentage of an occupation’s male incumbents either with at least a high school diploma or with 1949 incomes of \$3500 or more if neither percentage was known*. In fact, one would not go far awry in interpreting the untransformed values of Duncan’s SEI scale in this fashion, since the transformation required to effect this interpretation is roughly equal to the identity operator.

This interpretation of Duncan’s SEI scale is, we believe, novel, though Cain (1974, p. 1501) comes close to making it. And while it sounds ridiculous, that is not necessarily a disadvantage. It keeps one’s attention focused upon the essential feature of Duncan’s SEI scale, to wit, the particular way it glues education and income together to construct “socioeconomic” status. Beyond that, it makes clear the inherent uncertainty which necessarily surrounds any results obtained by the use of Duncan’s SEI scale: having used it, there is absolutely no way of knowing whether the observed effects are brought about by the economic rewards attached to occupations or by the skills required to pursue them. Instead, one must resort to casting the results in terms of an occupation’s “socioeconomic” status—a concept which seems more nearly contrived for convenience than a social fact in Durkheim’s sense. In keeping with the bulk of the literature on status attainment, we will continue throughout this paper to accept the *socioeconomic* interpretation of Duncan’s SEI scale, but it

should be obvious that the serious questions which can be raised about this or any other interpretation of Duncan's scale also make the interpretation of any findings based upon it problematical.

Occupation as a Contextual Variable

In recent years, a considerable amount of sociological inquiry has been directed toward detecting contextual effects on individual behavior. In this research tradition, individual behavior is seen in part as a function of the characteristics of the other individuals with whom the subject shares group memberships (see, e.g., Blau, 1957, 1960; Davis, Spaeth, and Huson, 1961; Tannenbaum and Bachman, 1964; Farkas, 1974). Research of this kind has not been without its critics, of which Hauser (1970a, 1970b, 1974) is by far the most outspoken.

There are several strategies of research for examining so-called group, contextual, and structural effects upon individual behavior. We need not detail these here, although we should note that the most general of all these models is the one embedded in the analysis of covariance. The particular strategy of interest in the present context is the one where the consequences for an individual's behavior of his membership in a social group or population aggregate are summarized by an indicator which reflects the average or some other measure of the central tendency on a particular trait of the individual members comprising the social groups or population aggregates to which he belongs. An example of this research strategy would be characterizing schoolchildren by the proportion of minority group members in the school they attend. Such a characterization of an individual has nothing to do with the structural features of the school he attends, such as library books per capita. Instead, it rests solely upon the individual characteristics of his fellow classmates.

The fundamental difficulty with this strategy for analyzing contextual or composi-

tional effects is put quite simply: there is *no logically conceivable way in which one could run an experiment to test for any observed effects*. This is the fundamental defect with the analysis of all compositional effects; Hauser (1970a) comes close to stating this principle, but does not make it as explicit as he might have. To illustrate this principle, we may pursue the foregoing example. Suppose we wanted to run an experiment to examine the effects of minority composition on school achievement among white students. Now obviously, our first step would be to make random assignments of white students to schools; in this way the white students in each school would be expected to have equivalent means and variances on all characteristics save those we experimentally manipulate; this is the advantage we realize from experimentation. So far, so good. Now we must construct our experimental variable. To do this, we can again make random assignments, this time of minority students, of subjects to schools. However, while we can make random assignments of minority students to schools so that their expected means and variances on all variables are equal from school to school, *we must assign them in differential numbers*. If we failed to do this, the schools would not differ in their minority composition and there would be no between-school variance in our experimental variable. The situation is now this: by making random assignments we have secured an expected equality from school to school in the means and variances of both white and minority students on all variables. However, the ratio of white to minority students varies from school to school and that, necessary to conduct the experiment at all, proves fatal.

Minority and white students *do not differ on their minority status alone*. They differ in their socioeconomic backgrounds, in the numbers of their siblings, in the quality of their experiences, and almost surely in their attitudes and values as well. Because the schools vary in their minority composition,

they will necessarily vary in every individual level correlate of minority status as well. Consequently, any observed effect of minority composition is confounded by every individual correlate of minority status and *there is no logically possible way of experimentally separating these confounding factors.* Having found an effect of minority composition, we can generate additional compositional effects by the carload lot. All we need to do is to refer to the individual level correlates of minority status. Furthermore, although it is not essential to the argument advanced herein, any attempt to separate these confounding factors via nonexperimental methods will certainly flounder on the barricade of multicollinearity, since the relevant associations are the typically high, ecological correlations across the units of the experiment—in this case schools.

The foregoing argument requires comment. First, there is nothing about it which denies the existence of contextual or compositional effects. Indeed, they may be large and substantial, but the argument clearly implies that *there is no logically possible way of isolating them.* Second, while *practically* speaking, most social-science findings are not subject to verification via experimentation, one can at least *logically* conceive of an experiment to test them. Contextual effects, however, are in a different ball park from most other social-science findings. There is no conceivable way of contriving an experiment to test them. That should give one pause, for it is far from clear that the limited resources for social-science research should be expended on discovering effects *whose causes can never be experimentally isolated.* Finally, while it is technically possible to detect a generalized contextual effect—indeed, the experiment outlined above could do that—the impossibility of specifying the precise causal force which generates the effect means that contextual analysis is profoundly and fundamentally irrelevant to policy decisions. Any effort to formulate policy on the

basis of presumed, *specific* contextual effects is foolhardy, for there is no way of knowing whether the effects are generated by the specific causes one has identified or by one of their confounding correlates.

Occupation is not inherently a contextual variable. Although other considerations enter the picture, one can at least think of detailed occupational groupings as clusters of jobs whose incumbents are mutually substitutable. Thus, it is the similarity in the work required by the jobs forming an occupational group, rather than the similarity in the personal characteristics of their incumbents, which delineates one occupation from another. However, once occupational information is scored with a socioeconomic index like Duncan's, occupation *is turned into a contextual variable.* All of the reservations that one might have about contextual analysis similarly apply to the analysis of occupational information coded in this particular way. One can, of course, *logically* conceive of an experiment to test for the effects of occupation as such; though practically such an experiment might be unfeasible, at least logically one can imagine randomly assigning subjects to occupational groups. But there is no logical way to conduct an experiment about *occupational status.* Duncan's SEI scale includes the education and income levels of *incumbents* as its factors; if one tried to run an experiment on occupations by randomly assigning subjects to occupations, the expected means and variances in their educational and income levels would be identical from one occupation to the next and the occupations would no longer be differentiated according to their socioeconomic status as it is measured by Duncan's scale. In order to keep the Duncan scale scores of the occupations differentiated, one would have to assign relatively more high school graduates and high-income earners to some occupations than to others. Once one has done this, the entire advantage of experimentation is lost and one's experiment would be confounded

by every individual level correlate of educational attainment and income. There is just no logically conceivable way to reproduce experimentally any results derived by scoring occupational data with Duncan's scale scores; this fact alone ought to give one pause before using such a scale, particularly in analyses which purport to be causal, rather than merely descriptive in character.

The educational component of Duncan's index strikes us as the most problematical in this regard. Although the income component refers to individuals and to total income, rather than earnings from one's main occupation, one can at least think of this component of Duncan's scale as a characteristic of *jobs* rather than of *people*. To the extent this is so, one could go ahead and make random assignments of persons to posts in an experimental situation without destroying the income differentiation of occupations. But education is another matter: it is attached to people not posts. There is no way around this fact and this component of Duncan's SEI scale indubitably means that its use, both conceptually and practically, serves to reduce "occupation" to a contextual variable. There is more than a little intellectual irony in the fact that perhaps the leading critic of contextual analysis (Hauser, 1970a, 1970b, 1974) is also one of the principal proponents of Duncan's SEI scale (Featherman, Jones, and Hauser, 1975; Featherman and Hauser, 1976). . . .

Prestige and Socioeconomic Status as Occupational Indicators

Prestige scores have two distinct advantages and one very definite disadvantage relative to socioeconomic indices in the study of status attainment and related phenomena. The most obvious advantage of prestige scores is that, unlike socioeconomic indices derived from the characteristics of an occupation's incumbents, they do not reduce occupation to a contextual variable. Because prestige scores

are *operationally* independent of the characteristics of an occupation's incumbents, one can logically conceive of an experiment in which subjects are randomly allocated to occupations differing in the prestige they are accorded by the general public. Of course, it may well be that the individual characteristics of an occupation's incumbents are a source of its prestige, in which case one consequence of the experiment might be to reduce the between-occupational variance in prestige scores, since occupations would no longer differ in the characteristics of their incumbents. If, however, the socioeconomic characteristics of an occupation's incumbents are merely *correlates* rather than *causes* of an occupation's prestige rating, then no change would be observed in their prestige ratings, a situation which would be obtained if prestige ratings are derived from, say, the desirability of the work performed in an occupation, the authority built into occupational positions regardless of their incumbents, and features of the typical work setting.

Another advantage of using prestige scores in the study of occupational stratification flows from its status as a well-defined analytical concept in stratification theory. Although Featherman and Hauser (1976, p. 404) conclude "that occupational prestige scores represent a congeries of salient dimensions or occupational characteristics," such a definition of occupational prestige has little to do with the concept as it is typically used in stratification theory. We would venture that the appropriate definition of occupational prestige is analytically parallel to Weber's definition of power. In this view, the relative prestige of two occupations may be defined as the expectation that a member of one will give (or receive) deference from a member of the other. The concept of expectation or probability is crucial to this definition, as it also is to Weber's definition of power, for it admits the possibility that some members of an occupation may receive deference from members of another occupation

while others will give deference to members of the same occupation. In keeping with this view of the relative prestige of two occupations, we can think of the overall prestige of an occupation as the expectation that one of its members will receive (or give) deference to a randomly selected member of any other occupation.

The foregoing definition of occupational prestige implies that it has both a formal and informal component—a part that is built into an occupation by virtue of its formal authority relations with other occupations and a part that devolves upon an occupation by virtue of the performance of its members in situations which are not organized by authority relations. Prestige, in this view, is not identical with power, but it does represent a significant resource—viz., command over the respect of others—which can be mobilized in the effort to secure desired outcomes in the face of competing alternatives. Whether or not occupational prestige conceived in this way is, in fact, what prestige scales measure is, of course, another question. One advantage of using occupational prestige scales, however, rests precisely on one's ability to raise this question intelligibly. Because one analytically knows what prestige is, one can query whether one has measured it satisfactorily. A parallel question cannot be posed of socioeconomic scales, since as best we can tell the socioeconomic status of an occupation is whatever is measured by a socioeconomic scale of occupations.

But whatever analytical advantage prestige scales of occupational status may have is in large measure undercut by their performance, relative to socioeconomic scales, in empirical research. As we have already noted, whatever it is that socioeconomic scales of occupational status measure more nearly governs the process of intergenerational occupational mobility and the entire process of status attainment than do the occupational differences reflected in prestige

scales. This is one very sound reason for preferring the former to the latter, even if one can be less than analytically clear about what it is that socioeconomic scales measure. It is, of course, possible that prestige scales perform poorly because they are inferior measures of the underlying analytical concept. We think a case to that effect could be sketched out, but space does not permit us to do so here.

NOTES

1. This is because Duncan's SEI scale is constructed from the regression equation,

$$\hat{P} = 0.55(E) + 0.59(I) - 6.0,$$

where E is the age-standardized percentage of the male experienced civilian labor force with 4 years of high school or more, and I is the age-standardized percentage of males who had incomes in 1949 of \$3500 or more.—ED.

REFERENCES

- Blau, P. M. (1957), "Formal organization: Dimensions of analysis," *American Journal of Sociology* 63, 58–69.
- Blau, P. M. (1960), "Structural effects," *American Sociological Review* 25, 178–193.
- Blau, P. M., and Duncan, O. D. (1967), *The American Occupational Structure*, Wiley, New York.
- Cain, G. G. (1974), "Review of *Socioeconomic Background and Achievement* by O. D. Duncan, D. L. Featherman, and B. Duncan," *American Journal of Sociology* 79, 1497–1509.
- Davis, J. A., Spaeth, J. L., and Huson, C. (1961), "A technique for analyzing the effects of group composition," *American Sociological Review* 26, 215–226.
- Duncan, O. D. (1961a), "A socioeconomic index for all occupations," in *Occupations and Social Status* (A. J. Reiss, Jr. et al., Eds.), pp. 109–138, The Free Press of Glencoe, New York.
- Duncan, O. D. (1961b), "Properties and characteristics of the socioeconomic index," in *Occupations and Social Status* (A. J. Reiss, Jr. et al., Eds.), pp. 139–161, The Free Press of Glencoe, New York.
- Duncan, O. D., Featherman, D. L., and Duncan, B. (1972), *Socioeconomic Background and Achievement*, Seminar, New York.

- Farkas, G. (1974), "Specification, residuals, and contextual effects," *Sociological Methods and Research* 2, 333–363.
- Featherman, D. L., and Hauser, R. M. (1976), "Prestige or socioeconomic scales in the study of occupational achievement?" *Sociological Methods and Research* 4, 403–422.
- Featherman, D. L., Jones, F. L., and Hauser, R. M. (1975), "Assumptions of social mobility research in the U.S.: The case of occupational status," *Social Science Research* 4, 329–360.
- Hauser, R. M. (1970a), "Context and consex: A cautionary tale," *American Journal of Sociology* 75, 645–664.
- Hauser, R. M. (1970b), "Hauser replies," *American Journal of Sociology* 76, 517–520.
- Hauser, R. M. (1974), "Contextual analysis revisited," *Sociological Methods and Research* 2, 365–375.
- Hauser, R. M., and Featherman, D. L. (1977), *The Process of Stratification: Trends and Analyses*, Academic Press, New York.
- Hodge, R. W. (1970), "Social integration, psychological well-being and their socioeconomic correlates," *Sociological Inquiry* 40, 182–206.
- Siegel, P. M. (1971), *Prestige in the American Occupational Structure*, unpublished Ph.D. dissertation, University of Chicago Library, Chicago.
- Stevens, G., and Featherman, D. L. (1981), "A revised socioeconomic index of occupational status," *Social Science Research* 10, 364–395.
- Tannenbaum, A. S., and Bachman, J. G. (1964), "Structural versus individual effects," *American Journal of Sociology* 69, 585–595.
- Treiman, D. J. (1977), *Occupational Prestige in Comparative Perspective*, Academic Press, New York.

24 ■ Robert M. Hauser and John Robert Warren

Socioeconomic Indexes for Occupations

A Review, Update, and Critique

There are several reasons to focus more attention on the collection, scaling, and analysis of occupational data than has recently been the case. First, job-holding is the most important social role held by most adults outside their family or household. When we meet someone new, often our first question is, "What do you do?" and that is a very good question. Job-holding defines how we spend much of our time, and it provides strong clues about the activities and circumstances in which that time is spent. Second, job-holding tells us about the technical and social skills that we bring to the labor market, and for most people job-holding delimits

current and future economic prospects. Thus, even for persons who are not attached to the labor market, past jobs or the jobs held by other members of the same family or household provide information about economic and social standing. Third, as market labor has become nearly universal among adult women as well as men, it is increasingly possible to characterize individuals in terms of their own current or past jobs. Fourth, once we have a good job description, it is possible to map jobs into many classifications, scales, and measures. Fifth, measurement of jobs and occupations does not entail the same problems of refusal, recall, reliability,

Robert M. Hauser and John Robert Warren. This chapter draws heavily on material in a previous publication: Robert M. Hauser and John Robert Warren, "Socioeconomic Indexes of Occupational Status: A Review, Update, and Critique," in *Sociological Methodology* (vol. 27), 1997, pp. 177–298, edited by Adrian Raftery. Used by permission of the American Sociological Association.

and stability as occur in the measurement of income or wealth. Job descriptions—contemporary or retrospective, from job-holders or from their family members—are imperfect, but the reliability and validity of carefully collected occupational data are high enough to support sustained analysis (Hauser, Sewell, and Warren 1994). Thus, even if we are limited to retrospective questions, we can confidently trace occupational trajectories across the adult years. The same cannot be said of earnings trajectories, let alone other components of personal or household income or wealth.

Conceptual Issues

It is important to distinguish between jobs and occupations. A job is a specific and sometimes unique bundle of activities carried out by a person in the expectation of economic remuneration. An occupation is an abstract category used to group and classify similar jobs. Such abstractions are often heterogeneous and idiosyncratic in construction, but they usually involve determinations of similarity in typical activities, in the sites where work is performed, in the form of job tenure, in the skill requirements of the job, or in the product or service that results from the job. There are multiple systems for the classification of jobs and complex interdependencies between occupational and industrial classifications. Most social scientific uses of occupational data are based either on (a) the classification systems of the U.S. Bureau of the Census, which are revised each decade at the time of the census, or (b) the *Dictionary of Occupational Titles*, which is produced by the Employment and Training Administration of the U.S. Department of Labor.

Some measures of social class reflect job or personal characteristics, whereas others depend on occupation. For example, consider two widely used conceptions of “social class.” Wright’s (1985, 88) class typology combines

concepts of ownership, authority, and expertise. It requires information about a person’s educational attainment as well as ownership, authority, supervision, and occupational classification. On the other hand, Erikson and Goldthorpe’s (1992, 38–39) “class schema” is ultimately a grouping of occupational categories based upon Goldthorpe and Hope’s (1974) study of occupational prestige in Great Britain (Goldthorpe 1980). Each author sees his scheme as a theoretically refined basis for identifying the membership of real and discrete social classes. In Wright’s neo-Marxian classification, the aim is to identify modes of labor exploitation in the relations of production. In Erikson and Goldthorpe’s neo-Weberian classification, the class categories are designed to identify distinct combinations of occupational function and employment status.

In our opinion, differences between the two class schemes and between them and our occupational status measures lie more in the proximity of constituent variables to jobs and persons than in other theoretical or conceptual distinctions that have been debated by their authors. Other things being equal, we should expect a classification based partly upon personal and job characteristics to be more direct and powerful in its influence than a classification based on occupational characteristics would be alone. Rather than relying on a predetermined combination of occupation and other social or economic characteristics, we suggest that investigators should use data on individual education and income, and on other job characteristics, as well as on occupational standing (see Jencks, Perman, and Rainwater 1988).

In working with measures of occupational social standing, we emphasize the social and economic grading of the occupational structure, rather than a priori constructions of distinct social classes. People are linked to jobs, not only through job-holding, but also through their relationships with others who hold or have held jobs. Jobs can be mapped

into standard occupational classifications, and the categories of those classifications may be linked to occupational characteristics. By working back through this series of linkages, we can describe people in terms of occupational characteristics. Such characteristics will be valid as descriptions of jobs only to the degree that occupations are homogeneous and the intervening maps and linkages are sound. In our view, the remarkable thing about this way of measuring social and economic characteristics is not that it is error prone, which would seem obvious, but that it has such high reliability and validity. That it does so is a social fact, which rests both on skill and care in classification and coding but also on strong uniformities in social structure.

Occupational Prestige and Socioeconomic Indexes of Occupational Status

What are the relevant status characteristics of occupations? Many discussions of occupations in the stratification system begin with the concept of occupational prestige, the general level of social standing enjoyed by the incumbents of an occupation. In the United States there have been three major national surveys of occupational prestige, the most recent of which was carried out in conjunction with the 1989 General Social Survey (GSS) of the National Opinion Research Center (NORC; Nakao and Treas 1994). The main problem with all of these occupational prestige ratings is that they lack criterion validity. Prestige is not as highly correlated with other variables as are other measures of occupational social standing, specifically, measures of the socioeconomic status of occupations, as indicated by the average educational attainment and income of occupational incumbents.

Duncan (1961) created the first socioeconomic index (SEI) of occupational status. For forty-five census occupation lines, he ran the linear regression of the percentage of

“good” or “excellent” ratings on measures of both occupational education and occupational income (see Duncan 1961 for details). This regression yielded roughly equal weights for the two regressors, a result that motivated some sociologists to characterize socioeconomic scales as a hybrid of “social status” (as indexed by occupational education) and economic status (see Hodge 1981).

The Duncan SEI has been updated or elaborated in several ways, and researchers should be cautious in using the updates because of their potential lack of comparability. Most recently, as part of their work with prestige scores obtained in the 1989 GSS, Nakao and Treas (1994) created socioeconomic scores for 1980-basis census occupational lines by regressing their prestige ratings on the characteristics of male and female occupational incumbents in the 1980 census. The obvious next step is to create another set of socioeconomic scores, using the 1989 prestige scores as a criterion, but based upon characteristics of the work force in the 1990 census.

In an earlier paper (see Hauser and Warren 1997), we made a special extract of occupational education and earnings from the 1990 census 5 percent public use sample. Throughout our analyses, we used the same definition of occupational education as Nakao and Treas (1994), namely, the percentage of people in an occupation who had completed one or more years of college. After experimenting with alternative treatments of earnings and income, we constructed the new socioeconomic indexes using occupational wage rates, whereas Duncan (1961) used the percentage of occupational incumbents who had reported incomes of \$3500 or more. Our prestige criterion was the percentage of prestige ratings above a fixed threshold. However, for statistical reasons, we used a logistic transformation of the prestige criterion and of the educational level and wage rate of each occupation (see Hauser and Warren 1997, 203–17).

We then constructed socioeconomic indexes for the total work force and, separately, for men and for women. Our purpose in creating gender-specific indexes was to compare the behavior of occupational characteristics between men and women, especially in relation to occupational prestige. We do not recommend routine use of the gender-specific indexes in research. Although the indexes for all workers, men, and women have roughly the same range and are in the same metric, their statistical properties differ. Findings based on the total, male, and female indexes are not strictly comparable (Warren, Sheridan, and Hauser 1998), and, where researchers choose to use a composite socioeconomic index, we recommend the index based on the characteristics of all workers.

Chastened and instructed by the example of Fox's (1991) and Friendly's (1991) re-analyses of Duncan's data, we also paid a good deal of attention to issues of fit and functional form. In the final set of regression analyses, we used several types of residual plots to identify influential outliers. Based on these findings, we deleted several occupations from the regression analyses used to estimate weights for the socioeconomic scores. Several of the largest and most influential exceptions to typical relationships among occupational education, wage rates, and prestige occur in common and visible jobs: business owners, farmers, clergy, secretaries, teachers, waiters and waitresses, janitors, and truck drivers. This finding reminds us that occupational prestige is by no means the same as occupational socioeconomic status, and we should respect both the theoretical and empirical distinctions between them.

Using estimates from our preferred models, we computed total-based, male-based, and female-based SEI scores for all occupations. The combined three sets of scores were transformed to range between 0 and 100. Our 1990-basis and 1980-basis total (TSEI), male (MSEI), and female (FSEI) scores for

all occupation lines, the socioeconomic components of those scores, and the 1989 Nakao-Treas prestige scores and ratings are available elsewhere (Hauser and Warren 1997; <http://www.ssc.wisc.edu/cde/cdewp/1996papers.htm>).

Structural Models of the Socioeconomic Index

Subsequent analyses—presented in detail in Hauser and Warren (1997)—reveal three potential weaknesses in composite socioeconomic indexes of occupational standing. First, gender differences appear both in the relationships between occupational socioeconomic standing and prestige and in the socioeconomic characteristics of occupational incumbents. Second, occupational wage rates appear to be far less highly correlated, both within and across generations, than occupational education. Third, the latter finding led us to wonder whether the use of prestige-validated socioeconomic indexes may overestimate the importance of the economic standing of occupations in the stratification process. Thus, we developed structural equation models in which the construction of socioeconomic indexes was embedded in the stratification process. These models were estimated using data for men and women in the 1994 GSS.

For example, one of our models considered relationships among father's occupational status, the status of a man or woman's first occupation, and the status of his or her current or last occupation. In this rudimentary model, we specified that the status of first occupation depends on that of father's occupation, whereas the status of current or last occupation depends upon father's occupation and first occupation. To be sure, this is scarcely a complete model of the stratification process, but it is sufficient to generate new estimates of the weights of the socioeconomic index components. At each of the three stages of the model—father's occupation, first occupation,

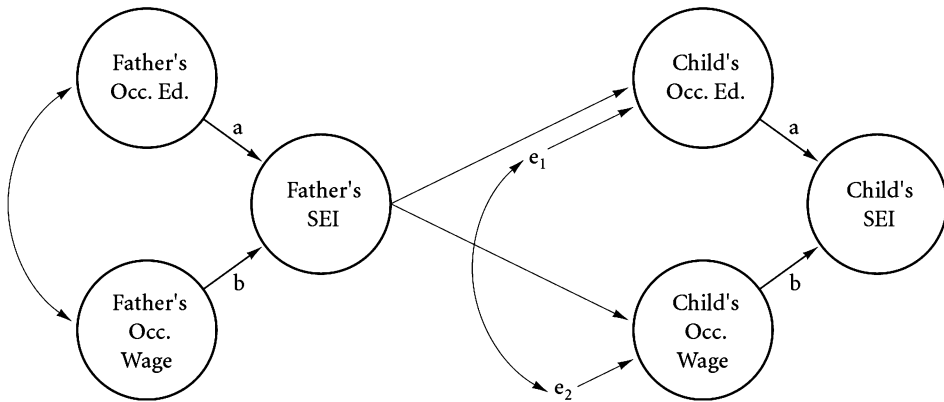


Figure 1. Illustrative Model of Intergenerational Stratification in Occupational Socioeconomic Status

and current or last occupation—we assume that an SEI composite is completely determined by measures of occupational education and occupational wages (Hauser and Warren 1997, 236). Figure 1 illustrates part of this model with measures of father's and child's occupational education and occupational wage. The effects of father on child identify the weights of occupational education, a , and occupational wage, b , in the SEI. To our surprise, we found that the weight of the wage rate is negligible, that is, $b = 0$. Thus, in the GSS data the process of occupational stratification is best described by relationships among occupation-based measures of educational attainment, not the combination of occupation-based measures of educational attainment and wage rates. Our findings from the socioeconomic model point to occupational differentiation by education as a central feature of the stratification process.

Discussion

We are thus led to question the value of traditional socioeconomic indexes of occupational standing, including those that we constructed. If the 1994 GSS data are a reliable guide, we would do better—in studies of the stratification process—to index occupations

by their educational level alone than by any of the usual, weighted combinations of educational level and earnings. However, given the modest sensitivity of occupational status correlations to differences in model specification, we would not suggest any wholesale effort to reevaluate previous findings about levels, trends, and differentials in occupational stratification. It would be sufficient, we think, to suggest that previously estimated levels of correlation are slightly too low. Finally, we would caution that our findings about the relative importance of occupational education and occupational wage rates are specific to models of the stratification process. Just as the relative weights of occupational education and wage rates differ between prestige and socioeconomic outcomes, so they may also differ across other outcomes (e.g., health, wellbeing, social participation, or political choice). If there is any general conclusion to be drawn from our analysis, it is that we ought to move toward a more specific and disaggregated appraisal of the effects of occupational characteristics on social, psychological, economic, political, and health outcomes. Although composite measures of occupational status may have heuristic uses, the global concept of occupational status is scientifically obsolete.

REFERENCES

- Duncan, Otis Dudley. 1961. "A Socioeconomic Index for All Occupations." Pages 109–138 in *Occupations and Social Status*, edited by Albert J. Reiss, Jr. New York: Free Press.
- Erikson, Robert, and John H. Goldthorpe. 1992. *The Constant Flux: A Study of Class Mobility in Industrial Societies*. Oxford: Clarendon Press.
- Fox, John. 1991. *Regression Diagnostics*, edited by Michael S. Lewis-Beck. Quantitative Applications in the Social Sciences, Vol. 07–079. Newbury Park, CA: Sage.
- Friendly, Michael. 1991. *SAS System for Statistical Graphics*. SAS Series in Statistical Applications. Cary, NC: SAS Institute.
- Goldthorpe, John H. 1980. *Social Mobility and Class Structure in Modern Britain*. Oxford: Clarendon Press.
- Goldthorpe, John, and Keith Hope. 1974. *The Social Grading of Occupations: A New Approach and Scale*. Oxford: Clarendon Press.
- Hauser, Robert M., and John Robert Warren. 1997. "Socioeconomic Indexes of Occupational Status: A Review, Update, and Critique." Pages 177–298 in *Sociological Methodology, 1997*, edited by Adrian Raftery. Cambridge: Blackwell Publishers.
- Hauser, Robert M., William H. Sewell, and John Robert Warren. 1994. "Education, Occupation, and Earnings in the Long Run: Men and Women from Adolescence to Midlife." Paper presented at the 1994 Meeting of the American Sociological Association. University of Wisconsin-Madison.
- Hodge, Robert W. 1981. "The Measurement of Occupational Status." *Social Science Research* 10:396–415.
- Jencks, Christopher, Lauri Perman, and Lee Rainwater. 1988. "What Is a Good Job? A New Measure of Labor-Market Success." *American Journal of Sociology* 93:1322–57.
- Nakao, Keiko, and Judith Treas. 1994. "Updating Occupational Prestige and Socioeconomic Scores: How the New Measures Measure Up." Pages 1–72 in *Sociological Methodology, 1994*, edited by Peter Marsden. Washington, D.C.: American Sociological Association.
- Warren, John Robert, Jennifer T. Sheridan, and Robert M. Hauser. 1998. "Choosing a Measure of Occupational Standing: How Useful are Composite Measures in Analyses of Gender Inequality in Occupational Attainment?" *Sociological Methods and Research* 27:3–76.
- Wright, Erik Olin. 1985. *Classes*. London: Verso.

■ THE NEW GRADATIONALISM? ■

25 ■ Aage B. Sørensen

Foundations of a Rent-Based Class Analysis

There is an enormous literature on the concept of class that consists mostly of debates about which properties should be included in the concept.

The theoretically most ambitious concept, the concept of class as exploitation originating with Marx, proposes a mechanism of how antagonistic interests emerge and therefore how class conflict is generated. However, the theory rests on a labor theory of value that has been abandoned by economic theory. The various attempts to resurrect the concept by invoking authority are unsatisfactory because it is not clear that authority is a source of exploitation and antagonistic interests.

There is another solution. This is to maintain Marx's insistence on property rights as the source of exploitation, but to not see all wealth as a source of exploitation. I propose instead to restrict exploitation to inequality generated by ownership or possession of *rent-producing assets*. Rent-producing assets or resources create inequalities where the advantage to the owner is obtained at the expense of nonowners. These nonowners would be better off if the rent-producing asset was redistributed or eliminated. A concept of *class as exploitation* based on the concept of rent is consistent with modern economic theory and

therefore avoids the problems of the labor theory of value.

Class and Wealth

Marx thought that classes were based on rights to the payments on wealth, and Weber thought property to be very important for the emergence of economic classes. Exploitation is a question of economic advantage obtained at the expense of someone else. The right to the returns on wealth is indeed essential for the distribution of these returns, as I will show.

Rights to returns may reflect legal ownership. However, rights to the advantage provided by assets or resources need not be legal rights to be effective. Following Barzel (1997), economic property rights are properly seen as reflecting an individual's ability to consume a good or asset directly or consume through exchange, that is, to control the use of a good or an asset. Such economic rights may be enforceable by law and are then stronger, but they need not be supported by the state to be effective. Property rights are not absolute and constant, and they can be changed through individual or collective action to protect and enhance the rights. Such action incurs transaction costs that are the costs of transfer, capture, and

protection of rights. Illustrations will be given below. When transaction costs are positive, rights are not perfectly delineated, and the transfer or protection of rights will be impeded or made impossible. Positive transaction costs may appear for a variety of reasons. Barzel (1997) emphasizes that some of the attributes of assets may be costly to measure and not fully known to actual or potential owners. These attributes are subject to capture by others who then obtain rights to the benefits from these attributes. Transfer of rights allowing an actor to realize the full value of his assets may be costly because mobility is costly or prevented by force. Collective action needed to rearrange property rights that create a monopoly is costly.

For example, in the modern corporation, stockholders do not own all of the assets of the organization but share it with other parties inside and outside the organization that have rights to gains from various attributes of the assets. Managers obtain some gains because stockholders cannot fully control their use of assets because of lack of information. Other employees may obtain advantages (to be discussed below), for example, by retaining control over their effort. That ownership is divided does not mean that the concept of property as the basis for exploitation should be abandoned, as Dahrendorf (1959) proposes. For example, the absence of individual legal property rights to productive assets in socialist society does not mean that individuals do not gain from controlling the use of an asset in such a society. Only their property rights are more restricted in these societies, and it may be difficult to identify those who obtain the gains (Barzel 1997).

The broader concept of property rights proposed by Barzel implies that individuals—even slaves—usually have some property rights to assets under some circumstances (Barzel 1997, p. 105). This means that all individuals will have some wealth, even if it consists only of their ability to execute a task that can be exchanged for a wage.

A simple formalization may be helpful. Denote with v_j the value of resource or asset j , where value is given by the returns to j over the lifetime of the asset. These returns are usually monetary, but could also be social or psychological. Further, let c_{ij} be the right of actor i to asset j or the control i exercises over j . Then the total wealth possessed by i will be:

$$w_i = \sum c_{ij} w_j$$

where w_i is the wealth of actor i .

Individuals maximize their wealth by maximizing the return on their assets, employing them in the production of goods and services. This usually means that they will need the use of other assets controlled by other actors. Therefore, they need to transact with these other actors. A farmer needs land to maximize the return on the efforts and skills he may devote to farming; a worker needs an employer, machines, and raw materials, to realize the value of her main asset, that is, her labor power. These assets will often be controlled or owned by other actors—the landlord owns the land, the capitalist the machines. These assets can be bought by the actor needing them, or they can be rented. Rental here means transfer of use rights to the asset.¹ The laborer can rent out her labor power to the capitalist in return for a wage, or the laborer can rent the capitalist in return for a profit to him. Such rentals are especially important with durable assets or resources, and even when assets are bought and owned, good accounting practice suggests calculating the payment of a rental to the owner.

The total wealth controlled by actors defines their class situation with respect to *class* as *life conditions*. The assets controlled will determine their incomes and the variability of their incomes. Workers will obtain wages as a result of their effort and skills, and their particular employment opportunities will be important for the variation in their earnings. Assets will be relevant for the respect and

prestige received from the community when knowledge of these assets permits a collective evaluation of the standing of actors. The assets controlled will shape opportunities for transactions with other actors and therefore preferences or economic interests in the meaning of interests suggested by Weber. By shaping welfare and well-being, as well as economic opportunities and the investments that maximize these opportunities, the total wealth and its composition create the behavioral dispositions that are accountable for the inoculation and socialization mechanisms associated with class as life conditions, which I will amplify further below.

When individuals need to transact with other actors to get access to assets they need to realize a return on their wealth, the actors may be able to control the supply of the needed asset. Costs of mobility or other costs may prevent access to alternative suppliers, the supply may be inherently limited by nature or the supplier of the needed access may have created a monopoly. This may allow actors controlling the needed asset to charge for use of the asset that is greater than the payment needed to cover the costs of the asset. For example, the owner of a mine in an isolated location may gain an advantage from lower wages because workers are not able to find alternative employment. Workers are thus prevented from realizing the returns on their labor that they could have obtained elsewhere, and the mine owner has lower costs of production and therefore greater gains from production. The advantage thus gained from effectively being able to control the supply of assets is an economic rent.

Rents may also reflect lack of full information. Executives of organizations may obtain benefits far in excess of what is needed to secure their employment because they are able to control cash flows that stockholders are not able to monitor. Or the supply of the asset will be limited because its availability depends on the presence of other specific assets. In general, rents are advantages that pre-

vent other actors from realizing the full return on their assets. Rents are crucial for the emergence of *exploitation classes* because those who benefit from rents have an interest in protecting their rights to the rent-producing assets, while those who are prevented from realizing the full return on their assets have an interest in eliminating the rents. Rents thus may create antagonistic interests and conflict.

To see how rents emerge, it is useful to consider the transactions involved in maximizing returns on productive assets more closely. The prices for the rentals of assets needed to maximize the returns on actor j 's resources are costs to j for those assets he does not own and returns to j for those assets he does own. This means his wealth is crucially dependent on the prices of assets relevant for him. These prices depend in the usual manner on supply and demand in the market. If the supply of a certain asset, for which actor j pays a rental, increases, the price will fall and actor j 's wealth increases since he has lower costs. Demand in a similar manner will influence the value of assets. This is the normal story.

Suppose now that actor k controls the supply of something j needs to employ her or his assets. Actor k may own land actor j needs access to in order for j to obtain a return on his labor and farming skills. They will negotiate a price for using the land and this price is a cost to j reducing the benefit he receives from his labor and skills. When negotiating the price, j and k will compare what other farmers pay for land. In the long run, competition will ensure that a price emerges that will ensure j a sufficient revenue to keep him alive and able to work the land and compensate the landlord for whatever costs he covers, for example, for fencing. Of course, j may try to buy the land instead of renting it, but for the eventual outcome this does not matter: the rental to the landowner is replaced by a rental of capital to finance the purchase, that is, of interest.

For the competitive equilibrium to occur it is important that the supply of land can vary in response to prices. It is also required that renters of land are mobile and thus able to take advantage of the rental offers provided by landowners. If these conditions are not met, if land is in fixed and limited supply or if renters of land—farmers or peasants—are prevented from being mobile by force or law, the owners of land can charge a rental price that is larger than the hypothetical competitive price that just covers their costs related to the ownership of land. The difference between the actual rental price and the competitive price is what is called an *economic rent*.

Rents are payments to assets that exceed the competitive price or the price sufficient to cover costs and therefore exceeding what is sufficient to bring about the employment of the asset. Thus, a rent on asset i can be defined as:

$$r_i = v_i^A - v_i^C$$

where v_i^A is the actual value of i and v_i^C is the value that would have emerged under competition and equal to the costs of making the asset available.² These values are given by the stream of income generated by the asset over time.

The existence of rents depends on the ability of the owner of the asset to control the supply. I have already alluded to the classic example: the tenancy contract associated with feudalism. Part of the benefit from the land goes to payment for the labor of the peasant, and another part of the benefit goes to payment for capital expenditures on the land by the landlord. The rent benefit obtained from a tenancy arrangement is the remainder, that is, the payment not needed to employ the peasant and keep the land fertile. It is an advantage going to the landlord because of his rights to the returns on the asset that he controls. But the rent benefit forces a disadvantage on the peasant, since he does not realize the full value of his labor and skill.

The association of rents with land is not required. Rent will emerge on all productive assets that are in fixed supply and that actors need to maximize their wealth; or rent may be present as a result of transaction costs involved in getting access to needed assets. Alfred Marshall (1949 [1920]) devoted much attention to the concept of rent and generalized its applicability to benefits received from any productive resource or asset. He showed that rents also may appear as payments for the use of capital and labor in restricted supply; as payments for the use of unique combinations of capitals and labors, such as those created by certain technologies; and as payments for unusual and rare individual abilities that cannot be developed by training alone (musical talents, artistic creativity, athletic ability, etc.). Rents may be created in employment relationships when workers control their effort in an attempt to increase the advantage obtained by the wage because cost of monitoring prevents the employer from adjusting wages to effort. In general, the salient property is that a component of the payment obtained from the asset, or its return, is in excess of what is needed to bring out the supply of the asset.

The association of land with rent is not only an accident of history, it also reflects that tenancy arrangements usually are long-term and rents therefore are long-term. In a competitive economy, rents may emerge in industrial production as a result of an innovation or an import restriction. However, when others discover that there is an excess profit or rent available from owning a particular resource these others increase the supply of the resource if they can. This reduces the excess profit and eventually makes it disappear. Marshall (1949 [1920]) calls such temporary rents *quasi rents*. These temporary rents are the typical rents in capitalist production and will become important in our discussion below.

For our analysis it is extremely important that rents are advantages to the owner of assets

that are not needed to bring about the use of these assets. If the competitive payment is enough to make the landlord willing to let the farmer use his land, then any excess is in a basic sense unnecessary. It is an advantage costing nothing. The farmer has a clear interest in reducing, and if possible eliminating, the rent. The landowner has an equally clear and opposite interest in preserving the advantage provided by the rent. Rents therefore create antagonistic interests. Certain rents are especially important for social structure and social change. These are enduring rents that, resulting from enduring property rights to rent producing assets, cause significant advantages and disadvantages. They are at the basis for class formation, as owners of such assets will protect their property rights to these assets, and nonowners will seek to eliminate these rights.

In summary, the individual's total wealth, as defined by her control of assets, will determine her life conditions and thus her class location in terms of class as life conditions. It will be argued below that the consequences of these conditions are not only dependent on the total wealth, but also on the overtime variability in the returns on that wealth (which define the variation in the value of the wealth). Part of the total wealth may generate benefits obtained at the expense of someone else, who would be better off with a different distribution of control or property rights to the various attributes of the assets. This *rent*-generating part defines class as exploitation. Below, I further develop these ideas, treating briefly first the idea of class as life conditions as total personal wealth and then treating in more detail the exploitation class concepts based on rights to returns from rent-producing wealth.

Wealth and Class As Life Conditions

As noted above, there is an abundance of research that shows that class as life conditions indeed is a powerful determinant of all kinds of outcomes. There is much less understand-

ing of how these outcomes come about. We have, of course, a rich literature on socialization that demonstrates that class is associated with important socialization differences, we know about important value differences among different classes, and we also know about a host of lifestyle differences associated with different classes. However, this only moves the question one level back. What is it about the living conditions of different classes that accounts for these differences?

I propose that the answer is lifetime wealth and the expected variation in returns on that wealth for incumbents of different classes. There is abundant evidence that social class accounts for more outcomes the more homogeneous class categories are with respect to a variety of resources, or their wealth. It is important to consider not the cross-sectional distribution of income, but the long-term wealth profile that determines what economists call *permanent income and consumption patterns*. A person who obtains a higher education will orient her lifestyle not to the level of income in her youth, but to the long-term expected living conditions corresponding to the wealth associated with her human capital.

Further, the variation in the returns on the wealth is important, particularly for the socialization patterns that emerge in different classes. An older literature found strong differences between social classes in what was called "the ability to defer gratification" (see, e.g., Schneider and Lysgaard 1953). This literature was largely dismissed in the seventies because it was seen as reflecting an attempt to "blame the victim" (see Ryan 1971). More recently, psychologists and economists have suggested a different formulation of the same phenomena (see Ainslie 1992). People discount future rewards, often at very high rates. In particular, there are strong differences among social classes or different socioeconomic levels in time orientation, with persons at low socioeconomic levels having a much shorter time horizon than others. Those with high discount rates invest less in

their health and education, and in the health and education of their children.

These differences among classes in time orientation or deferred gratification patterns reflect the level of uncertainty in living conditions or the variability of returns. Such uncertainty is not the fault of the “victim,” but is a rational reaction to the expected high uncertainty of returns. Banks also charge a higher interest rate with uncertain investments, and banks presumably are acting rationally. The impact of uncertainty on people’s investments in themselves, and their children, should be greater, the lower the overall level of resources. Fewer resources give less of a buffer.

A person’s total wealth has two main components. One part is personal, human, and physical wealth that is acquired mostly outside of the labor market in and from families and schools, but some is acquired from on-the-job training. The other component is wealth acquired from employment relations.

The personal part of wealth that exists, independent of the actor’s employment relationships, has several components. The amount of human capital obtained through investments in training and health is particularly important. There may also be skills and abilities that command rents.

Finally, the amount of wealth obviously depends on the endowments of physical capital provided by the family of origin and augmented by the person through entrepreneurship and investment independently of his involvement with the labor market. This variation in endowments creates different incentives for investments in human capital and the like, and these differences explain the much emphasized nonvertical nature of the Goldthorpe class scheme, for example, the life conditions of farmers. This component of personal wealth is obviously of major importance for a full analysis of the class structure.

Individuals also obtain wealth from their employment relationships. They may have access to on-the-job training opportunities

that increase their human capital. A component of the human capital acquired on the job may be specific to the job and the firm and give bargaining power to the worker. Because of specific human capital or collective action, the worker may gain above-market wages, increasing the value of his labor assets, and thus obtain a rent. The employment relationship will in these circumstances be closed in contrast to the open employment relationship characteristic of the competitive market. The resulting increase in the expected duration of employment relationships is crucially important for the variability of returns and therefore for the consequences of differences in wealth, because the shorter the employment relationships, the more variable will be the returns on wealth. The duration of the employment relation also is important for the amount of wealth obtained in the relationship. We should therefore expect persons in stable employment relationships to invest more in themselves and in their children. Professionals having large amounts of human capital and stable employment relationships should invest the most. Therefore, what Erikson and Goldthorpe (1992) call the “service class” should be especially successful securing their children’s future. Though no specific test of the idea proposed here exists, an abundance of research on social mobility and inequality of educational opportunity demonstrates the ability of professionals and others with high levels of human capital and enduring employment relations to secure the success of their offspring.

Wealth, Rents, and Exploitation

The issue for the formulation of a theory of exploitation is to define a process by which some holder of an economic property right obtains an advantage at the expense of persons without these rights. As shown above, wealth transfers made possible by the acquisition of rights to assets generating economic rents satisfy this requirement.

Rents satisfy the requirements of the structural theory of inequality. Rents are created by social relationships of ownership of rent-producing assets (with the obvious exceptions of rents on natural abilities, to be treated later). Advantaged exploitation classes are positions in the social structure that allow individuals to gain control or economic property rights over assets or attributes of assets that generate rents; disadvantaged exploitation classes are defined by the absence of these rights. Changing the property relations that generate rents will change the distribution of wealth and hence the class structure.

The holder of a rent-producing asset has an interest in securing the continued flow of benefits, and those denied the benefits, a clear interest in obtaining the benefit by acquiring it, or by destroying the social organizations that create the rents. When actors act on their interest, they create social organization and processes to protect or destroy rent benefits. These arrangements, well described by neo-Weberians, are processes of closure and usurpation (Parkin 1979) and the processes of moving from awareness of interests, through development of consciousness, to acting in pursuit of these interests (Giddens 1973). For the scenario to unfold, not only membership, but also interests must be enduring.

The distinction between temporary rents and enduring rents is very important for the analysis of class-formation processes. Class formation not only depends on stability of membership in structural locations providing antagonistic interests, as pointed out by Goldthorpe and Giddens. Class formation also depends on the rate of change in advantages or disadvantages provided by rents. This immediately suggests that structural conflict or class conflict should be more prevalent under feudalism than under capitalism, for rents are more permanent under feudalism. No revolution has occurred in an advanced capitalist society.

The importance of the distribution over time, of the advantages provided by rents is often ignored. A cross-sectional inequality

does not necessarily imply a longer-term advantage provided by an enduring rent. For example, according to human capital theory, the higher incomes of the higher educated compensate for higher training costs and do not create a permanent advantage over the lifetime of the person. Thus, skills acquired according to the mechanism proposed by human capital theory do not create rents and therefore not classes. This is generally ignored in the so-called new class theory that sees classes emerge on the basis of skills and education (Gouldner 1979; Konrad and Szélenyi 1979). Education can, of course, create rents, but a measure of educational attainment as used, for example, in Wright's class scheme cannot separate the rent from the human capital component. The role of education in class analysis will be discussed further below.

The types of social organization and processes that emerge around rent-producing assets differ according to which type of asset is being considered. Feudalism can be described as an elaborate organization for the distribution of rent benefits based in land and mercantilism as an extension of the arrangements to cover industrial production. In modern industrial society, there are three main types of rents to be considered, already identified by Marshall (1949 [1920]): (1) monopoly rents, based on monopolization of the supply of an asset, for example when a cable company gains a monopoly from a local government on distributing TV signals; (2) composite rents, formed by unique combinations of productive rents or asset specificity, for example when a worker has acquired skills only employable in a particular job; and (3) rents based on natural abilities and talents, for example the height and ball-catching ability to make a professional basketball team. I will consider some of the main properties of each of these types.

Monopoly Rents

"Artificial" or social constraints on production create monopoly rents. The monopoly

may have emerged “naturally” because of increasing returns to scale creating prohibitive costs of entering production for others, as in the production of automobiles. Often monopolies are created by governments as licenses or patents. Finally, social associations, such as trade unions or industry associations, who agree to regulate the production of something, create monopolies. In all these cases the supply of a product will not be sensitive to price, and rents will appear and persist unless the monopoly is broken.

For the sake of clarity, assume that the monopolist is working with production conditions that generate constant returns to scale so that average costs equal marginal costs. Nothing essential in the present argument depends on this assumption. Under perfect competition, output and price would be q_c and p_c and the price will correspond to the cost of the product. The monopolist is able to charge a price p_m above the price p_c that would prevail with perfect competition. This will cause an increase in revenue per unit produced that is an increase in the income of the producer over and above the amount needed to bring forth the production, that is, a rent. In addition to the creation of the rent and the corresponding increase in inequality, there will be a reduction in the wealth of society, for less is produced at the price p_m . This is the “deadweight loss” caused by monopoly rents and represents a welfare cost to society, which is a social waste of resources.

The increase in revenue to the monopolist is, of course, an advantage others might want. If others therefore successfully enter the market, the resulting competition might eventually erase the monopoly rent, lowering the price to p_c and increasing the quantity to q_c . When this happens, the temporary advantage to the initial producer is a quasi rent.

This scenario, of course, assumes that others can enter production. If there are prohibitive entry costs created by production technologies, governments or trade associations, competition will be about obtaining the mo-

nopoly. Such competition is the typical case of *rent seeking*, that is, zero-sum competition over rent-producing assets. The efforts and other costs involved in trying to acquire the rent-producing property or resource of course reduce the benefit of the monopoly. Indeed, those who wish to acquire the monopoly should be willing to pay the equivalent of benefits to obtain it, so that the rent benefit completely disappears. The costs of rent seeking do not increase the production of society and therefore represent wasted resources (Tullock 1980). This waste is in addition to the waste represented by the deadweight loss.

The nature of the rent seeking depends on whether the monopoly can be traded in the market or not. If it can be traded, the sale may create a large transfer of wealth to those who obtained the monopoly first, and subsequent owners will not realize a rent. For example, it is often argued that rents received by farmers, as agricultural subsidies, produce higher land values and therefore higher interest payments, eliminating the initial advantage. Once established, the rent creating monopoly is difficult to eliminate, even when the monopoly is fully capitalized and the rent has disappeared. Clearly the new owners are strongly interested in receiving the rents they have paid for, even though the advantage to them has disappeared. An example, described by Tullock (1980), is a taxi medallion system, similar to the one in New York City. The medallions are sold, producing large gains to the initial owners, but only normal rates of returns to subsequent owners. Their existence creates a welfare loss to consumers. This loss can be reduced only by removing the restriction on taxi driving, something that is almost impossible to do without forcing losses on the present owners of medallions.

A variety of monopoly rents emerge in the labor market. *Employment rents* emerge when employment and jobs are closed to outsiders by the collective action of unions, by government approved certification of professions, and by other occupational licenses. Unions

create rents when they close shops or ration employment through apprenticeship systems. Unions may also significantly alter the distribution of rents when they obtain egalitarian wage systems that will increase the wages to the least productive with the lowest market wages (for evidence see Freeman and Medoff 1984). Professional associations create rents when they obtain certification limiting employment to the properly certified or when they gain control over the recruitment to the profession through control over educational institutions; medical schools are a good example of this. In general, educational credentials, used as rationing devices for employment or for access to employment-specific education for professions, create monopoly rents to those holding the credentials. Credentials will be further discussed below.

Employment rents create rent seeking as zero-sum competition for positional goods (Hirsch 1976) in what I have called “vacancy competition” (Sørensen 1983). Employment rents are not only monopoly rents. Positions may also be closed without the assistance of outside agencies like unions or professional associations. In internal labor markets, closed positions can be created without collective agreements because of the existence of composite rents created by asset specificity, for example specific skills, to be discussed below. . . .

Composite Rents or Rents on Asset Specificity

When two separate assets or resources are so specific to each other that payment to their joint use exceeds the payment to each resource in separate use, composite rents emerge. Marshall’s prime example of composite rents is the joint advantage to owners and employees of an advantageous market position (Marshall 1949 [1920]). A specific example is the joint rent received when a mill is built on a water stream, to the owner of the mill and the owner of the stream. If there is only one site for the location of the mill, then the rents to the mill owner and to the owner

of the water source cannot be separated: “There is nothing but ‘higgling and bargaining’ to settle how the excess of the value of the two together over that which the site has for other purposes shall go to the owner of the latter” (Marshall 1949 [1920], p. 520).

Composite rents emerge from what in the literature on transaction costs is called asset specificity. They emerge, for example, when workers have obtained specific on-the-job training and therefore are more productive in one firm than in another (Becker 1964). Also, monitoring and agency problems may create composite rents. The composite rent creates a joint advantage that would disappear if the match between the firm and the worker is dissolved, so that employment relations become closed. There are two types of solutions to these problems.

First, the composite rent could be eliminated by organizing production so that the transaction cost problems disappear and employment relations become open. With respect to specific skills, this would imply eliminating the use of such skills. Such deskilling by eliminating the need for specific skills differs from the original deskilling idea made prominent by Braverman (1974), which suggests that capitalism will try to eliminate the need for all skills in the labor market. A general trend toward deskilling has never been established, despite many attempts, nor does it make theoretical sense that employers inevitably stand to gain by reducing the general level of skills required. However, reducing composite rents due to specific on-the-job training would be a plausible strategy.

The second solution is to reduce the importance of composite rents, without destroying closed employment relations to outsiders, by using organizational devices that increase effort. A large organizational literature on internal labor markets may be seen as analyzing organizational solutions to the problem of increasing the firm’s share of composite rents. A prominent solution is the creation of promotion schemes to elicit effort.

Promotion schemes capitalize on the interdependence of effort created by zero-sum competition over the wage and earnings differential provided by promotion ladders that are positional goods.

Promotion ladders create cross-sectional inequality. It is this inequality that is attributed to class by the Wright class scheme as an effect of authority. However, promotion ladders may be designed to provide less than the market wage at the start of the career, legitimized by training and higher than market wages at the end of the career, according to the deferred payment theory (Lazear 1995). This pattern ties the worker to the firm and preserves the composite rent: she receives positive and negative rents depending on her seniority in the system. This does not mean that workers will obtain an overall surplus over their lifetime. As with investments in education, the cross-sectional distribution does not inform about the long-term advantage obtained. If the promotion ladders work as intended, they elicit effort and capture composite rents for the employer, quite contrary to the usual interpretation of the authority effect.

The interpretation of promotion systems as rent-capturing systems also implies that employers have an incentive to default on the positive rents the worker is to receive at the end of the career by dismissing the worker when these positive rents emerge. This will create “reputational” problems for individual firms, but if many firms collude in the practice, the reputational effects are diminished.

Another device to elicit effort and capture a larger share of the composite rent is the incentive wage systems. By paying employees more than their market wage, firms increase effort since workers will be reluctant to shirk out of fear of losing their jobs. Wright (1979) used such an efficiency wage explanation for the income advantage of “semi-autonomous” employers. In his latest class scheme, Wright proposes a similar explanation for the wage advantage of managers,

called a “loyalty rent.” The efficiency wage explanation is used by Krueger and Summers (1987) to account for the persistent wage differentials across industries that cannot be attributed to unmeasured worker characteristics or compensating differentials.

Rents on Natural and Cultural Endowments

Marshall (1949 [1920]) suggests that rents emerge on “free gifts of nature” in the form of genetic endowments that result in the ability to produce something in demand. The rents directly reflect genetic endowment, as when genes are responsible for certain physical attributes facilitating certain tasks; for example, height for basketball players. Or, the rents obtain indirectly when an individual endowment facilitates training for certain skills, as in academic achievement. In the latter case, the endowment need not be genetic. Cultural endowments are important for learning, but hard to learn for those not socialized into a given culture, or who lack the requisite cultural capital (Bourdieu and Passeron 1990 [1970]). Cultural capital thus may be seen as a source of rent similar to genetic endowment.

It may seem paradoxical to include rents on individual endowments as a source of structural inequality. However, these rents have important social consequences that connect to the class analysis literature. In particular, rents on natural and cultural endowments have important consequences for the emergence of credentials.

General ability creates higher productivity and higher wages in many employments. Higher productivity may alternatively be obtained by training. The training costs needed by the less able create a surplus for the able, assuming that equally productive able and untrained workers receive the same wage as the less able and trained workers. With an expansion of demand for credentials—that is, an increasing demand for education—the rents become larger for the more able. They

therefore seek even more education and higher and more expensive credentials. This self-stimulating demand is the main thesis of the credentialism literature (e.g., Collins 1979). The larger rents provide an incentive for institutions of higher education to increase tuition costs. They therefore share in the rents produced by credentials. Those rents then permit the hiring of prestigious faculty to train the easily trained, enhancing the reputation of these institutions and further increasing the rent on the credentials they confer.

Those possessing high credentials wish to secure an advantage to their offspring. This is facilitated by making cultural capital relevant for training. However, the very existence of credentials also is important. The superior ability of one's offspring cannot be secured, but much can be done to secure a valuable credential for the offspring by facilitating access to institutions providing valuable credentials. In the absence of such credentials, less able offspring from high-status backgrounds might have to compete for valued employment with the more able from more humble origins. The monopoly on employment ensured by the credential protects the less able from high-status backgrounds from being outcompeted by the more able from lower-status backgrounds. Credentials thus increase the ability of high-status groups to confer their advantage to their less able offspring and increase the advantage to their more able offspring. There are strong incentives for high-status groups to create credentials and closure as emphasized by the neo-Weberians (Parkin 1979; Murphy 1988).

In general, the differential rents generated by individual endowments imply that increasing equality of educational opportunity through educational expansion should increase the rents on natural and cultural endowments. Such policies therefore should be strongly supported by those who already obtain considerable rents on their endowments, such as professors.

Exploitation Classes and Collective Action in Modern Capitalism

Rent-seeking activities create lobbies to influence the regulatory activities, subsidies and welfare policies of governments. Social movements lobby to improve the welfare of the disadvantaged by granting them rents. Major corporations lobby through campaign contributions designed to obtain the type of policies and regulations that increase the rents to these corporations. Rents divide owners of different productive assets—as when owners of land are in conflict with owners of industrial productive assets about corn tariffs—and they unite workers and capitalists to preserve import regulations and trade barriers that create rents to certain firms and industries.

That class action in modern capitalist society is about rent seeking and the protection of property rights to rent-producing assets clearly creates a different conception to what Marx had in mind when he analyzed capitalism in the late nineteenth century. Marx's conception, for example, does not require monopolies and asset specificities for the creation of advantage. The surplus created by labor will be a universal feature of capitalism, which will derive its nature from the relentless pursuits of ever-falling rates of profit. However, when the labor theory of value is abandoned, it is impossible to sustain the idea that there is a permanent "hidden" form for surplus in capitalist production in the manner conceived of by Karl Marx. The main class actions will be rent seeking, the protection of existing rents, and the destruction of rents.

It is an interesting question whether rent seeking, rent protection, and the destruction of rents might sustain Marx's grand scenario for the development of capitalism. Marx was certainly right about the dynamics of advanced capitalism. The engine of this dynamic is the pursuit of acquiring rent-producing assets through innovation and product development and by creating demand through advertising for profitable

products. The relentless pursuit of advantages that exceed above-market returns—through the reorganization of firms and corporations, sometimes in the form of mergers and acquisitions, sometimes in the form of divestment—is also a pronounced feature of modern capitalism. These processes result in quasi rents that are usually quickly eliminated by competition. Individual capitalists gain and lose, and some obtain great fortunes. Even though their fortunes result from quasi rents, they are not destroyed by the elimination of these quasi rents. The process expands markets and produces globalism. The story is well known and well described.

Enduring rents to individual owners of capital require some type of collective action. The main form of collective action among capitalists is the establishment of cartels. Cartels may, of course, be hindered by government antitrust regulations. They may be effective despite such obstacles by various types of network organization among boards of directors and the like, but the incentive to break an agreement is always present. A more effective strategy for obtaining enduring rents is to obtain help from the state to preserve an advantage: the granting of a license or some other form of protection from the entry of competitors. An army of lobbyists tries to obtain such advantages by informing legislators about the consequences and advantages of their actions. The rationale for state creation of monopoly rents is usually that some public benefit will be obtained by the regulation that otherwise will be lost: competent doctors, safe cars, and the family farm.

Can Marx's scenario for the class structure of advanced capitalism be sustained with the conception of rent-based classes as exploitation? Marx's emiseration prediction is usually taken as a main reason for revising his theory. Clearly the idea of increased absolute poverty of the working class, caused by increased exploitation, has been rejected by the economic growth that has occurred since Marx wrote. For a long period, also a decrease in relative

inequality was observed in most societies. This decrease was replaced by an increase in inequality in the United States and many other advanced societies in the early 1980s, an increase that has continued since then. Nevertheless, it is not possible to sustain the idea that we find an increased polarization and homogenization of the working class. There is, however, substantial recent evidence that shows that capital has become very effective at eliminating the advantages of the working class in terms of rents obtained in the labor market. Eliminating these advantages has contributed to the increase in inequality.

Capital will gain by the destruction of monopoly rents in the labor market and by increasing its share of composite rents or destroying the source of these composite rents. The elimination of rents in the labor market benefits the capitalist when he benefits from the increased efficiency of production. He further benefits when his wealth is dependent on valuation of how efficiently he produces. The stock market provides this valuation. The stock market has, in the period where inequality has increased, very much increased the wealth of stockholders and rewarded rent elimination in the labor market. There are several main ways in which this has been achieved, and all have resulted in increased inequality: (1) eliminating rents created by collective action, in unions, (2) eliminating internal labor markets and composite rents, and (3) lowering the real value of the minimum wage.

Unions create rents in two ways. They may provide significant wage premiums for workers covered by union contracts. Nonunion workers may also obtain benefits when employers try to avoid unions. These benefits tend to accrue to workers who are highly skilled. As shown by Freeman and Medoff (1984), though unions do provide benefits, the rents are quite modest. The main effect of unions is to reduce wage inequality. Unions are especially effective at decreasing the wage spread between more and less productive workers. Unions may create substan-

tial rents to low-skilled or otherwise less productive workers.

A well-known major change in the labor market has been the reduction in union power. This is a reduction both in the number of workers that are union members and in the ability of unions to obtain wage increases and secure bargaining agreements. The reduction in membership has been from about one-third of the nonagricultural labor force to now 16 percent. The influence of the unions on the wage structure is far greater than its membership (Mitchell 1985). However, the decline in membership has also been accompanied by declining union power. The evidence is the increase in the number of concessions, the decline in the number of strikes, and the moderation of union demands (Mitchell 1985).

Closed employment and composite rents are widespread also in industries and firms without union presence, in internal labor markets, and for groups of workers traditionally not unionized, such as many white-collar groups. The composite rents obtained in these settings are eliminated by layoffs. Layoffs without recall reduce job security, but not necessarily employment. However, the loss of job security means also the loss of whatever rents the worker has obtained. With job security a worker can never do worse than his present job. If a better job comes along, he can move to this job, and the timing of this move need not have anything to do with increases in productivity. Therefore a system of closed employment, as in internal labor markets, produces career patterns that represent increases in rents only and not increases in productivity. These career structures are destroyed by downsizing.

Composite rents associated with internal labor markets can also be eliminated by job redesign and other changes in production technology. Or, they can be eliminated by removing asset specificity through outsourcing and subcontracting for labor. There is much talk about such changes, including how they

could encourage the evolution of new types of employment relationships.

The elimination of employment rents through downsizing and job redesign often means that workers are forced to look for new jobs in the labor market without much choice of which job to accept. This should mean that the next job after the downsizing is likely to be a worse job. It also means that the match between the downsized worker's productivity and wage is likely to differ from previous employment. There should be a closer match between actual individual productivity and wage level as a result of the job displacement. There is some evidence that suggests a tighter relationship between wages and productivity in the 1980s than in the 1970s (Levy and Murnane 1992; Mitchell 1985). Holzer (1990) thus finds a better match than Medoff and Abraham (1981), but the two studies are not very comparable. Juhn, Murphy, and Pierce (1993) find the increase in inequality driven by increased returns to unmeasured skills.

Consistent with the idea of a stronger link between wages and personal endowments, we also observe a marked increase in within-occupation inequality. This is true for all occupations, but it is especially true for managers and sales personnel. In fact, for men the overall Gini coefficient rose from .315 to .332 between 1980 and 1989, but in the managerial and sales occupations combined, it increased from .322 to .353. For all other occupations, the Gini increases from .302 to .312 (Ryscavage and Henle 1990, p. 11). As inequality increased, structural locations seemed less relevant for explaining the variation in earnings.

The increase in inequality is very much driven by an increase in wages and earnings of the highest-paid workers and stagnation or decline for others. The stagnation and decline follow from the rent destruction. The increases for the highest-paid result from rent sharing with capital and may be legitimized by arguments that top managers were underpaid in the 1970s and therefore did not have enough of an incentive for doing their very

best; in particular, they may have been more tolerant of rents to other employees (see, e.g., Jensen and Murphy 1990).

Stock market valuation clearly has been important for changing this situation. Another mechanism to increase manager incentive has come about through leveraged buyouts that force managers to squeeze all slack out of the firm to meet debt obligations. Leveraged buyouts also make top managers much wealthier. Finally, the increased competitiveness may have increased the rents on the abilities that boards of directors believe are needed in tough managers.

The declining real value of the minimum wage, until quite recently, also reduces employment rents for those less productive workers paid more than their competitive wage because of the minimum wage. This brings more poverty, for nothing guarantees that a competitive wage moves a worker above the poverty line. The rent destruction in the labor markets, except perhaps for the highest-paid managers, leaves a labor market more flexible and more fluid, for fewer groups have anything to protect. The result is less structure, meaning less positional inequality, but more inequality overall. Thus, while greater homogenization overall may not have resulted from these recent trends, the destruction of rents in the labor market has created a labor market with fewer structural supports for the returns to labor. The idea of a homogenization of the working class can be sustained if it refers to the availability of structural advantages making earnings from work less dependent on individual endowments and more dependent on occupational choice and collective action.

Nothing guarantees that efficient labor markets create good lives. Rents are required in modern society to provide decent standards of living for the poorest part of the population. These rents are provided from the state in the form of income support and other welfare goods. The modern welfare state provides required support, but also creates an arena for

rent seeking by all, including the middle strata with effective interest groups. It is beyond the scope of this discussion to deal with rent protection and rent seeking in the welfare state. Elsewhere (Sørensen 1998), I have provided a treatment of the breakdown of traditional norms around the provision of welfare goods and resulting increased rent seeking.

Conclusion

A sound basis for class concepts should be based on property rights to assets and resources that generate economic benefits. Property rights should be conceived of broadly. They are economic property rights defined as the ability to receive the return on an asset, directly or indirectly through exchange (Barzel 1997). Some of these rights may be supported by the state, and they are then legal rights, but people also obtain advantages from rights that are not legally enforceable. Property rights define an actor's wealth and I suggest that the *class as life conditions* reflects a person's total wealth. Part of this wealth may be in assets that generate returns or payments that are rents. Rent distribution creates *exploitation classes* that may engage in collective action.

Class as life condition is a very useful concept for analyses of how patterns of attitudes, behaviors, and socialization vary by location in social structure. A prominent example is the class concept proposed by Goldthorpe (1987; see also Erikson and Goldthorpe 1992). A major objective for constructing class schemes that account for different living conditions is to identify homogeneous groupings with respect to total wealth, type of wealth, and the variability of wealth over time. Such groupings will differ in the amount and type of investments they make in themselves and their children. We therefore obtain class schemes that include nonvertical dimensions reflecting the type of wealth possessed and its variability over time, as generated, for example, by the stability of employment relationships.

The present proposal overcomes the evident problem associated with Weberian and neo-Weberian class analysis where there is no proposal for why anyone should be upset about their position in society and engage in class formation. Enduring rents identify antagonistic interests. Those who do not own a rent-producing asset suffer a disadvantage as the result of the rent. It is in their interest to eliminate the rent, and in the interest of the rent receiver to protect the advantage. The proposal I present here provides new insights. The concept of quasi rents suggests that monopoly rents often are transitory and the associated interests therefore not enduring. Thus, not only will stability of membership in structural locations and closure be important for class formation, but variations over time in rent advantages are important for predicting class formation. Rents provide a new interpretation of credentialism as a device to preserve and transmit advantages from one generation to the next with uncertainty about the ability of offspring.

The rent-based concept of class as exploitation provides an explanation for the recent increase in earnings inequality and for the practice of downsizing to destroy rents in the labor market.

The main class action will be by actors to seek rents, to protect rent privileges, and to destroy rents in structural locations, such as internal labor markets. The argument here implies that it is to the advantage of the capitalist class to produce a labor market conforming to the assumption of neoclassical economics, and I have tried to show that capitalism in the last decades has been successful in eliminating rents to labor. Eliminating rents in the labor market creates more efficient labor markets—that is, labor markets with less structure and more fluidity. A rents-free labor market will be one where simple class schemes are increasingly less applicable. The destruction of rents also creates more inequality within the labor market and produces more wealth that accrues to some of those owning means of

production—for example, capitalists—whether they are old families, new entrepreneurs, pension funds or graduate students with mutual funds. The resulting society conforms to Marx's predictions about the nature of advanced capitalism: “The bourgeoisie, whenever it has got the upper hand, has put an end to all feudal, patriarchal, idyllic relations. . . . It has resolved personal worth into exchange value, and in place of the numberless infeasible chartered freedoms, has set up a single, unconscionable freedom—Free Trade” (Marx 1959a [1848], p. 323).

Thus, the main prediction about the development of capitalism from rent-based class theory is that rents will disappear from structural locations in the labor market. This will result in a structureless society, without the nooks and crannies of social structure we have come to expect because feudalism is slow to disappear. The result is the transfer of wealth to those who have rights to rent-producing assets, even though these assets usually are quasi rents, for the wealth created by quasi rents is not destroyed when the rent is destroyed. As a result, we see increasing wealth inequality (Wolff 1995).

If Marx's grand scenario for advanced capitalism is interpreted as having to do with the distribution of rents, it is sustained. Rent seeking creates the dynamics of capitalism, and the destruction of rents in the labor market creates a structurally more homogenous working class, that is, a working class without structural supports for its welfare.

NOTES

1. There is a possible confusion between rentals and rents. “Rental” refers to the transfer of use rights to an asset from one actor to another for a payment (a wage for labor, or interest for capital). These payments constitute returns to the holder of the benefit right to the asset. A component of this return may be an economic rent, to be discussed in detail below.

2. Rent is a component of what we ordinarily call profit, but profit as usually calculated includes a payment to capital, or interest, earned as payment for past savings and a component of wage to the

owner of the asset for his management of the asset. The latter components are not part of the rent received when interest and wages equal market rates of return.

REFERENCES

- Ainslie, George, 1992 *Piconomics*, New York, Cambridge University Press.
- Barzel, Yoram, 1997 *Economic Analysis of Property Rights*, 2nd edn., New York, Cambridge University Press.
- Becker, Gary S., 1964 *Human Capital*, New York, National Bureau of Economic Research.
- Bourdieu, Pierre, and Jean-Claude Passeron, 1990 [1970] *Reproduction in Education, Society and Culture*, translated by Richard Nice, London, Sage Publications.
- Braverman, Harry, 1974 *Labor and Monopoly Capital*, New York, Monthly Review Press.
- Collins, Randall, 1979 *The Credential Society: An Historical Sociology of Education and Stratification*, New York, Academic Press.
- Erikson, Robert, and John H. Goldthorpe, 1992 *The Constant Flux: A Study of Class Mobility in Industrial Societies*, Oxford, Oxford University Press.
- Freeman, Richard, and James L. Medoff, 1984 *What Do Unions Do?*, New York, Basic Books.
- Giddens, Anthony 1973 *The Class Structure of the Advanced Societies*, London, Hutchinson.
- Goldthorpe, John H. 1987 *Social Mobility and Class Structure in Modern Britain*, 2nd edn., Oxford, Clarendon Press.
- Gouldner, Alvin, 1979 *The Future of Intellectuals and the Rise of the New Class*, New York, Seabury Press.
- Hirsch, Fred, 1976 *The Social Limits to Growth*, Cambridge, MA, Harvard University Press.
- Holzer, Harry J., 1990 "The Determinants of Employee Productivity and Earnings," *Industrial Relations* 29, pp. 403–22.
- Jensen, Michael C., and Kevin J. Murphy, 1990 "Performance Pay and Top-Management Incentives," *Journal of Political Economy* 98, pp. 225–65.
- Juhn, Chichi, Kevin M. Murphy, and Brooks Pierce, 1993 "Wage Inequality and the Rise in Returns to Skills," *Journal of Political Economy* 101, pp. 410–42.
- Konrad, György, and Ivan Szélenyi, 1979 *The Intellectuals on the Road to Class Power*, New York, Harcourt Brace Jovanovich.
- Krueger, Alan B., and Lawrence H. Summers, 1987 "Reflections on the Inter-Industry Wage Structure," pp. 17–47 in *Unemployment and the Structure of Labor Markets*, edited by Kevin Lang and Jonathan S. Leonard, Oxford, Basil Blackwell.
- Lazear, Edward P., 1995 *Personnel Economics*, Cambridge, MA, MIT Press.
- Levy, Frank, and Richard J. Murnane, 1992 "U.S. Earnings Levels and Earnings Inequality: A Review of Recent Trends and Proposed Explanations," *Journal of Economic Literature* 30, pp. 1,333–81.
- Marshall, Alfred, 1949 [1920] *Principles of Economics*, 8th edn., London, Macmillan.
- Marx, Karl, 1959a [1848] "The Communist Manifesto," pp. 315–55 in *Capital[illegible] The Communist Manifesto, and Other Writings by Karl Marx*, edited by Ma[illegible] Eastman, New York, The Modern Library.
- Medoff, James J., and Katherine Abraham, 1981 "Are Those Paid More Really More Productive?," *Journal of Human Resources* 16, pp. 186–216.
- Mitchell, Daniel J. B., 1985 "Shifting Norms in Wage Determination," *Brookings Papers on Economic Activity*, Washington, DC, Brookings Institution.
- Murphy, Raymond, 1988 *Social Closure: The Theory of Monopolization and Exclusion*, Oxford, Clarendon Press.
- Parkin, Frank, 1979, *Marxism and Class Theory: A Bourgeois Critique*, New York, Columbia University Press.
- _____. 1992 *Durkheim*, Oxford, Oxford University Press.
- Ryan, William, 1971 *Blaming the Victim*, New York, Orbach and Chambers.
- Schneider, Louis, and Sverre Lysgaard, 1953 "The Deferred Gratification Pattern," *American Sociological Review* 18:2, pp. 142–9.
- Sørensen, Aage B., 1983 "Processes of Allocation to Open and Closed Positions in Social Structure," *Zeitschrift für Soziologie* 12, pp. 203–24.
- _____. 1998 "On Kings, Pietism and Rent-seeking in Scandinavian Welfare States," *Acta Sociologica* 41:4, pp. 363–76.
- Tullock, Gordon, 1980 "The Transitional Gains Trap," pp. 211–21 in *Toward a Theory of the Rent Seeking Society*, edited by James S. Buchanan, Robert D. Tollison, and Gordon Tullock, College Station, TX, Texas A&M University Press.
- Wolff, Edward N., 1995 *Top Heavy: A Study of Increasing Inequality of Wealth in America*, Washington, DC, Brookings Institution.
- Wright, Erik Olin, 1979 *Class Structure and Income Determination*, New York, Academic Press.

26 ■ Amartya K. Sen

From Income Inequality to Economic Inequality

Introduction

I begin by recounting a true story—a rather trivial and innocuous story, as it happens, but one with something of a lesson. Some years ago, when I went to give a lecture at another campus, I chose “Economic Inequality” as the title of my talk. On arrival, I found the campus covered with posters announcing that I was speaking on “Income Inequality.” When I grumbled about it slightly, I encountered gentle, but genuine, amazement that I wanted to fuss about such “an insignificant difference.” Indeed, the identification of economic inequality with income inequality is fairly standard, and the two are often seen as effectively synonymous in the economic literature. If you tell someone that you are working on economic inequality, it is quite commonly assumed that you are studying income distribution.

This implicit identification can be found in the philosophical literature as well. For example, in his interesting and important paper “Equality as a Moral Ideal,” Harry Frankfurt (1987), the distinguished philosopher, provides a closely reasoned critique of what he calls economic egalitarianism, defining it as “the doctrine that there should be no inequalities in the distribution of money” (p. 21).

The distinction, however, is important. Many of the criticisms of economic egalitarianism as a value or a goal apply much more readily to the narrow concept of income inequality than it does to the broader notions of economic inequality. For example, giving a larger share of income to a person with more needs, say due to a disability, can be

seen as militating against the principle of equalizing incomes, but it does not go against the broader precepts of economic equality since the greater need for economic resources due to the disability must be taken into account in judging the requirements of economic equality.

The subject of this paper is precisely the difference between economic inequality and income inequality. It will be argued that we ought to pay much more attention than we conventionally do to economic inequality in an appropriately broad sense, taking note of the fact that income inequality, on which economic analysis of inequality so often concentrates, gives a very inadequate and biased view of inequalities, even of those inequalities that can be powerfully influenced by economic policy. There is a serious gulf here, and the distinction, I would argue, is of considerable importance for economic practice as well as for economic theory. I shall also present some empirical examples, involving the United States as well as other countries, to illustrate the force of this distinction. The more difficult issue concerns the problems involved in having an appropriately broad notion of economic inequality that is both theoretically adequate and empirically usable. This question, too, I shall briefly try to address.

The Need for Going Beyond Income Inequality

A convenient point of departure is A. B. Atkinson’s (1970) pioneering move in the

measurement of inequality. He assessed inequality of incomes by bringing in an overall social objective function and measured inequality of an income distribution through the social loss (in terms of equivalent income) from that distribution in comparison with a corresponding equal distribution. However, he took the individuals to be symmetrical and also did not explicitly consider what the individuals respectively get out of their incomes and other circumstances.

There is a case for going beyond this structure and for examining the nature of individual advantages themselves as the constituent elements of social welfare (or, more generally, of social objectives). In this context, we have to take note of the heterogeneities of the individuals and of their respective nonincome circumstances.

The important point to note is that the valuation of income is entirely as a means to other ends and also that it is one means among others. A more inclusive list of means has been used by John Rawls in his theory of justice through his concentration on primary goods, which include rights, liberties and opportunities, income and wealth, and the social bases of self-respect” (Rawls 1971, pp. 60–65). Income is, of course, a crucially important means, but its importance lies in the fact that it helps the person to do things that she values doing and to achieve states of being that she has reasons to desire. The worth of incomes cannot stand separated from these deeper concerns, and a society that respects individual well-being and freedom must take note of these concerns in making interpersonal comparisons as well as social evaluations.

The relationship between income (and other resources) on the one hand and individual achievements and freedoms on the other is not constant. Different types of contingencies lead to systematic variations in the conversion of incomes into the distinct functionings we can achieve (i.e., the various things we can do or be), and that affects the

lifestyles we can enjoy. There are at least five important sources of parametric variation.

(1) *Personal heterogeneities*: People have disparate physical characteristics connected with disability, illness, age, or gender, making their needs diverse. For example, an ill person may need more income to fight her illness than a person without such an illness would need. While the compensation needed for disadvantages will vary, some disadvantages may not be correctable even with more expenditure on treatment or care.

(2) *Environmental diversities*: Variations in environmental conditions, such as climatic circumstances (temperature ranges, rainfall, flooding, and so on), can influence what a person gets out of a given level of income.

(3) *Variations in social climate*: The conversion of personal incomes and resources into functionings is influenced also by social conditions, including public health care and epidemiology, public educational arrangements, and the prevalence or absence of crime and violence in the particular location. Aside from public facilities, the nature of community relationships can be very important, as the recent literature on social capital has tended to emphasize.¹

(4) *Differences in relational perspectives*: The commodity requirements of established patterns of behavior may vary between communities, depending on conventions and customs. For example, being relatively poor in a rich community can prevent a person from achieving some elementary functionings (such as taking part in the life of the community) even though her income, in absolute terms, may be much higher than the level of income at which members of poorer communities can function with great ease and success. For example, to be able to “appear in public without shame” may require higher standards of clothing and other visible consumption in a richer society than in a poorer one (as Adam Smith [1776] had noted more than two centuries ago). The same parametric variability may apply to the personal re-

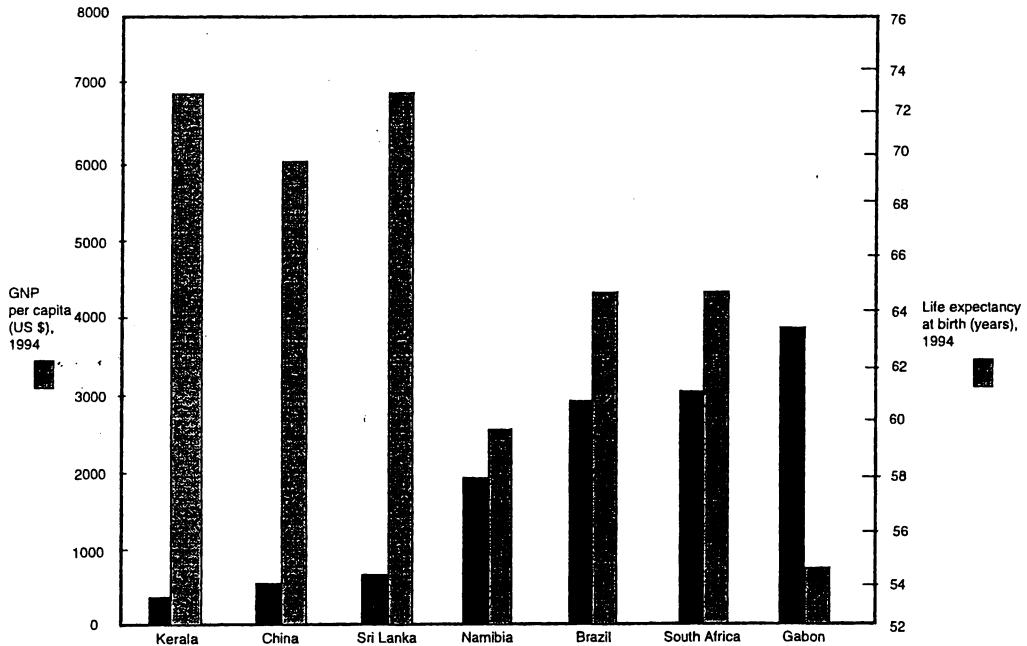


Figure 1. GNP per Capita (US \$) and Life Expectancy at Birth, 1994.

Sources: Country data, 1994, World Bank (1996); Kerala data, life expectancy, 1989–1993, sample registration system reported in Government of India (1997a); domestic product per capita, 1992–1993, Government of India (1997b).

sources needed for the fulfillment of self-respect. This is primarily an intersocietal variation rather than an interindividual variation within a given society, but the two issues are frequently interlinked.

(5) *Distribution within the family*: Incomes earned by one or more members of a family are shared by all, nonearners as well as earners. The family is, thus, the basic unit for consideration of incomes from the point of view of their use. The well-being or freedom of individuals in a family will depend on how the family income is used in furtherance of the interests and objectives of different members of the family. Thus, intrafamily distribution of incomes is quite a crucial parametric variable in linking individual achievements and opportunities with the overall level of family income. Distributional rules followed within the family (e.g., related to gender or age or perceived needs) can make a major difference to the attainments and predicaments of individual members.

Illustrations of Contrasts

I have presented elsewhere empirical examples of different types that illustrate the variability of the relation between incomes and achievements (Sen 1981, 1985a, 1995, 1998). I shall take the liberty of dwelling on a few such illustrations to indicate what kind of contrasts may be involved.

Figure 1 presents the gross national product (GNP) per head and life expectancy at birth of six countries (China, Sri Lanka, Namibia, Brazil, South Africa, and Gabon) and one sizeable state (Kerala) within a country (India).² The income-poor people of Kerala or China or Sri Lanka enjoy enormously higher levels of life expectancy than do the much richer populations of Brazil, South Africa, and Namibia, not to mention Gabon. Since life expectancy variations relate to a variety of economic influences, including epidemiological policies, health care, educational facilities, and so on, the reach of

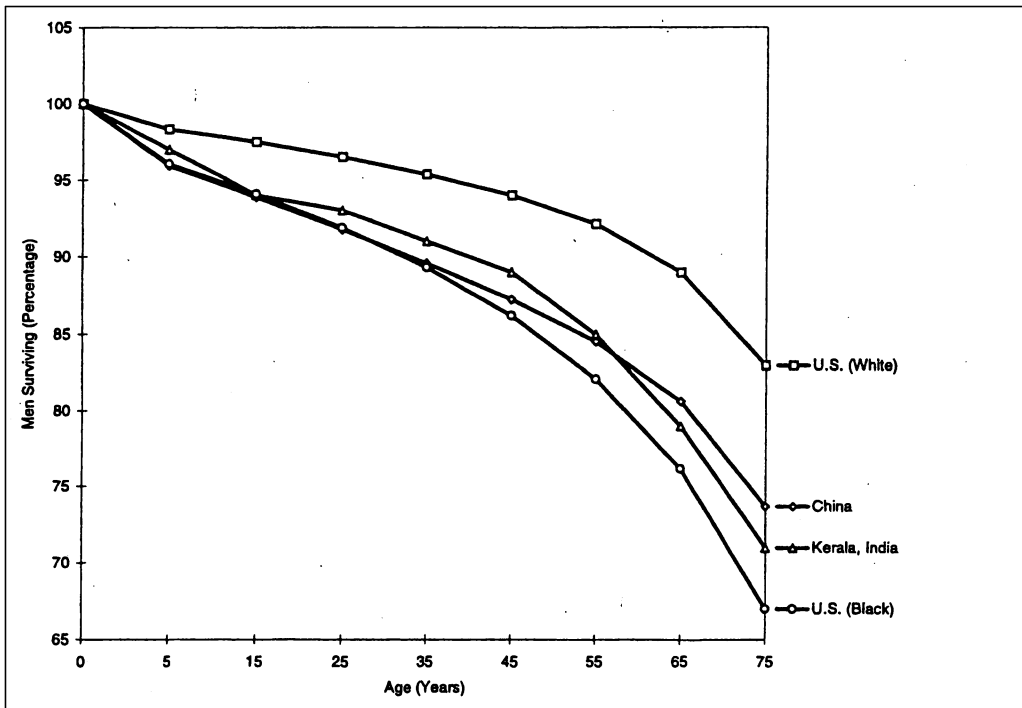


Figure 2. Variations in Male Survival Rates by Region

Sources: U.S., 1991–1993, U.S. Department of Health and Human Services (1996); Kerala, 1991, Government of India (1991); China, 1992, WHO (1994).

economic opportunities is much broader than that of income alone. I have had the occasion to discuss elsewhere how public policies in particular have been quite crucial in influencing the quality of life and longevity of different populations (see Sen 1981; Drèze and Sen 1989). In terms of inequality analysis, even the direction of the inequality points oppositely when we compare Kerala, China, and Sri Lanka on one side with Brazil, South Africa, Namibia, and Gabon on the other.

Figures 2 and 3 make a related, but differently focused, comparison, bringing in the U.S. itself. Even though the income per capita of African Americans is considerably lower than that of the American white population, African Americans are of course a great many times richer in income terms than the people of China or Kerala (even after correcting for cost-of-living differences). In this context, it is interesting to

compare the survival prospects of African Americans vis-à-vis the immensely poorer Chinese or Indians in Kerala. American blacks do much better in terms of survival at low age groups (particularly in terms of infant mortality), but the picture changes over the years.

It turns out that, in fact, the Chinese and the Keralites decisively outlive American black men in terms of surviving to older age groups. Even American black women end up having similar survival patterns for high ages as the Chinese and decidedly lower survival rates than the Indians in Kerala. So it is not only the case that American blacks suffer from relative deprivation in the income space (vis-à-vis American whites), they are absolutely more deprived than the much poorer Indians in Kerala and the Chinese (in the case of men) in terms of living to ripe, old ages. In explaining these differences between living standards judged by income per

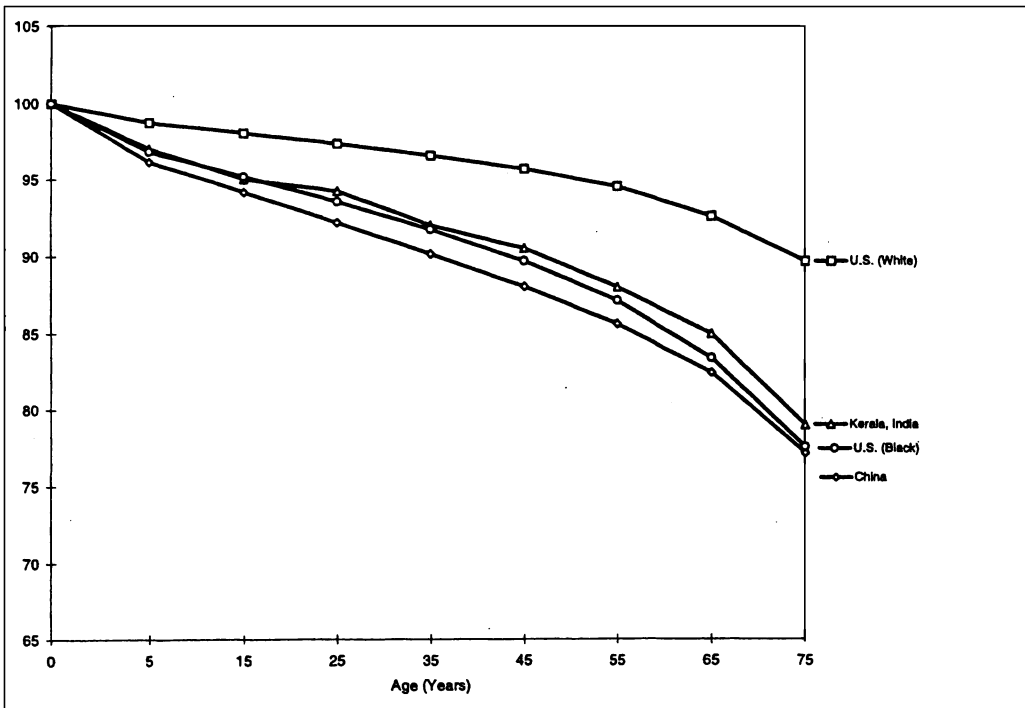


Figure 3. Variations in Female Survival Rates by Region

Sources: U.S., 1991–1993, U.S. Department of Health and Human Services (1996); Kerala, 1991, Government of India (1991); China, 1992, WHO (1994).

head and that judged by the ability to survive to higher ages, a number of causal issues are relevant (including medical insurance, public health care, elementary education, law and order, etc.) that are not unrelated to economic policies and programs.³

Figure 4 compares, for different states within India, the values of gross domestic product (GDP) per capita, literacy (female and male), life expectancy at birth (female and male), and total fertility rate. The last has eventual importance for population growth, but its inclusion here is mainly for its immediate role, at high levels, as a major restraint on the freedom and wellbeing that young women can enjoy when battered by continual bearing and rearing of children. Since the last is viewed as a negative influence on the quality of life, it is measured in the opposite (downward) direction from the zero line.

It is readily seen (as can also be confirmed by standard measures of statistical relations)

that the relative values of GDP per capita figures are much at variance with the nonincome indices of aspects of quality of life (female literacy, male literacy, female life expectancy, male life expectancy, and low fertility rate), which all move very closely together. For example, the GDP figures would put Haryana and Punjab very much higher than Tamil Nadu and Kerala, but in terms of aspects of quality of life, exactly the opposite is the case.

As these illustrations exemplify (Figures 1–4) and as can be confirmed by other statistics (see, e.g., Sen 1985a, 1995, 1998, and literature cited there), there are substantial differences between the income-based view and the nonincome indicators of quality of life. Inequality comparisons will yield very different results depending on whether we concentrate only on incomes or also on the impact of other economic and social influences on the quality of life.

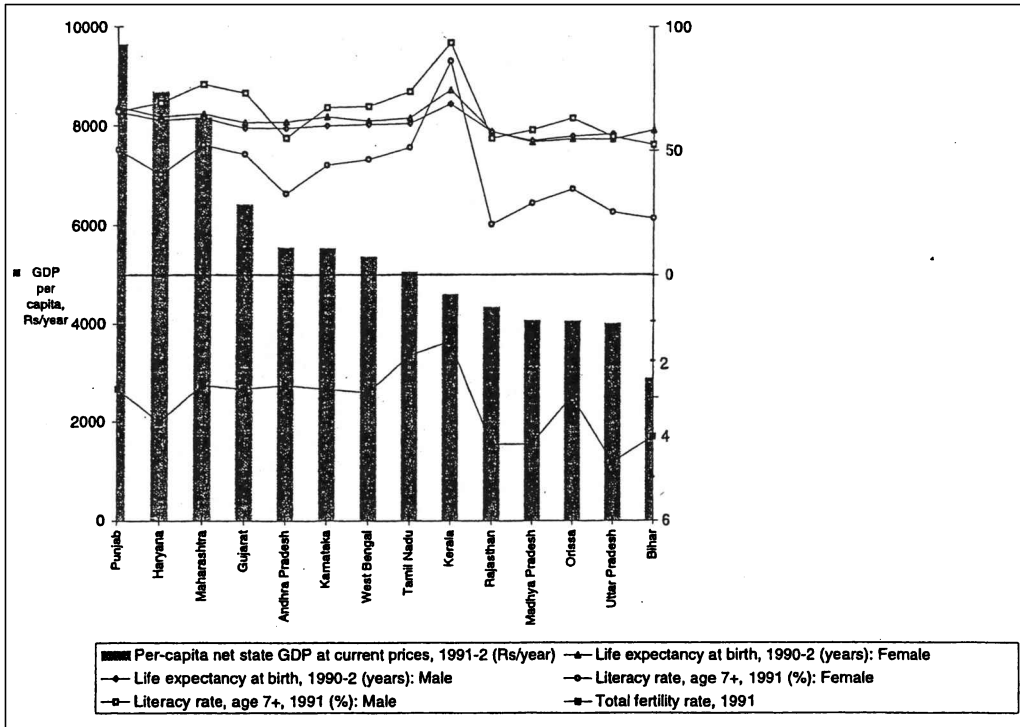


Figure 4. Selected Indicators for Indian States

Sources: Drèze and Sen (1995, Table A.3)

A further issue, which I shall not take up in this paper (but that I have addressed elsewhere, particularly in Sen [1997]), concerns the severely negative impact of unemployment, especially persistent unemployment, on the lives that people can live. This is an especially important issue for the assessment of quality of life and inequality in contemporary Europe. Even though unemployment benefits and social security may reduce the impact of the extraordinary levels of high unemployment on European income inequality in particular, the persistence of unemployment leads to many other kinds of deprivation (see Sen 1997 and literature cited therein) that are not reflected at all in the income statistics. An over-concentration on income inequality alone has permitted greater social and political tolerance of unemployment in Europe (and even some economic smugness vis-à-vis the achievement of low unemployment levels in the U.S.) that can-

not be justified if a broader view of economic inequality is taken.

Interpersonal Utility Comparisons and Inequality

The illustrations just presented of contrast between income and achievement deal with particular classes of indicators of quality of life (longevity, survival, literacy, fertility, employment status). Illustrations can also be provided to exemplify variability in the relation between income and other substantive achievements such as being healthy, being well-nourished, taking part in the life of the community, and so on.

The acceptance of variability between income and achievement is not, however, an adequate ground for a definitive rejection of income inequality as the center of our attention in inequality assessment—without considering whether an alternative approach

would be workable and satisfactory. Practical economics, no less than politics, is the art of the possible, and that issue remains, even when the need for going beyond income inequality is well accepted. Can we really get an alternative, practically usable approach based on the broader concentration on functionings rather than incomes?

Before taking on this issue fully, I would like to examine a related question, proposing a different alternative to the focus on incomes. Are we not likely, it may be sensibly asked, to be served better by opting for a more familiar notion, like utility, in shifting away from income inequality? Why not the inequality of utilities as the central focus of attention for inequality analysis? Indeed, just such a focus has been proposed and elegantly explored already by James Meade (1976) in his exploration of “the just economy.” So the utility-based evaluation of inequality should be examined first, before stepping on to the less tried, and perhaps more hazardous, field of functionings (or the freedom to function).

The possibility of interpersonal comparisons of utilities was, of course, famously challenged by Lionel Robbins (1938) and others in the high days of simple positivist criticism of utilitarian welfare economics. Certainly, the claim to a high scientific status of utility comparisons is compromised by many practical difficulties in relating observations to firm and indisputable conclusions regarding interpersonal rankings of utilities and utility differences. On the other hand, comparisons of pleasures and happiness are made in our day-to-day reflections and discourse, and there is considerable discipline in the making of such comparisons. Indeed, as Donald Davidson (1986) has pointed out, the nature of our understanding and communication regarding intrapersonal comparisons of states of happiness and desires are not radically different from the corresponding interpersonal exercises. Also, interpersonal comparison of utilities need not take an all-or-nothing form, and it is possible to

have “partial interpersonal comparability” with a rigorous analytical structure (see Sen 1970a, b).

The difficult issue in basing inequality analysis on interpersonal comparisons is not so much the impossibility of making such comparisons but the possibility of being misled by such comparisons (particularly about important differences in the substantive deals that people get and the real predicaments from which they suffer). Our ability to take pleasure in very adverse circumstances tends to adapt to the hardship of circumstances so that the badly placed underdogs do not typically spend their lives weeping over what they have missed. People learn to make the most of small opportunities and to cut desires to size, that is, to levels that are realistic under the circumstances. Thus, in the scale of pleasures and desire fulfillment, the deprivation of the persistent underdog finds rather muffled and muted expression. Deprived people, varying from subjugated housewives in sexist societies to the hopelessly poor in strongly stratified economies, come to terms with their deprivation, and the psychological indicators of pleasure or desire fulfillment may fail to reflect the extent of real deprivation that these people suffer.

There is, I believe, force in this criticism of relying on interpersonal comparison of pleasures and desire fulfillment for making judgments about inequality or injustice. However, this critique does not touch at all the more modern definition of utility as a numerical representation of individual choice behavior. In this interpretation, to say that a person gets more utility from x than from y is not essentially different from saying that, given the straightforward choice between x and y , the person would choose x . The malleability or adaptation of pleasure-taking ability need not compromise the perspective of utility as real-valued representation of preference.

However, if we see utility only as a numerical representation of each person’s choice

behavior, there is, then, no basis here for interpersonal comparisons of utility since each person's choice behavior is a distinct and separate entity. My choices may well reveal that I prefer a banana to an apple, but no choice of mine would, in any obvious sense, reveal whether I prefer to be someone else. Interpersonal comparisons deal with objects of comparison that are not objects of actual choice.

This point is often missed when it is presumed that similarity of choice behavior over commodity space must reveal a congruence of utilities. It is often presumed that, when two persons are observed to have the same demand function, then they must be seen as having the same level of interpersonally comparable utility for any given commodity bundle. Indeed, much of real-income comparison proceeds on the basis of identifying individual advantages with the commodity basket enjoyed, evaluated by a shared preference relation, and that procedure is not illegitimate for making situational comparisons of different persons' opulence. But to interpret them as utility comparison, going beyond opulence, would be a complete *non sequitur*. If instead of assuming that each person gets the same utility as others do from the same commodity bundle, it were assumed that one gets exactly one tenth of the utility that another gets from each respective bundle, then that too would be perfectly consistent with all the behavioral observations (including the shared demand function). Congruent demand functions tell us nothing about the congruence of utility functions, and this follows generally from the fact that the observations on which demand functions are based do not lend themselves to any presumption about interpersonal comparisons of well-being (only of commodity holdings and opulence).

This must not be seen as just a fussy difficulty of theoretical interest; it can make a very big difference in practice as well. Even if a person who is disabled or ill or depressed

happens to have the same demand function as another who is not disadvantaged in this way, it would be quite absurd to assume that she is having exactly the same utility or well-being from a given commodity bundle as the other can get from it. To attribute the same utility function to each and to treat that as the basis of interpersonal comparison for the analysis of inequality or injustice would be both epistemologically unsound and ethically unfair.

Utility cannot, therefore, serve as a satisfactory basis for interpersonal comparison for inequality analysis, and this holds no matter whether we interpret utility as pleasure or as desire fulfillment or as a numerical representation of choice behavior. Indeed, as I have tried to discuss, attempts to make that use arbitrarily can be pernicious for judgments of equity and justice.

Quality of Life, Functionings, and Capabilities

The choice of nonutility variables in terms of which inequality can be judged has been a matter of some interest in recent years, and such concepts as the quality of life or freedom of living and other such notions have been invoked.⁴ It is, however, important to emphasize that focusing on the quality of life rather than on income or wealth or on psychological satisfaction is not new in economics. Indeed, the origin of the subject of economics was strongly motivated by the need to study the assessment of and causal influences on the conditions of living. The motivation is stated explicitly, with reasoned justification, by Aristotle, but it is also strongly reflected in the early writings on national accounts and economic prosperity by William Petty, Gregory King, Francois Quesnay, Antoine Lavoisier, Joseph Louis Lagrange, and others. While the national accounts devised by these pioneers established the foundations of the modern concept of income, the focus of their attention was never confined to this

one concept. They were also very aware of the basic issue that the importance of income is instrumental and circumstantially contingent rather than intrinsic and categorical.

In traditional welfare economics, there has been interest both in individual utilities and in individual incomes. When individuals are taken to be symmetrical, the two are closely linked. John Rawls has pointed to the important issue that income is not the only versatile means that facilitates a person's pursuit of his or her respective objectives. He has focused instead, as was stated earlier, on the broader category of primary goods, which are general-purpose means that help anyone to promote his or her ends (including "rights, liberties and opportunities, income and wealth, and the social bases of self-respect").

The concentration on primary goods in the Rawlsian framework relates to his accounting of individual advantage in terms of the opportunities they enjoy to pursue their respective objectives. Rawls's Difference Principle, which is part of his theory of justice as fairness, assesses efficiency as well as equity in terms of the respective holdings of primary goods, represented by an index.

The broadening of the narrow concentration on incomes alone involved in this move is significant, but this widening of the informational focus from incomes to primary goods is not adequate to deal with all the relevant variations in the relationship between resources and functionings. Primary goods themselves are mainly various types of general resources, and the use of these resources to generate the capability to do things is subject to distinct types of variations (as has been already discussed), including personal heterogeneities, environmental diversities, variations in social climate, and differences in relational perspective. We can have complete equality of the chosen index of primary goods, and yet some people may be immensely more deprived than others because of age, disabili-

ties, proneness to illness, epidemiological conditions, and so on.

I have tried to argue for some time now (Sen 1980, 1985a, b, 1992) that, for many purposes, the appropriate space is neither that of utilities (as claimed by welfarists) nor that of primary goods (as demanded by Rawls). If the object is to concentrate on the individual's real opportunity to pursue her objectives, then account would have to be taken not only of the primary goods the person holds but also of the relevant personal characteristics that govern the conversion of primary goods into the person's ability to promote her ends. For example, a person who is disabled may have a larger basket of primary goods and yet have less chance to lead a normal life (or to pursue her objectives) than an able-bodied person with a smaller basket of primary goods. Similarly, an older person or a person more prone to illness can be more disadvantaged in a generally accepted sense even with a larger bundle of primary goods.

The concept of functionings, which has distinctly Aristotelian roots, reflects the various things a person may value doing or being. The valued functionings may vary from such elementary ones as being adequately nourished and being free from avoidable disease to very complex activities or personal states, such as being able to take part in the life of the community and having self-respect.

The extent of each functioning enjoyed by a person may be represented by a real number and, when this is the case, a person's actual achievement is given by a functioning vector in an n -dimensional space of n functionings (presuming finiteness of distinct functionings). When numerical representation of each functioning is not possible, the analysis has to be done in terms of the more general framework of seeing the functioning achievements as a functioning n -tuple and the capability set as a set of such n -tuples in the appropriate space (this will, then, not be

a vector space). The set of alternative functioning vectors available to her for choice is called her capability set. While the combination of functionings (strictly, n -tuples) a person undertakes reflects her achievements, the capability set represents the freedom to achieve: the alternative functioning combinations from which this person can choose.

Figure 5 illustrates a functioning space (two dimensional), with the capability set of a person being given by region K , and from this capability set K , the person chooses one functioning vector x (though this need not necessarily be unique). It may be useful to think of choice in this space in terms of an indifference map of valued living defined over the functioning vectors, and x can then be seen as belonging to the highest reachable indifference curve (as indicated).⁵ The focus of this capability approach could be either on the realized functionings (what a person is actually able to do) or on the set of alternatives she has (her real opportunities).

I shall not go into the details of the approach here (which I have tried to present elsewhere [Sen 1985a, 1992, 1993]). But it is useful to ask whether the focus on capability is likely to be very different from that on functionings. The capability approach can be used either with a focus on what options a

person has—given by the whole capability set—or on the actual functioning combination she chooses—given by the chosen functioning vector. In the former procedure, what may be called the options application, the focus can be on the entire K , whereas in the latter, the choice application, the concentration is more narrowly on x . The options application is directly concerned with the freedom to choose over various alternatives, whereas the choice application is involved with the alternative that is actually chosen. Both the versions of the capability approach have been used in the literature, and sometimes they have been combined.

By a well-established tradition in economics, the real value of a set of options lies in the best use that can be made of them and, given maximizing behavior and the absence of uncertainty, the use that is actually made. The use value of the opportunity, then, lies derivatively on the value of one element of it (to wit, the best option or the actually chosen option). In this case, the focusing on chosen functioning vector coincides with concentration on the capability set. With this type of elementary evaluation, the two uses of the capability approach share not only the identification of a relevant space (that of functionings) but also the focal variable in that space (the chosen functioning vector).

However, the options application can be used in other ways as well since the value of a set need not invariably be identified with the value of the best, or the chosen, element of it. It is possible to attach importance to having opportunities that are not taken up. This is a natural direction to go if the process through which outcomes are generated is of significance of its own. Indeed, choosing itself can be seen as a valuable functioning and having an x when there is no alternative may be sensibly

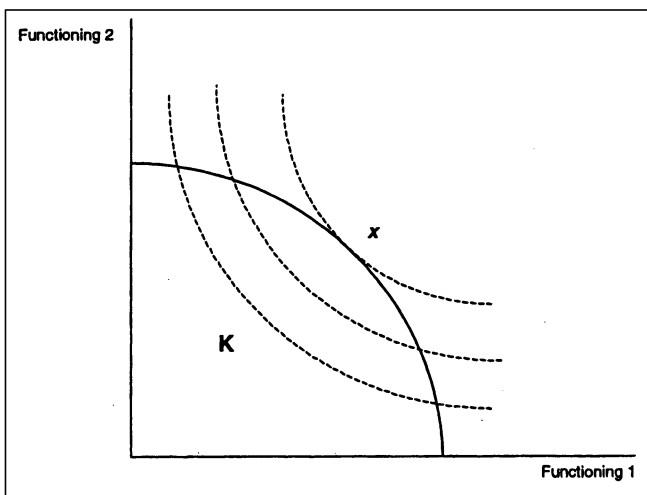


Figure 5. Functioning Vectors and the Preference Map

distinguished from choosing x when substantial alternatives exist.⁶ The importance of this type of consideration lies more in drawing attention to broader concerns than in offering a quick resolution of interpersonal comparison of freedoms (and thus of overall individual advantages that take note of the significance of freedom).

Weights, Valuations, and Explicitness

I turn now to a crucial methodological issue that has received much attention in recent discussions involving the capability approach and related proposals. The heterogeneity of functionings involves the need to weigh them against one another. This would apply to all approaches geared to functionings, whether the concentration is on realized functioning vectors x (as with the choice application) or on the capability sets K (as with the options application).⁷

Is this weighting requirement a special difficulty associated with the capability approach? This cannot be the case since heterogeneity of factors that influence individual advantage is a pervasive feature of actual evaluation. While we can decide to close our eyes to this issue by simply assuming that there is something homogeneous (e.g., income or utility) in terms of which everyone's overall advantage can be judged and interpersonally compared (and that variations of needs, personal circumstances, etc., can be, correspondingly, assumed away), this does not resolve the problem—it only evades it.

Comparisons of real income involve reduction of bundles of different commodities into points on a real line and, in judging comparative individual advantages, there is the further problem of interpersonal comparisons taking note of variations of individual conditions and circumstances. As was discussed earlier (see section 4), even when each person's preference is taken to be the ultimate arbitrator of well-being for that person and even when, to take a very special case, every-

one has the same demand function or preference map, the comparison of market valuations of commodity bundles (or their relative placing on a shared system of indifference map in the commodity space) may tell us rather little about interpersonal comparisons of well-being.

In evaluative traditions involving fuller specification, considerable heterogeneity is explicitly admitted. For example, in Rawlsian analysis, primary goods are taken to be constitutively diverse (including “rights, liberties and opportunities, income and wealth, and the social bases of self-respect”), and Rawls (1971) deals with them through an overall index of primary goods holdings. While a similar exercise of judging over a space with heterogeneity is involved both in the Rawlsian approach and in the use of functionings, the former is informationally poorer, for reasons discussed already, because of the parametric variation of resources and primary goods vis-à-vis the opportunity of achieving high quality of living.

The problem of valuation is not, however, one of an all-or-nothing kind. Some judgments, with incomplete reach, follow immediately from the specification of a focal space. When some functionings are selected as significant, such a focal space is specified, and the relation of dominance itself leads to a partial ordering over the alternative states of affairs. If person i has more of a significant functioning than person j and at least as much of all such functionings, then i clearly has a higher valued functioning vector than j has. This partial ordering can be extended by further specifying the possible weights. A unique set of weights will, of course, be sufficient to generate a complete order, but it is typically not necessary. Given a range of weights on which there is agreement (i.e., when it is agreed that the weights are to be chosen from a specified range, even without any agreement as to the exact point on that range), there will be a partial ordering based on the intersection of rankings (see Figure 6).

This partial ordering will get systematically extended as the range is made more and more narrow. Somewhere in the process of narrowing the range, possibly well before the weights are unique, the partial ordering will become complete. But even with an incomplete ordering, many decision problems can be adequately resolved, and even those that are not fully resolved can be substantially simplified (through the rejection of dominated alternatives).

It is thus crucial to ask, in any evaluative exercise of this kind, how the weights are to be selected. This judgmental exercise can be resolved only through reasoned evaluation. For a given person who is making his or her own judgments, the selection of weights will require reflection rather than interpersonal agreement or a consensus. However, in arriving at an agreed range for social evaluation (e.g., in social studies of poverty), there has to be some kind of a reasoned consensus on weights or at least on a range of weights. This is a social choice exercise and requires public discussion and a democratic understanding and acceptance. It is not a special problem that is associated only with the use of the functioning space.

A Concluding Remark

The argument for shifting our attention from income inequality to economic inequality relates to the presence of causal influences on individual well-being and freedom that are economic in nature but that are not captured by the simple statistics of incomes and commodity holdings. The case for such broadening of informational focus also entails the need to pay evaluative attention to heterogeneous magnitudes and calls for the derivation of partial orderings based on explicit or implicit public acceptance. The normative force of this acceptance rests substantially on the quality and reach of public discussions on matters of central social concern. The subject of this essay, though nominally about inequality, is ultimately as much about the nature and importance of public discussion on social evaluation.

In matters of public judgment, there is no real escape from the evaluative need for public discussion. The work of public valuation cannot be replaced by some "super clever" assumption. Some assumptions that give the appearance of working very well operate

through hiding the choice of values and weights in some constructed opaqueness. For example, the assumption, often implicitly made, that two persons with the same demand function must have the same relation between commodity bundles and well-being (no matter whether one is ill and the other not, one disabled and the other not, etc.) is basically a way of evading the consideration of significant influences on well-being. That evasion becomes transparent, as I have tried to illustrate, when we supplement income and commodity data by information of other types (including matters of life and death).

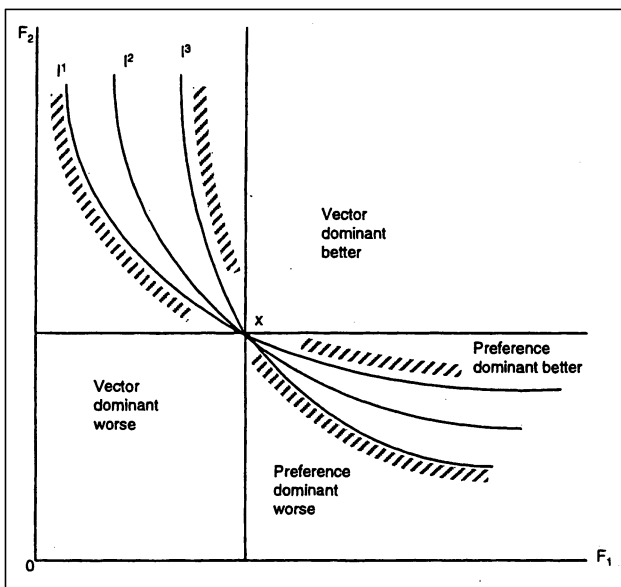


Figure 6. Functioning Vectors and Dominance

The exercise need not, however, be as exacting as it may first appear and as it certainly would be if we were not to settle for anything less than getting complete orderings of interpersonal advantages and inequalities. Our values about inequality aversion are not typically of the fine-tuning variety, getting the level of inequality “just right,” taking note of all its pros and cons. Rather, the engagement is mainly about the avoidance of substantial inequalities and serious injustice.

As material for public discussion and for informed consensus or acceptance, the need is not so much for a complete ordering of interpersonal advantages and of levels of inequality (which would be inevitably based on some crude assumptions and evasions) but for usable partial orderings that capture the big inequalities in a clear way, taking note of the various significant concerns that go well beyond the commodity space. The focus has to be on the reach and relevance of partial orderings that can be cogently derived and used. Insistence on completeness can be an enemy of informed and democratic decision making.

NOTES

1. See particularly Coleman (1986) and Putnam, Leonardi, and Nanetti (1993).

2. While Kerala is merely a state rather than a country, nevertheless, with its population of 29 million, it is larger than the majority of countries in the world.

3. On this, see Sen (1993) and also the medical literature cited therein. See also the discussion that American black men from the Harlem district of rich New York fall (in terms of survival) not only behind the Chinese or the Indians in Kerala but also behind the famished population of Bangladesh.

4. See the literature considered in Nussbaum and Sen (1993).

5. The use of such an indifference map in explaining valuation of functionings may be of considerable pedagogic value, especially in moving from the familiarity of the commodity space to the unaccustomed functioning space. It is, nevertheless, important to recognize that the nature of the indifference map in

the functioning space may not altogether mirror what we standardly presume in the case of commodity space. In particular, there may be considerable areas of incompleteness as well as fuzziness (see Sen 1985a). The recent literature on fuzzy set theory can be helpful in analyzing the valuation of functioning vectors and capability sets (see particularly Chiappero Martinetti 1994, 1997; Delbono 1989; Cerioli and Zani 1990; Balestrino 1994; Balestrino and Chiappero Martinetti 1994; Ok 1995; Casini and Bernetti 1997; among other contributions).

6. See Sen (1985a, b). There remains the more difficult issue of determining how this process consideration should be incorporated. For various alternative proposals and also axiomatized formulas, see Suppes (1987), Pattanaik and Xu (1990), Sen (1991), Foster (1993), Arrow (1995), Herrero (1995), Puppe (1995), among others.

7. In the latter case, there is the further task of comparing sets rather than points in this space, and it involves the additional issue that the importance of freedom can stretch well beyond the value of the particular element that is chosen (except in the special case of elementary evaluation).

REFERENCES

- Arneson, R. 1989. Equality and equality of opportunity for welfare. *Philosophical Studies* 56:77–93.
- Arneson, R. 1990. Liberalism, distributive subjectivism, and equal opportunity for welfare. *Philosophy and Public Affairs* 19:158–94.
- Arrow, K. J. 1993. *Social choice and individual values*. 2nd edition. New York: Wiley.
- Arrow, K. J. 1995. A note on freedom and flexibility. In *Choice, welfare and development: A festschrift for Amartya K. Sen*, edited by K. Basu, P. Pattanaik, and K. Suzumura. Oxford, UK: Clarendon Press.
- Atkinson, A. B. 1970. On the measurement of inequality. *Journal of Economic Theory* 2:244–63.
- Balestrino, A. 1994. Poverty and functionings: Issues in measurement and public action. *Giornale degli Economisti e Annali di Economia* 53:389–406.
- Balestrino, A., and E. Chiappero Martinetti. 1994. Poverty, differentiated needs, and information. Mimeographed, University of Pisa and University of Pavia.
- Casini, L., and I. Bernetti. 1997. Environment, sustainability, and Sen's theory. *Notizi di Politeia* 12:55–78.
- Cerioli, A., and S. Zani. 1990. A fuzzy approach to the measurement of poverty. In *Income and wealth distribution, inequality and poverty*, edited by C. Dagum and M. Zenga. New York: Springer.

- Chiappero Martinetti, E. 1994. A new approach to evaluation of well-being and poverty by fuzzy set theory. *Giornale degli Economisti e Annali di Economia* 53:367–88.
- Chiappero Martinetti, E. 1997. Standard of living evaluation based on Sen's approach: Some methodological suggestions. *Notizie di Politeia*, 12:37–53.
- Coleman, J. S. 1986. *Individual interests and collective action: Selected essays*. Cambridge, UK: Cambridge University Press.
- Davidson, Donald. 1986. Judging interpersonal interests. In *Foundations of social choice theory*, edited by J. Elster and A. Hylland. Cambridge, UK: Cambridge University Press, pp. 195–211.
- Delbono, F. 1989. Poverta come incapacita: Premesse teoriche, identificazione e misurazione. *Rivista Internazionale di Scienze Sociali* 97.
- Drèze, J., and A. K. Sen. 1989. *Hunger and public action*. Oxford, UK: Clarendon Press.
- Drèze, J., and A. K. Sen. 1995. *India: Economic development and social opportunity*. Oxford, UK: Clarendon Press.
- Foster, J. E. 1993. Notes on effective freedom. Unpublished paper, presented at the Stanford Workshop on Economic Theories of Inequality, sponsored by the MacArthur Foundation.
- Frankfurt, H. 1987. Equality as a moral ideal. *Ethics* 98:21–43.
- Government of India. 1991. *Sample registration system: Fertility and mortality indicators*. New Delhi.
- Government of India. 1997a. *Women in India: A statistical profile*. New Delhi: Department of Education.
- Government of India. 1997b. *Economic survey 1996–97*. New Delhi: Ministry of Finance.
- Herrero, Carmen. 1995. Capabilities and utilities. Unpublished paper, University of Alicante and IVIE, Spain.
- Meade, J. E. 1976. *The just economy*. London: Allen and Unwin.
- Nussbaum, M., and A. K. Sen, eds. 1993. *The quality of life*. Oxford, UK: Clarendon Press.
- Ok, E. 1995. Fuzzy measurement of income inequality: A class of fuzzy inequality measures. *Social Choice and Welfare* 12:111–36.
- Pattanaik, P., and Y. Xu. 1990. On ranking opportunity sets in terms of freedom of choice. *Recherches Economiques de Louvain* 56:383–90.
- Puppe, Clemens. 1995. Freedom of choice and rational decisions. *Social Choice and Welfare* 12:137–53.
- Putnam, R., R. Leonardi, and R. Y. Nanetti. 1993. *Making democracy work: Civic traditions in modern Italy*. Princeton, NJ: Princeton University Press.
- Rawls, J. 1971. *A theory of justice*. Cambridge, MA: Harvard University Press.
- Robbins, L. 1938. Interpersonal comparisons of utility: A comment. *Economic Journal* 48:635–41.
- Sen, A. K. 1970a. *Collective choice and social welfare*. San Francisco: Holden Day. Republished Amsterdam: North-Holland, 1979.
- Sen, A. K. 1970b. Interpersonal aggregation and partial comparability. *Econometrica* 38:393–409.
- Sen, A. K. 1980. Equality of what? In *Tanner lectures on human values*, edited by S. McMurrin. Cambridge, UK: Cambridge University Press.
- Sen, A. K. 1981. Public action and the quality of life in developing countries. *Oxford Bulletin of Economics and Statistics* 43:287–319.
- Sen, A. K. 1985a. *Commodities and capabilities*. Amsterdam: North-Holland.
- Sen, A. K. 1985b. Well-being, agency and freedom: The Dewey lectures 1984. *Journal of Philosophy* 82:169–221.
- Sen, A. K. 1991. Welfare, preference and freedom. *Journal of Econometrics* 50:15–29.
- Sen, A. K. 1992. *Inequality reexamined*. Oxford: Oxford University Press and Cambridge, MA: Harvard University Press.
- Sen, A. K. 1993. The economics of life and death. *Scientific American* 268(May):40–7.
- Sen, A. K. 1995. Demography and welfare economics. *Empirica* 22:1–21.
- Sen, A. K. 1997. Inequality, unemployment and contemporary Europe. Paper presented at the Calouste Gulbenkian Conference on A Joint Europe, May 1997. *International Labour Review*. In press.
- Sen, A. K. 1998. Mortality as an indicator of economic success or failure. (Text of the first Innocenti Lecture, UNICEF, Florence, 1995.) *Economic Journal*. In press.
- Smith, A. 1776. *An inquiry into the nature and causes of the wealth of nations*, edited by R. H. Campbell and A. S. Skinner. Oxford: Clarendon Press.
- Suppes, P. 1987. Maximizing freedom of decision: An axiomatic analysis. In *Arrow and the foundations of economic policy*, edited by G. R. Feiwel. London: Macmillan.
- U.S. Department of Health and Human Services. 1996. *Health, United States 1995*. Hyattsville, MD: U.S. Department of Health and Human Services.
- WHO. 1994. *World health statistics annual 1994*. Geneva: World Health Organization.
- World Bank. 1996. *World development report*. New York: Oxford University Press.

27 ■ Kim A. Weeden, Young-Mi Kim,
Matthew Di Carlo, and David B. Grusky

Is the Labor Market Becoming More or Less Gradational?

The spectacular takeoff in earnings inequality within the United States has led to much debate about its sources and causes. It is hardly surprising that such a sudden takeoff in inequality, one that reversed what had been a decades-long decline, would generate so much research on its causes. Important as the resulting causal agenda has been, it is perhaps equally important to examine the *consequences* of the takeoff, especially its consequences for the class structure. The purpose of this chapter is to ask whether the transition to a high-inequality regime in the United States has strengthened or weakened the class structure.

How could we know so little about such consequences despite decades of research on the takeoff? The simple answer is that the research literature on earnings inequality has been dominated by economists and that economists have never regarded classes as an important prism through which inequality might be viewed. This answer, while perhaps nominally on the mark, is nonetheless silent on the more fundamental matter of why economists have shown so little interest in the sociological model of class. It might well be imagined that economists would find the sociological model of increasing interest as the old income paradigm is called into question and multidimensional representations of inequality gain sup-

port (e.g., Sen, Ch. 26; Bourguignon 2006). Indeed, insofar as social classes are understood as institutionalized packages of valued goods, they then become an omnibus inequality measure that is pregnant with information on precisely those noneconomic variables (e.g., working conditions, life chances, lifestyles) that a multidimensional approach emphasizes (see Grusky and Weeden 2007).

The bivariate class-earnings distribution accordingly tells us whether earnings are becoming more or less tightly associated with the various inequality dimensions that social classes signal. If both earnings inequality and the earnings-class association are increasing, it means that the rich are not just growing ever richer, but they are also more likely to be advantaged on other inequality dimensions. Conversely, it means that the poor are not just growing ever poorer, but they are also more likely to be disadvantaged on other inequality dimensions. These results would imply that a stark winner-take-all form of inequality is emerging. It follows that economists, at least those with multidimensionalist sympathies, should care about the class-earnings distribution and whether it reveals a form of inequality more problematic than univariate analyses of earnings alone can reveal.

For sociologists, there is likewise much interest in assessing whether the inequality

This is a commissioned chapter that draws heavily on previously published material (Kim A. Weeden, Young-Mi Kim, Matthew Di Carlo, and David B. Grusky, 2007, "Social Class and Earnings Inequality," *American Behavioral Scientist* 50, pp. 702–736).

dimensions are crystallizing, but the class-earnings association is further of interest because it speaks to the viability of class models themselves. That is, because class models are viable insofar as the various dimensions of inequality cohere into classes comprising distinct packages of goods, the class analyst has an interest in demonstrating that some amount of coherence indeed obtains. It is accordingly common for class analysts to attempt to validate their models by demonstrating that class and earnings are closely related (e.g., Beck 1992; Hout, Brooks, and Manza 1993). Put differently, sociologists care about crystallization not just because it reveals whether inequality is becoming an all-or-nothing affair, but also because a class representation of inequality is methodologically feasible only insofar as inequality is quite crystallized. If it is, inequality becomes a lumpy, group-determined phenomenon rather than a more continuous and individualistic one. As Giddens (Ch. 14) famously (if awkwardly) put it, “class structuration” is well developed insofar as class position is determinative of a host of life conditions, including most obviously earnings.

We will carry out our analyses in the context of a model that examines labor market structure at both the big-class and micro-class levels (see Weeden et al. 2007; Weeden and Grusky 2005a, 2005b, 2005c; Grusky 2005a, 2005b; Grusky and Sørensen, Ch. 17). In operationalizing big classes, we take the usual approach of aggregating “similar” detailed occupations, with such aggregations yielding categories like professional, manager, clerical worker, craft worker, or farmer. As we noted above, these aggregations are often justified on the argument that class members share broadly similar life conditions (e.g., prestige, earnings, cultural capital), although reference is sometimes additionally or alternatively made to the similarity of their on-the-job working conditions and employment contracts (e.g., Lockwood 1989; Erikson and Goldthorpe, Ch.

47). In this sense, the categories of big-class schemes are largely statistical constructions of academics or Census Bureau officials, whereas the detailed occupations or “micro-classes” (e.g., lawyer, secretary, carpenter) that comprise big-class categories are more deeply institutionalized constructions of employers, employees, and many others. This institutionalization takes the form of (1) workers representing their career aspirations in occupational terms, (2) professional and vocational schools training workers for occupationally defined skills, (3) professional associations and labor unions forming around occupational designations, and (4) employers constructing, advertising, and remunerating jobs in terms of occupational labels (e.g., Treiman 1977; Wilensky 1966).

We simply don’t know whether the takeoff in earnings inequality has strengthened big-class distinctions, occupational distinctions, or both. This is not to suggest that big classes and detailed occupations have been ignored altogether. To the contrary, many analyses of earnings inequality have featured big classes of one sort or another (often referred to as “occupations”), while a handful of analyses have brought in detailed occupations in some way (Groshen 1991; Howell and Wolff 1991; Levy 1998; Murphy and Welch 1993). These efforts fall short for our purposes either because (1) only a single time period is analyzed, or (2) the decomposition is carried out in terms of big classes alone or detailed occupations alone, not both at once.

The structure of the takeoff is best understood, we will argue, by recognizing that detailed occupations are nested in big classes and then apportioning the earnings inequality into three components: (1) the between-class (BC) component (i.e., inequality between big classes); (2) the between-occupation/within-class (BO/WC) component (i.e., inequality between the detailed occupations that comprise each big class); and (3) the within-occupation (WO) component (i.e., inequality within detailed occupations). The BO/WC

and WO components are conflated as “within-class inequality” in analyses that use big classes alone, and the BC and BO/WC components are conflated as “between-occupation inequality” in analyses that use detailed occupations alone. We argue below that these components are driven by very different mechanisms and are therefore usefully distinguished in understanding trends.

The further virtue of distinguishing these three components is that it allows us to identify whether trends in inequality are playing out in ways that strengthen either big classes or occupations, strengthen neither, or strengthen both. As an example of one possible outcome, Figure 1 graphs a constellation of changes that, taken together, are distinctly pro-class in their implications. In this stylized figure, the labor market has just three big classes (i.e., BC1, BC2, and BC3), with three micro-classes (i.e., MC1–MC3, MC4–MC6, MC7–MC9) then nested in each of these big classes. Within each micro-class, the $\log(\text{wages})$ of incumbents are uniformly distributed, an implausible but presentationally convenient functional form. The two vertical axes reveal a pattern of change in which the WO component is shrinking, the BO/WC component is shrinking, and the BC component is growing. The second axis (pertaining to “time 2”) thus suggests a three-class society marked by substantial between-class differences in $\log(\text{wages})$ and trivial within-class differences. By contrast, the transition to a micro-class regime, as represented in Figure 2, has the WO component again decreasing, whereas the BO/WC and BC components now move in directions opposite to those shown in Figure 1. This change effectively yields a labor market with nine small classes rather than three big ones.

It is quite plausible that we will instead observe roughly commensurate increases in the WO, BO/WC, and BC components. This type of change simply “stretches out” the overall distribution and leaves the relative sizes of the WO, BO/WC, and BC compo-

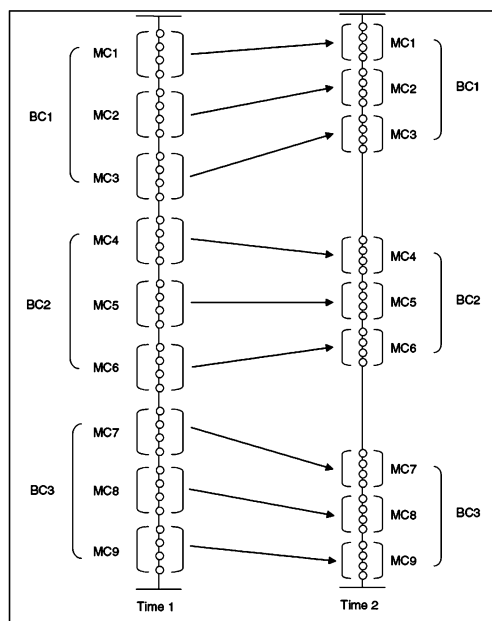


Figure 1. The Transition to a Big-Class Regime
 Note: BC=big class; MC=micro class

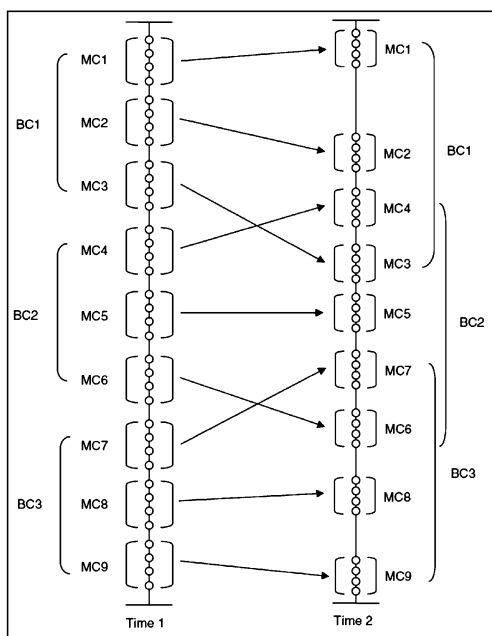


Figure 2. The Transition to a Micro-Class Regime
 Note: BC=big class; MC=micro class

nents unchanged (see Allison 1978; also, Sen 1973). If an absolutist tack is taken, all that matters is that such change increases the earnings heterogeneity within classes and occupations, thereby undermining the class principle (e.g., Kim and Sakamoto 2005).¹ The relativist counterargument is that (1) the social meaning of inequality is gradually recalibrated to adjust to an increased dispersion in earnings, and (2) a commensurate change in all components is therefore a class-neutral change (see Blau and Kahn 2002). This counterargument suggests that one should care principally about the relative sizes of the WO, BO/WC, and BC components. We will not privilege either argument but rather will cater to both by presenting evidence on relative as well as absolute change in the inequality components.

Sources of Trends

It is instructive to identify the mechanisms that underlie change in the three components of inequality and thus strengthen or weaken the class principle. As will become evident, these mechanisms differ across the three components, implying that rates of change may also differ across the components.

The BC Trend

The between-class (BC) component is well studied within the literature and hence can be dispensed with quickly. In understanding the BC component, it is relevant that entry into different big classes is linked to different levels of schooling, with the professional and managerial classes usually requiring a college or more advanced degree, the sales and clerical classes typically requiring at least some college education, and the manual classes (i.e., craft, operative, laborer) serving as the default location for those with high school degrees or less. This link between big-class categories and educational levels is hardly perfect, not just because conventional big-class schemes typically make distinctions be-

tween categories that have roughly similar educational requirements (e.g., sales, clerical), but also because entry requirements for most big classes are not *explicitly* tied to specific educational degrees (e.g., Weeden and Grusky 2005a; Grusky and Sørensen, Ch. 17). Most obviously, entry into the craft class is not limited to those with high school degrees, nor is entry into the operative or laboring classes reserved exclusively for those who lack such degrees. Even so, there is clearly a strong association between class and educational qualifications, and the BC component should therefore fluctuate with changes in the returns to different levels of schooling. If big classes are effectively just education groups in disguise, the BC trend line should be quite flat in the 1970s by virtue of the relatively large supply of college-educated workers but then increase rapidly in the 1980s as the demand for college-educated workers begins to outstrip the supply (see Katz and Murphy 1992; Levy and Murnane 1992).

In the foregoing story, classes are mere statistical entities that stand in as markers of educational qualifications, and the trend in the BC component simply shadows the well-known trend in returns to education. Could one imagine a story about the BC component in which classes are more than such indirect “stand-ins?” This type of story becomes available in societies where big classes are deeply institutionalized and their representatives participate formally in collective bargaining. For example, big classes in Sweden are conventionally represented as meaningful wage-setting actors, not just statistical categories that indirectly signal changes in market-driven returns to schooling. The BC component will increase in Sweden whenever the upper class does unusually well in its collective bargaining or the lower class does unusually poorly in its collective bargaining. To some extent, big unions in the United States engage in similar collective bargaining and wage-setting efforts, but they do not map onto entire big

classes nearly as neatly as in the Swedish case, nor do they command anything approaching the wage-setting power of their Swedish counterparts. It follows that the BC component will fluctuate in the United States largely in response to changes in the returns to education or changes in the educational composition of classes.

The BO/WC Trend

The direction of the BO/WC trend does not fall out quite so directly from an existing empirical literature (cf. Weeden, Ch. 18). We can, however, lay out two simple hypotheses about this trend, the first directly related to demand shifts that favor high-skill labor, and the second directly related to the spread of occupationalization and occupational closure.

The first of these two stories, a within-class variant of the standard demand-side account, begins with the presumption that each big class is a mixture of high-skill and low-skill occupations. Clearly, the skill differential within classes will be much smaller than the skill differential for the full labor market, but one would nonetheless anticipate a nontrivial differential within even the most homogeneous classes. If the demand for labor within the high-skill occupations of each big class outstrips the supply of labor for those occupations, their earnings will be driven up and the BO/WC component will accordingly grow. For example, the relative demand for high-skill computer programmers and related technical occupations within the professional class clearly took off during the 1980s, whereas the demand for other less skilled occupations within the professional class (e.g., photographers) appears to have increased more slowly. The increasing demand for computer programmers and related technical occupations can itself be understood as part of the shift toward an “information economy” driven by technological change and outsourcing of low-skill work.

Should we anticipate a quick increase in the supply of high-skill labor in response to

these higher wages? Although one might expect relatively quick responses to demand-induced wage premia that emerge either within occupations or between classes, there is good reason to believe that interoccupational wage premia will be preserved over the longer run. This is because occupations have often established control over entry in ways that make slow-growth policies viable. That is, many high-skill occupations control entry via licenses, credentials, training requirements, and certification, thus allowing them to protect the wages of current incumbents by restricting supply (see Weeden, Ch. 18). In this sense, a demand-shift story becomes especially plausible when applied to the BO/WC component, as here one finds institutionalized mechanisms of closure that give real staying power to demand-induced wage premia.

The second main source of BO/WC trend also references these institutionalized mechanisms of closure but turns attention to the gradual diffusion of such mechanisms over the last quarter century. The process of occupationalization, one of the main rent-generating forces of our time, can raise the wages of occupational incumbents by (1) increasing the occupation’s control over the supply of labor, and (2) reducing competition from other potential providers of the product or service and therefore increasing demand (see Weeden, Ch. 18; Abbott 1988). For example, the American Bar Association (ABA) is not only involved in establishing certification requirements and hence better controlling the supply of labor, but is also involved in ensuring that lawyers, rather than representatives of other occupations (e.g., paralegals, accountants), maintain as much control as possible over the dispute-adjudicating niche in the labor market.

If closure movements were a random process to which all occupations, both high skill and low skill alike, were equally subjected, then they would not have any implications for within-class inequality. However, high-skill occupations appear to have been

especially successful over the last quarter century in securing closure of this sort, thereby raising their already high wages yet higher. This process has been especially prominent within the professions (i.e., “professionalization”) and other nonmanual big classes. Although some high-skill occupations in the manual sector have attempted to follow suit with professionalization projects of their own (e.g., truck drivers, electricians), these projects are neither as common nor as successful as those within the nonmanual sector.

If professionalization has been the method of choice for securing the rewards of social closure within the nonmanual sector, craft unionization is the corresponding method of choice within the manual sector. The latter method has fared less well of late than has professionalization. As craft unions weaken, it has become difficult for them to maintain control over the supply of labor and to capture the rent that such control makes possible, and the wage gap between the privileged and less privileged craft occupations has accordingly begun to close. The BO/WC trend may therefore be suppressed by such countervailing “deoccupationalizing” effects in the manual big classes.

The forces making for BO/WC trend are therefore deeply dependent on the institutional mechanisms by which occupationalization is either weakened or strengthened. Indeed, even when one understands the BO/WC trend line in terms of a standard demand-side account, the forces of occupationalization are still relevant because they affect the long-term staying power of demand-induced wage premia. The importance of occupationalization for the BO/WC component leads us to hypothesize that this component will particularly take off in those sectors of the labor market, such as the nonmanual sector, in which occupationalization has been unusually successful.

The WO Trend

The final component of interest is the within-occupation variance (WO). In the

economics literature, the closest analogue to the within-occupation variance is what has conventionally been termed the “within-group” variance, where this refers to the residual variance in earnings that falls out of standard Mincer-type human capital regressions (using such variables as age, schooling, and interactions between age and schooling). Although the groups that are referenced by such methods are statistical cross-classifications rather than institutionalized occupations (cf. Groshen 1991), it is still worth examining the ways in which the within-group variance is presumed to be generated.

The stylized fact with which this literature begins is that most of the growth in overall wage inequality is attributable to growth in such residual inequality (e.g., Acemoglu 2002; DiNardo, Fortin, and Lemieux 1996; Juhn, Murphy, and Pierce 1993; Katz and Autor 1999; cf. Card and DiNardo 2002; Lemieux 2003). This growth is frequently, although not exclusively, attributed to the rising demand for skill, where the skills of interest are not indexed by observable schooling but by unobservable social, cognitive, or personality traits (see Levy and Murnane 1992, pp. 1365–7). The demand for skilled labor started expanding in the 1970s, yet the effect of this expansion on the wage premium for observable schooling was suppressed until the 1980s, when finally the supply of college-schooled labor no longer kept pace with the growing demand for such labor. By contrast, the relative supply of unobservable skills may be presumed to be constant over time, meaning that the wage premium for such skills should steadily expand and generate a long-term upward trend in residual inequality over the 1970s and beyond (see especially Acemoglu 2002).

The rising demand for unobservable skills, which is typically treated as exogenous, can itself be attributed to any number of causes, including skill-biased technological change and shifts in the structure of final demand (also see Stone 2004; Frank and Cook 1995).

Whatever its source, a rising payoff to unobservable skills will generate a growth in within-group inequality, just as a rising payoff to observable skills generates a growth in between-group inequality. This type of account has not been subjected to anything approaching a definitive test, yet it remains attractive because of its consistency with competitive wage theory (see Levy and Murnane 1992).

We are not averse to borrowing this economic account and using it to explain trends in within-occupation inequality (WO). It is possible, however, that much of what appears to be growing demand for high-skill labor within the residual statistical groups of Mincer-type regression models is just growing demand for particular high-skill occupations, not growing demand for high-skill tasks within all occupations. If so, we need only invoke a story about increasing demand for high-skill occupations, a story already outlined in the preceding section on the BO/WC trend. Once the BO/WC inequality is stripped out of the residual, it may well be that the remaining WO component suggests much less in the way of growing demand for high-skill labor.

The WO component is not, however, likely to be entirely stable once the BO/WC component is parsed out. To the contrary, it is plausible that complementary increases in the WO component will be generated by rising demand for unobservable skill, just as the standard economic account would have it. This type of account is best suited to explaining a WO takeoff in those sectors of the labor market, such as the nonmanual sector, in which the demand for skill has substantially increased. The introduction of computers and related technical innovations has presumably transformed the nonmanual workplace in especially revolutionary (and skill-demanding) ways. At the same time, the demand for interactive or “soft” skills may also have grown among service and manual laborers (Howell and Wolff 1991; Fernandez 2001), but we

doubt that this soft-skill demand shift in the manual sector is as pronounced as the computer-based demand shift in the nonmanual sector.

Data and Methods

We analyze earnings data collected in the May supplement of the Current Population Surveys (CPS) from 1973 to 1978 and in the “outgoing rotation group” (ORG) supplements of the monthly CPS from 1979 to 2004 (Bureau of Labor Statistics [BLS], years vary). These surveys are nationally representative samples of households that contain information on the usual pay of the main job that all household members held in the week prior to the survey.

Data Processing

In preparing the May/ORG files for the analysis, we were required to make a host of technical and conceptual decisions that affect the measurement of inequality and may affect observed trends in inequality as well. We follow best practice in the earnings literature whenever a consensus on best practice has been reached and otherwise proceed by carrying out analyses with two or more operational alternatives. Because our main conclusions were usually unaffected by such operational decisions, we report but one set of results in this paper (see Weeden et al. 2007 for details).

Throughout the analyses presented here, earnings are measured as hourly wages. In the May/ORG CPS, workers who indicate that they are paid on an hourly basis are asked to report their usual hourly wages, and workers who indicate some other pay periodicity report their usual weekly wages. We follow here conventional practice of calculating hourly wages for nonhourly workers by dividing weekly wages by the number of hours usually worked at the main job. We then convert hourly wages into constant 2000 dollars using the personal consumption

expenditures index from NIPA (Bureau of Economic Analysis, 2005). When wages have been truncated by the BLS to maintain confidentiality (i.e., “topcoding”), we have applied the conventional multiplier of 1.4 (Lemieux 2003; Card and DiNardo 2002). We rely on unedited measures of earnings because, unlike the edited earnings, these are available in the May 1973–1978 surveys and are unaffected by changes in the mid–1990s to the BLS procedures for allocating earnings (see Hirsch and Schumacher 2003; Lemieux 2003). This decision forces us to exclude the 1994–1995 ORG files in which workers with edited earnings are impossible to identify (see Lemieux 2003, pp. 6–7).

We include all wage and salary workers between the ages of 16 and 65 who report positive hours worked at their main jobs and for whom a valid occupation code is available. Following conventional practice, we exclude workers whose hourly wages fall below \$1 or above \$100 in 1979 dollars or who are self-employed, the latter because they are not asked the earnings questions in the ORG supplements (Angrist and Krueger 2000; Card and DiNardo 2002; Mishel, Bernstein, and Schmitt 2001; cf. Bollinger and Chandra 2005). We do not, however, exclude part-time workers, as analysts of wage inequality often do. Instead, we weight workers by the number of hours they usually work, thereby obtaining a wage distribution representative of all hours worked in the economy (see e.g., DiNardo, Fortin, and Lemieux 1996; Lemieux 2003; Card and DiNardo 2002). This practice may be understood as a compromise between excluding part-time workers and simply ignoring the distinction between part-time and full-time work.

After imposing all restrictions, our May surveys contain on average 20,254 men and 15,776 women per year, and our ORG surveys contain on average 64,048 men and 62,181 women per year. The final sample pooled over years includes 1,618,174 men and 1,587,005 women.

Occupation and Class Schemes

For our reported analyses, we rely on Standard Occupation Classification (SOC) codes, thus allowing the scheme to change with the introduction of new versions of the SOC in the 1983, 1992, and 2003 surveys. We have also examined trends with a uniform scheme created by back-coding all CPS data between 1983 and 2004 into the 1970 SOC codes (see Weeden et al. 2007; also, Weeden 2005a, 2005b; also Weeden and Grusky 2005a, 2005b, 2005c). This back-coding procedure uses a set of sex-specific weights, calculated from double-coded data sets, to assign earlier SOC codes to data originally arrayed in a later SOC scheme.

Which of these two occupation schemes is to be preferred? With the back-coded scheme, the occupational structure is treated as frozen in time as of 1970, and the rise of new occupations and related changes in occupational boundaries are thus ignored. The indigenous occupation scheme, by contrast, provides the best available characterization of the occupational structure in a given year and presumably delivers the most accurate estimate of absolute levels of within-occupation inequality. However, where the goal is to obtain estimates of *trend* in the various components of inequality, one should not necessarily default to this form of accuracy. If each new occupational classification scheme comes on the heels of real change in institutionalized boundaries of the occupational structure, analysis of the data arrayed in the indigenous scheme will indeed offer a better assessment of trends in micro-class inequality. The 1970-basis scheme will overestimate the increase in within-occupation inequality and, conversely, underestimate the increase in between-occupation inequality, at least if the earnings profiles of newly formed occupations diverge from those of their “parent” occupations. If, on the other hand, each new occupation scheme merely improves the measurement of an underlying

ing occupational structure that has remained largely constant, the indigenous occupation scheme will overestimate the increase in between-occupation inequality.

It is likely that both accounts are partly true. That is, some of the changes in a new occupation scheme may capture newly institutionalized occupational boundaries (e.g., systems analysts), whereas others may capture occupational distinctions that existed in the past but had simply been ignored by the classifiers (e.g., distinctions among managers). Because there is no solution to this dilemma, we have estimated trends in the various components of inequality with both indigenous and 1970-basis schemes, allowing us to gauge the sensitivity of our results to our measurement decisions. We report only the indigenous-scheme decompositions here in the interest of brevity.

We have carried out analyses with two conventional big-class schemes: the Erikson-Goldthorpe (EG) scheme (Erikson and Goldthorpe, Ch. 47), and the Featherman-Hauser (FH) scheme (Featherman and Hauser 1978). In the following graphs and tables, we feature the FH scheme because our prior research suggests that it better captures the heterogeneity in many variables, including income (Weeden and Grusky 2005a, 2005b, 2005c). We have, however, replicated all of our analyses with the EG scheme and secured very similar results.

The FH scheme, which is based on aggregate Census Bureau categories, is more detailed than the EG scheme and is devised specifically for the U.S. occupational structure. In its commonly applied version, the FH scheme consists of twelve classes, from which we have excised two (self-employed professionals, self-employed managers). The remaining ten FH categories are employed professionals, employed managers, sales workers, clerical workers, craft workers, operatives, service workers, laborers, farmers, and farm laborers. We refer the reader to Weeden et al. (2007) for a discussion of the

algorithm that was used to translate SOC codes into the FH scheme.

Methods

We carried out all analyses with two measures of inequality, the variance of $\log(\text{wages})$ and Theil's index, both of which can readily be decomposed into BC, BO/WC, and WO components. However, we only present decompositions based on the variance of $\log(\text{wages})$, given that the analogous decompositions using the Theil index yielded similar conclusions. We present separate decompositions of inequality for men and women throughout.

The total variance in $\log(\text{wages})$ can be decomposed by fitting two models. The first model regresses $\log(\text{wages})$ on the categories of a big-class scheme, and the second model regresses $\log(\text{wages})$ on the categories of an occupation scheme, where occupations are nested in big classes. These two models allow us to implement our desired three-way decomposition: The BC component is the total variance in $\log(\text{wages})$ minus the variance of the residuals of the big-class model; the BO/WC component is the difference between the variance of the residuals of the occupation and big-class models; and the WO component is the variance of the residuals from the occupation model.

Results

The first set of three-way decompositions is presented in Figure 3 (for men) and Figure 4 (for women). Because the relative rate of change can be difficult to gauge via inspection alone, we report in Table 1 the estimates secured by imposing, for men and women separately, a linear trend on the variance in $\log(\text{wages})$. The first model in Table 1 fits a single slope and intercept to the trend lines, while the second model fits a single slope but allows the intercept to shift with each change in the occupational classification scheme (in 1983, 1992, and 2003). The simpler model

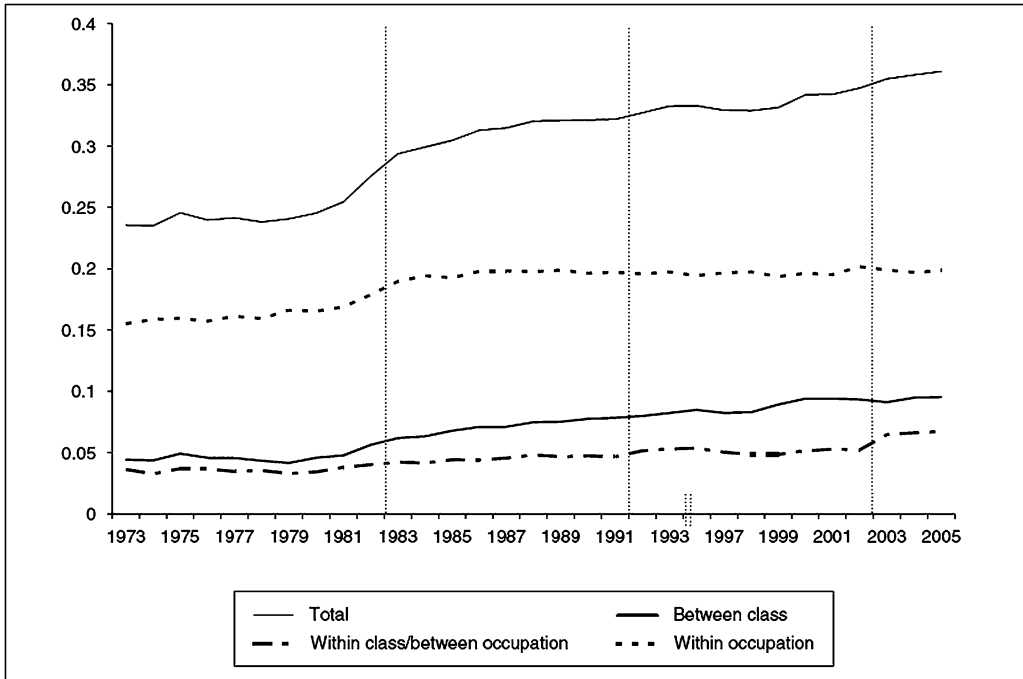


Figure 3. Variance-Based Decomposition of Men's Wage Inequality into Between-Class, Within-Class/Between-Occupation, and Within-Occupation Components, 1973 to 2005

Note: Vertical lines indicate new Standard Occupation Classification scheme; hash marks indicate a break in the time series.

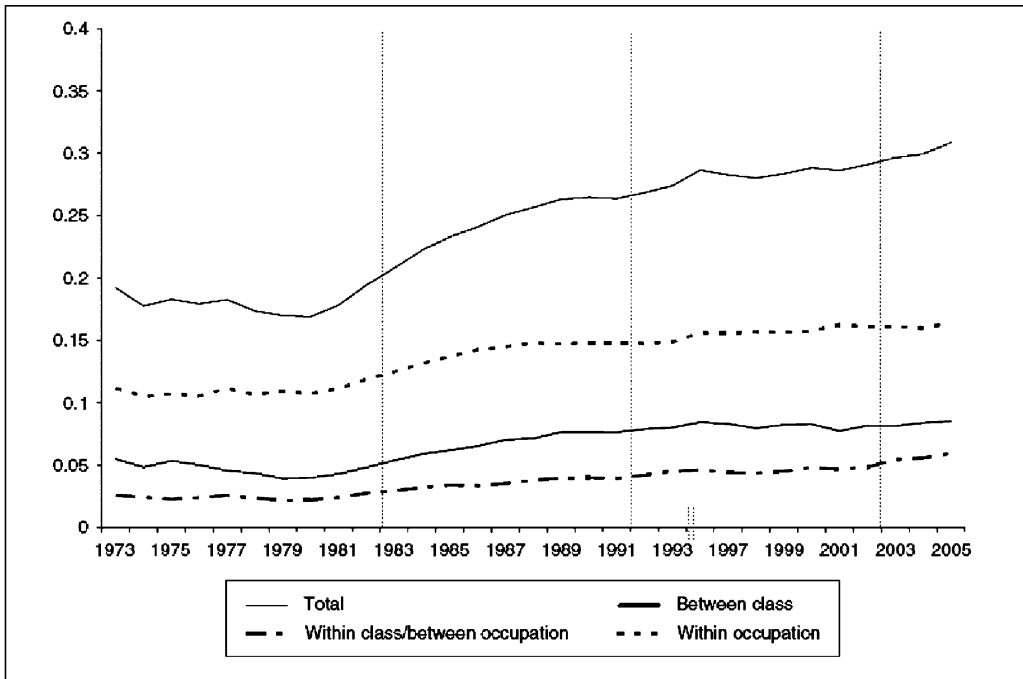


Figure 4. Variance-Based Decomposition of Women's Wage Inequality into Between-Class, Within-Class/Between-Occupation, and Within-Occupation Components, 1973 to 2005

Note: Vertical lines indicate new Standard Occupation Classification scheme; hash marks indicate a break in the time series.

Table 1. Estimated Linear Effect of Time on the Components of Wage Inequality, by Sex

| | Intercept | Slope Coefficient | Average Percentage Change per Year | Percentage of Total Change |
|--|-------------------|-------------------|------------------------------------|----------------------------|
| A: Model 1: Changing occupational structure | | | | |
| Men | | | | |
| Total | 0.2374 (0.005) | 0.0041 (0.000) | 1.7 | |
| Between class | 0.0411 (0.002) | 0.0019 (0.000) | 4.5 | 45.0 |
| Between occupation/ within class | 0.0323 (0.002) | 0.0009 (0.000) | 2.7 | 21.4 |
| Within occupation | 0.1640 (0.003) | 0.0014 (0.000) | 0.8 | 33.6 |
| Women | | | | |
| Total | 0.1684 (0.005) | 0.0046 (0.000) | 2.7 | |
| Between class | 0.0436 (0.002) | 0.0015 (0.000) | 3.3 | 31.6 |
| Between occupation/ within class | 0.0198 (0.001) | 0.0011 (0.000) | 5.4 | 23.5 |
| Within occupation | 0.1050 (0.002) | 0.0021 (0.000) | 2.0 | 44.7 |
| B: Model 2: Constant occupational structure | | | | |
| Men | | | | |
| Total | 0.2335 (0.003) | 0.0026 (0.000) | 1.1 | |
| Between class | 0.0402 (0.001) | 0.0014 (0.000) | 3.4 | 51.7 |
| Between occupation/ within class | 0.0346 (0.001) | 0.0003 (0.000) | 0.8 | 10.3 |
| Within occupation | 0.1586 (0.002) | 0.0010 (0.000) | 0.6 | 37.5 |
| Women | | | | |
| Total | 0.1691 (0.005) | 0.0024 (0.001) | 1.4 | |
| Between class | 0.0449 (0.002) | 0.0004 (0.000) | 0.8 | 15.0 |
| Between occupation/ within class | 0.0217 (0.001) | 0.0005 (0.000) | 2.4 | 21.7 |
| Within occupation | 0.1025 (0.002) | 0.0015 (0.000) | 1.5 | 62.9 |

Note: Numbers in parentheses are standard errors. Model 1 fits a simple linear trend on the variance in logged wages. Model 2 fits a uniform linear trend with intercept adjustments at each change in the occupational classification scheme.

(i.e., Model 1) corresponds to the assumption that new classification schemes are introduced in response to real changes in institutionalized occupational boundaries (i.e., “changing occupational structure”). If such changes are indeed real, then any increase in the explanatory power of these more refined classifications bespeaks a meaningful change in how inequality is apportioned within and between occupations. By contrast, the second model corresponds to the assumption that new classification schemes capture occupa-

tional distinctions that were always present, but were ignored in the more primitive classifications of prior years (i.e., “constant occupational structure”). It is appropriate under this assumption to parse out the classification effects and then fit a pooled slope coefficient.

The results of Table 1 pertain to absolute changes alone. As we have noted, the relative size of the BC, BO/WC, and WO components is also of interest, implying that one should condition on the overall growth in inequality and ask how it was distributed

Table 2. Predicted Change in Proportion of Total Variance Attributable to the Components of Inequality, 1973 to 2005

| | Model 1 | | Model 2 | |
|---------------------------------|---------|-------|---------|-------|
| | 1973 | 2005 | 1973 | 2005 |
| Men | | | | |
| Total variance | 0.237 | 0.369 | 0.233 | 0.317 |
| Percentage of total | | | | |
| Between class | 17.3 | 27.2 | 17.2 | 26.3 |
| Between occupation/within class | 13.6 | 16.4 | 14.8 | 13.6 |
| Within occupation | 69.1 | 56.4 | 67.9 | 60.0 |
| Women | | | | |
| Total variance | 0.168 | 0.315 | 0.169 | 0.246 |
| Percentage of total | | | | |
| Between class | 25.9 | 28.6 | 26.5 | 22.9 |
| Between occupation/within class | 11.8 | 17.2 | 12.8 | 15.6 |
| Within occupation | 62.3 | 54.2 | 60.6 | 61.5 |

Note: Model 1 fits a simple linear trend on the variance in logged wages. Model 2 fits a uniform linear trend with intercept adjustments at each change in the occupational classification scheme.

across the three components. We thus present in Table 2 measures of relative change under the models of Table 1. These measures, calculated for 1973 and 2005, are simply the percentage of the total variance in $\log(\text{wages})$ that each component represents.

We can turn now to addressing the main questions at hand. The BC trend line, which we consider first, is distinguished by a steep takeoff in the 1980s and steady increase thereafter. The between-class component accounts for 45.0–51.7 percent of the total change in male inequality (Table 1), and the between-class share of the total male variance increased from 17.2–17.3 percent in 1973 to 26.3–27.2 percent in 2005 (Table 2). For women, the pooled BC trend line is less steep, and the BC component accordingly accounts for less of the total change in inequality (see Table 1). This apparent gender difference should not be overinterpreted. Because the BC series for women declines steeply in the 1970s, the pooled coefficient reported in Table 2 averages across the decline and the post-1982 takeoff, yielding an attenuated

estimate of the latter. If one instead fits a linear trend to the post-1982 data, the male and female series uniformly show a rapid increase in the BC component. This growth in between-class inequality is hardly surprising. Insofar as big classes are merely education groups in disguise, the well-known increase in the payoff to schooling should be expressed as a takeoff in the BC trend line over the last two decades, just as we find here.

We consider next the trend in WC/BO inequality. In this case, our results cannot be straightforwardly read off from prior research, because big classes and detailed occupations have not before been analyzed in tandem. For both women and men, we find that the WC/BO trend line does not show the characteristic steep takeoff in the early 1980s, nor does it move sharply upward at any point thereafter. Except for a classification-induced shift in 2003, the increase in WC/BO inequality is slow and gradual, and it translates into a slightly higher share of the total variance by 2005, at least under Model 1 (see Table 2). The WC/BO component accounts

for 10.8–16.5 percent of the total growth in male inequality and for 19.1–19.9 percent of the total growth in female inequality (see Table 1). We thus find that the WC/BO component, like the BC component, is increasing both in absolute and relative terms, although the rate of increase for the WC/BO component is less dramatic than the corresponding rate for the BC component.

We had predicted that the upward trend in the between-occupation component would be especially pronounced in the nonmanual classes because they contain precisely those high-skill occupations (e.g., computer programmer) that might profit from demand explosions. Furthermore, the nonmanual classes are the home ground of closure movements that allow high-skill occupations to control supply and demand, thus generating rent and opening up an earnings gap relative to occupations that have not successfully occupationalized. Although we will not provide a direct test of this hypothesis here, we have elsewhere (Weeden et al. 2007) examined whether the between-occupation trends are playing out differently in different big classes. For women and men alike, we found that the nonmanual trend lines tend to move upward at a faster pace than the manual trend lines, the latter being mired for the most part at very low levels over the last thirty years. The most dramatic takeoff in between-occupation variance was found in the nonmanual class of sales workers. When we examined the particular occupations within this class that are rising or falling in earnings, we found some evidence of high-skill winners (e.g., sales engineers, stock and bond salespeople, advertising agents) and low-skill losers (e.g., retail sales clerks, cashiers, sales representatives), just as a demand-shift hypothesis would have it.

The third series in our analysis, the WO series, is distinguished by important gender differences in the absolute trend. Among male workers, the WO component is increasing, but fails to match the fast-paced

growth of the BC component or even the WC/BO component. Indeed, the WO series suggests a one-time takeoff in the early 1980s, followed by near stasis thereafter. The relatively slow pace of change means that only a minority share of the total change in male inequality is attributable to a growth in WO (see Table 1). This result is inconsistent with the long-standing claim that rising “within-group” inequality is the driving force behind the overall growth of inequality (cf. Lemieux 2003). It does, however, lend support to our alternative suggestion that much of what conventional Mincer-type formulations represent as growth in within-group inequality is in fact growth in the WC/BO component.

The relatively slow pace of change in the WO component for males translates, furthermore, into a decline in the percentage of the total variance that is generated within occupations. It may be recalled in this context that the BC and WC/BO components both increased in absolute as well as relative terms. Clearly, these two components had to be taking relative share from the WO component, and Table 2 reveals precisely this relationship. The within-occupation share declines from 69.1 percent to 56.4 percent when Model 1 estimates are used.

The female WO series, by contrast, shows clear evidence of ongoing increase after the early 1980s, with a steeper overall slope coefficient as a result. In fact, the WO component accounts for more of the total change in variance than the BC component, especially under the assumption of a constant occupational structure (see Model 2, Table 1). There are a host of potential explanations for the gender disparity in the WO series. The task of adjudicating among them is well beyond the scope of this paper, but it is worth noting that the disparity is potentially consistent with an account emphasizing skill-biased technological change. That is, many female-typed occupations (e.g., secretary) have developed an especially strong

reliance on computer, interactive, and other unobservable skills, whereas male-typed clerical occupations (e.g., mail sorter) have not been as frequently affected by skill-demanding technological change (Welch 2000; Weinberg 2000). The rapid rise in within-occupation inequality among women may arise from such new unobserved skills opening up new inequalities.

Although the trend in the WO component is steeper for women than for men, it is not so much steeper as to require a gender-specific account of trends in the WO share of inequality. Under Model 1, the WO share of the total variance declines from 62.3 percent to 54.2 percent for women, a decline that is only slightly less dramatic than that for men. By contrast, the WO share of inequality increases under Model 2, but this result is an artifact of gender differences in the pre-1982 trend. If only post-1982 data are analyzed, the Model 2 estimates show that the WO share of total inequality is declining for women as well as for men.

We argued in our introductory comments that rising demand for unobservable skills should play out principally in the nonmanual sector, where the workplace has been transformed dramatically by the introduction of computers, more advanced telecommunications, and related technical innovations. For reasons of space, we won't present the requisite decompositions here, but we have shown elsewhere that, as anticipated, the WO trends lines are increasing in the nonmanual classes yet stable in the manual classes (see Weeden et al. 2007 for details).

Discussion

We conclude by asking how our results speak to the relationship between class and earnings. Has the ongoing increase in earnings inequality played out in ways that support the rise of big classes? Does the takeoff in inequality instead support the rise of micro-classes? Are both types of class principles si-

multaneously supported? Or is earnings inequality becoming an increasingly individualistic affair that occurs mainly within groups rather than between them? By posing questions of this sort, we have sought to move away from the field's long-standing focus on the *causes* of earnings inequality, turning instead to questions about its *effects* on the structure of social classes.

We begin by assessing whether the foregoing results are consistent with big-class formulations (see Figure 1). Although there is much here that big-class enthusiasts can embrace, our data are not working exclusively or unambiguously in support of the big-class principle. The core results to which big-class adherents will inevitably point are the growing absolute size of the between-class component and the growing between-class share of the total variance. Important though these results are, it must also be noted that interoccupational cleavages within some big classes are growing as well, a result that is inconsistent with simple big-class stories. The double-edged sword of rising earnings inequality appears to be *simultaneous* growth in distinctions between classes and within them. These within-class cleavages have become so prominent in some classes, such as the FH service class, that it has to be asked whether this is still a single class, if ever it was.

How, then, does the micro-class story fare? Here again, one can find results that are consistent with a micro-class account, especially the steady growth in within-class occupational distinctions. Moreover, the within-occupation component to inequality is declining in relative size for both men and women, at least after 1982. We can conclude that the class principle, as expressed via big classes and micro-classes together, accounts for a growing share of the total variance in earnings, while the residual of "individualized inequality" is declining in relative size. The earnings distribution thus suggests a new type of dual closure in which two types of nested social groupings, big-classes and

detailed occupations, figure increasingly prominently.

Although the dual-closure account is a descriptively accurate account of recent trends, it would be heroic to suppose that inequality in the future will necessarily take on an increasingly simple dual-closure form. This is because, without understanding the mechanisms behind recent trends, a simple-minded extrapolation is obviously too bold. There are two questions about mechanisms that are especially important in this regard. The first pertains to the mechanisms underlying the trend in big-class formation, and the second pertains to the mechanisms underlying the trend in micro-class formation. We review each of these two questions below.

On the big-class side of the equation, one has to ask whether the rising payoff to schooling is indeed the main mechanism increasing BC inequality. If it is, big-classes are reduced to indirect beneficiaries of trends that ultimately support the development of “education classes” (e.g., Meyer, Ch. 105; Brooks, Ch. 33). The key question in this regard, and one which remains unanswered by our analyses, is whether the between-class component is increasing even in the presence of controls for education. For purposes of illustration, suppose that there are just two big classes (manual and nonmanual) and two education groups (educated and uneducated), as indicated in Figures 5a and 5b. We can then distinguish between (1) pro-education change in which the net education cleavage is widening and the net class cleavage is not (i.e., Figure 5a), and (2) pro-class change in which the net class cleavage is widening and the net education cleavage is not (i.e., Figure 5b). If the growth in between-class inequality is driven exclusively by the rising payoff to education, then education groups rather than big classes are arguably emerging as the principal form of extra-individual structure in the contemporary labor market.

The dual-closure account is also potentially vulnerable on the micro-class front.

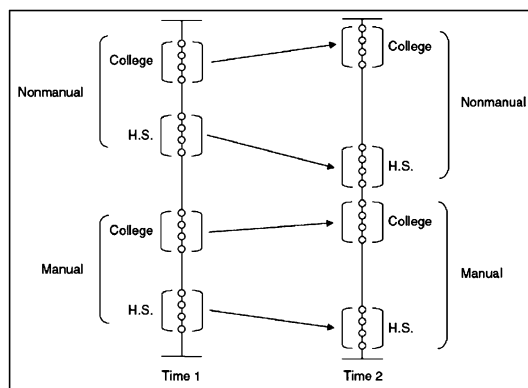


Figure 5a. The Transition to an Education Regime

Note: H.S.=high school

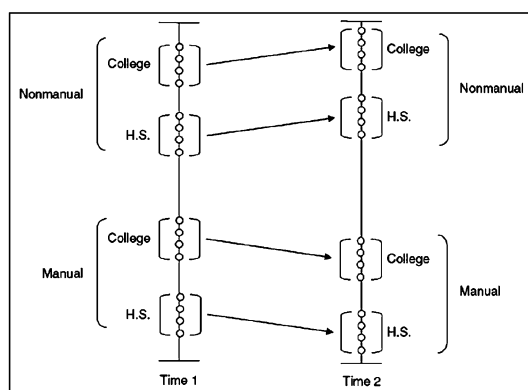


Figure 5b. The Transition to a Big-Class Regime

Note: H.S.=high school

Here, the skeptic might point out that the relative decline in within-occupation variance is playing out mainly in the manual sector, not within the ever-growing ranks of the nonmanual sector (see Weeden et al. 2007). It follows that, as the manual sector continues to shrink, the forces that expand WO inequality (e.g., increasing returns to unobservable skills) may come to dominate those that contract such inequality. By this account, WO inequality in nonmanual occupations will eventually control trends, and the overall decline in WO inequality reported here will therefore soon stall.

The foregoing skeptic’s account treats the steep rise of WO inequality within the nonmanual sector at face value. Might we instead interpret this steep rise as an artifact of

increasingly error-ridden occupational classifications? The occupational structure within the nonmanual sector is so dynamic that there is inevitable ambiguity in distinguishing true increases in within-occupation inequality from those generated by the emergence of new suboccupational distinctions. When, for example, the earnings of cosmetic dentists grow spectacularly while those of other dentists do not, does this bespeak a true increase in within-occupation inequality? Or does it instead reveal incipient suboccupational distinctions that might be formalized in separate occupational associations with differentiated training and credentialing? In practice, we have resolved this question by conditioning on formal occupational classifications, thus allowing for new occupations only when the SOC recognizes them. We suspect, however, that the SOC does not always keep pace with the dynamism of the nonmanual sector, meaning that some of the apparent takeoff in WO inequality within this sector is just classification error. If this interpretation is on the mark, the rise of group-based inequality is even more extreme than our results imply.

It bears reiterating that these ambiguities are especially consequential when one attempts to predict whether inequality in the future will take on increasingly individualized or class forms. If one instead settles for a purely descriptive account of recent trends, the ambiguity fades somewhat and the merit of a class account of inequality is less contestable. Although the majority of earnings inequality is still generated within occupations, the relatively rapid growth of the between-class and between-occupation components implies that the labor market is not only becoming more unequal but lumpier as well.

NOTES

1. The increase in the BC component under this scenario is a compensating form of pro-class change.

REFERENCES

- Abbott, A. (1988). *The system of professions: An essay on the division of expert labor*. Chicago: University of Chicago Press.
- Acemoglu, D. (2002). "Technical change, inequality, and the labor market." *Journal of Economic Literature*, 40, 7–72.
- Allison, P. D. (1978). "Measures of inequality." *American Sociological Review*, 43, 865–80.
- Angrist, J. D., and A. B. Krueger. (2000). "Empirical strategies in labor economics." In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics* (vol. 3A). Amsterdam: Elsevier Science.
- Beck, U. (1992). *Risk society: Towards a new modernity*. M. Ritter (Trans.). London: Sage.
- Blau, F. D., and L. M. Kahn. (2002). *At home and abroad: U.S. labor market performance in comparative perspective*. New York: Russell Sage Foundation.
- Bollinger, C. R., and A. Chandra. (2005). "Iatrogenic specification error: A cautionary tale of data cleaning." *Journal of Labor Economics*, 23, 235–57.
- Bourguignon, F. (2006). "Expanding the Income-Poverty Paradigm." In D. B. Grusky and R. Kanbur (Eds.), *Poverty and Inequality*. Stanford, CA: Stanford University Press.
- Bureau of Economic Analysis (2005). "National income and products accounts, personal consumption expenditures." Retrieved November 23, 2005, from <http://www.bea.gov/bea/dn/nipaweb/Index.asp>.
- Bureau of Labor Statistics. (Years vary.) *May and ORG Supplements to the Current Population Survey*. Available in machine-readable form from National Bureau of Economic Research.
- Card, D., and J. E. DiNardo. (2002). "Skill-biased technological change and rising wage inequality: Some problems and puzzles." *Journal of Labor Economics*, 20(4), 733–83.
- DiNardo, J., N. M. Fortin, and T. Lemieux. (1996). "Labor market institutions and the distribution of wages, 1973–1992: A semi-parametric approach." *Econometrica*, 65, 1001–46.
- Featherman, D. L., and R. M. Hauser. (1978). *Opportunity and change*. New York: Academic.
- Fernandez, R. (2001). "Skill-biased technological change and wage inequality: Evidence from a plant retooling." *American Journal of Sociology*, 107, 273–320.
- Frank, R. H., and P. J. Cook. (1995). *The winner-take-all society*. New York: Free Press.
- Groshen, E. (1991). "Sources of intra-industry wage dispersion: How much do employers matter?" *Quarterly Journal of Economics*, 106, 869–84.
- Grusky, D. B. (in collaboration with Gabriela Galescu). 2005a. "Foundations of class analysis: A

- Durkheimian perspective." In E. O. Wright (Ed.), *Approaches to class analysis*. Cambridge: Cambridge University Press.
- Grusky, D. B. (2005b). "Is Durkheim a class analyst?" In J. Alexander and P. Smith (Eds.), *The Cambridge companion to Durkheim*. Cambridge: Cambridge University Press.
- Grusky, D. B., and K. Weeden. (2007). "Measuring poverty: The case for a sociological approach." Ch. 2 in N. Kakwani and J. Silber (Eds.), *The many dimensions of poverty*. Hampshire, UK: Palgrave MacMillan.
- Hirsch, B., and E. Schumacher. (2004). "Match bias in wage gap estimates due to earnings imputation." *Journal of Labor Economics*, 22, 689–722.
- Hout, M., C. Brooks, and J. Manza. (1993). "Persistence of classes in postindustrial societies." *International Sociology*, 8, 259–78.
- Howell, D. R., and E. N. Wolff. (1991). "Trends in the growth and distribution of skills in the U.S. workplace, 1960–1985." *Industrial and Labor Relations Review*, 44, 486–502.
- Juhn, C., K. M. Murphy, and B. Pierce. (1993). "Wage inequality and the rise in returns to skill." *Journal of Political Economy*, 101, 410–42.
- Katz, L. F., and D. H. Autor. (1999). "Changes in the wage structure and earnings inequality." In O. C. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics* (vol. 3). Amsterdam: Elsevier Science.
- Katz, L. F., and K. Murphy. (1992). "Changes in relative wages, 1963–1987: Supply and demand factors." *Quarterly Journal of Economics*, 107, 35–78.
- Kim, C., and A. Sakamoto. (2005). "Bringing inequality back in: The role of occupational structure in growing wage dispersion." Paper presented at the RC 28 Meetings of the International Sociological Association, Los Angeles, August 2005.
- Lemieux, T. (2003). "Residual wage inequality: A re-examination." Unpublished manuscript. University of British Columbia.
- Levy, F. (1998). *The new dollars and dreams: American incomes and economic change*. New York: Russell Sage.
- Levy, F., and R. J. Murnane. (1992). "U.S. earnings levels and earnings inequality: A review of recent trends and proposed explanations." *Journal of Economic Literature*, 30, 1333–81.
- Lockwood, D. (1989). *The black coated worker: A study in class consciousness* (2nd ed.). Oxford: Clarendon Press.
- Mishel, L., J. Bernstein, and J. Schmitt. (2001). *State of working America: 2000/2001*. Ithaca, NY: Cornell University Press.
- Murphy, K., and F. Welch. (1993). "Occupational change and the demand for skill." *American Economic Review*, 83, 122–26.
- Sen, A. (1973). *On economic inequality*. Oxford: Oxford University Press.
- Stone, K. (2004). *From widgets to digits: Employment regulation for the changing workplace*. Cambridge, UK: Cambridge University Press.
- Treiman, D. J. (1977). *Occupational prestige in comparative perspective*. New York: Academic Press.
- Weeden, K. A. (2002). "Why do some occupations pay more than others? Social closure and earnings inequality in the United States." *American Journal of Sociology*, 108, 55–101.
- _____. (2005a). "Algorithm for translating 1980 and 1990 U.S. census occupation codes into 1970 codes." Unpublished manuscript. Department of Sociology, Cornell University.
- _____. (2005b). "Algorithm for backcoding 2000 census occupation codes into 1990 census occupation codes." Unpublished manuscript. Department of Sociology, Cornell University.
- Weeden, K. A., and D. B. Grusky. (2005a). "The case for a new class map." *American Journal of Sociology*, 111, 141–212.
- _____. (2005b). "Is inequality becoming less organized?" Unpublished manuscript. Department of Sociology, Cornell University.
- _____. (2005c). "Are there any big classes at all?" In D. Bills (Ed.), *The shape of social inequality: stratification and ethnicity in comparative perspective. Research in Social Stratification and Mobility* (vol. 22). Amsterdam: Elsevier.
- Weeden, K. A., Y. Kim, M. DiCarlo, and D. B. Grusky. (2007). "Social class and earnings inequality." *American Behavioral Scientist*, 50, 702–36.
- Weinberg, B. A. (2000). "Computer use and the demand for female workers." *Industrial and Labor Relations Review*, 53, 290–308.
- Welch, F. (2000). "Growth in women's relative wages and inequality among men: one phenomenon or two?" *American Economic Review*, 90, 444–9.
- Wilensky, H. L. (1966). "The professionalization of everyone?" *American Journal of Sociology*, 70, 137–58.

PART IV

Inequality at the Extremes

The Ruling Class, Elites, and the Upper Class

■ CLASSIC STATEMENTS ■

28 ■ Gaetano Mosca

The Ruling Class

1. Among the constant facts and tendencies that are to be found in all political organisms, one is so obvious that it is apparent to the most casual eye. In all societies—from societies that are very meagerly developed and have barely attained the dawns of civilization, down to the most advanced and powerful societies—two classes of people appear—a class that rules and a class that is ruled. The first class, always the less numerous, performs all political functions, monopolizes power and enjoys the advantages that power brings, whereas the second, the more numerous class, is directed and controlled by the first, in a manner that is now more or less legal, now more or less arbitrary and violent, and supplies the first, in appearance at least, with material means of subsistence and with the instrumentalities that are essential to the vitality of the political organism.

In practical life we all recognize the existence of this ruling class (or political class, as we have elsewhere chosen to define it).¹ We all know that, in our own country, whichever it may be, the management of public affairs is in the hands of a minority of influential persons, to which management, willingly or unwillingly, the majority defer. We know that

the same thing goes on in neighboring countries, and in fact we should be put to it to conceive of a real world otherwise organized—a world in which all men would be directly subject to a single person without relationships of superiority or subordination, or in which all men would share equally in the direction of political affairs. If we reason otherwise in theory, that is due partly to inveterate habits that we follow in our thinking and partly to the exaggerated importance that we attach to two political facts that loom far larger in appearance than they are in reality.

The first of these facts—and one has only to open one's eyes to see it—is that in every political organism there is one individual who is chief among the leaders of the ruling class as a whole and stands, as we say, at the helm of the state. That person is not always the person who holds supreme power according to law. At times, alongside of the hereditary king or emperor there is a prime minister or a major-domo who wields an actual power that is greater than the sovereign's. At other times, in place of the elected president the influential politician who has procured the president's election will govern. Under

Gaetano Mosca. "The Ruling Class," in Arthur Livingston, ed., *The Ruling Class*, translated by Hannah D. Kahn (New York; McGraw Hill, 1939), pp. 50–54, 56–62, 65–66. Copyright © by Elizabeth Abbot.

special circumstances there may be, instead of a single person, two or three who discharge the functions of supreme control.

The second fact, too, is readily discernible. Whatever the type of political organization, pressures arising from the discontent of the masses who are governed, from the passions by which they are swayed, exert a certain amount of influence on the policies of the ruling, the political, class.

But the man who is at the head of the state would certainly not be able to govern without the support of a numerous class to enforce respect for his orders and to have them carried out; and granting that he can make one individual, or indeed many individuals, in the ruling class feel the weight of his power, he certainly cannot be at odds with the class as a whole or do away with it. Even if that were possible, he would at once be forced to create another class, without the support of which action on his part would be completely paralyzed. On the other hand, granting that the discontent of the masses might succeed in deposing a ruling class, inevitably, as we shall later show, there would have to be another organized minority within the masses themselves to discharge the functions of a ruling class. Otherwise all organization, and the whole social structure, would be destroyed.

2. From the point of view of scientific research the real superiority of the concept of the ruling, or political, class lies in the fact that the varying structure of ruling classes has a preponderant importance in determining the political type, and also the level of civilization, of the different peoples. According to a manner of classifying forms of government that is still in vogue, Turkey and Russia were both, up to a few years ago, absolute monarchies, England and Italy were constitutional, or limited, monarchies, and France and the United States were classed as republics. The classification was based on the fact that, in the first two countries mentioned, headship in the state was hereditary and the chief was nominally omnipotent; in the second two, his office is

hereditary but his powers and prerogatives are limited; in the last two, he is elected.

That classification is obviously superficial. Absolutisms though they were, there was little in common between the manners in which Russia and Turkey were managed politically, the levels of civilization in the two countries and the organization of their ruling classes being vastly different. On the same basis, the regime in Italy, a monarchy, is much more similar to the regime in France, a republic, than it is to the regime in England, also a monarchy; and there are important differences between the political organizations of the United States and France, though both countries are republics.

As we have already suggested, ingrained habits of thinking have long stood, as they still stand, in the way of scientific progress in this matter. The classification mentioned above, which divides governments into absolute monarchies, limited monarchies and republics, was devised by Montesquieu and was intended to replace the classical categories of Aristotle, who divided governments into monarchies, aristocracies and democracies. What Aristotle called a democracy was simply an aristocracy of fairly broad membership. Aristotle himself was in a position to observe that in every Greek state, whether aristocratic or democratic, there was always one person or more who had a preponderant influence. Between the day of Polybius and the day of Montesquieu, many writers perfected Aristotle's classification by introducing into it the concept of "mixed" governments. Later on the modern democratic theory, which had its source in Rousseau, took its stand upon the concept that the majority of the citizens in any state can participate, and in fact *ought* to participate, in its political life, and the doctrine of popular sovereignty still holds sway over many minds in spite of the fact that modern scholarship is making it increasingly clear that democratic, monarchical and aristocratic principles function side by side in every political organism. We shall not stop to refute this democratic theory here, since that is the

task of this work as a whole. Besides, it would be hard to destroy in a few pages a whole system of ideas that has become firmly rooted in the human mind. As Las Casas aptly wrote in his life of Christopher Columbus, it is often much harder to unlearn than to learn.

3. We think it may be desirable, nevertheless, to reply at this point to an objection which might very readily be made to our point of view. If it is easy to understand that a single individual cannot command a group without finding within the group a minority to support him, it is rather difficult to grant, as a constant and natural fact, that minorities rule majorities, rather than majorities minorities. But that is one of the points—so numerous in all the other sciences—where the first impression one has of things is contrary to what they are in reality. In reality the domination of an organized minority, obeying a single impulse, over the unorganized majority is inevitable. The power of any minority is irresistible as against each single individual in the majority, who stands alone before the totality of the organized minority. At the same time, the minority is organized for the very reason that it is a minority. A hundred men acting uniformly in concert, with a common understanding, will triumph over a thousand men who are not in accord and can therefore be dealt with one by one. Meanwhile it will be easier for the former to act in concert and have a mutual understanding simply because they are a hundred and not a thousand. It follows that the larger the political community, the smaller will the proportion of the governing minority to the governed majority be, and the more difficult will it be for the majority to organize for reaction against the minority.

However, in addition to the great advantage accruing to them from the fact of being organized, ruling minorities are usually so constituted that the individuals who make them up are distinguished from the mass of the governed by qualities that give them a certain material, intellectual or even moral superiority; or else they are the heirs of indi-

viduals who possessed such qualities. In other words, members of a ruling minority regularly have some attribute, real or apparent, which is highly esteemed and very influential in the society in which they live.

4. In primitive societies that are still in the early stages of organization, military valor is the quality that most readily opens access to the ruling, or political, class. In societies of advanced civilization, war is the exceptional condition. It may be regarded as virtually normal in societies that are in the initial stages of their development; and the individuals who show the greatest ability in war easily gain supremacy over their fellows—the bravest become chiefs. The fact is constant, but the forms it may assume, in one set of circumstances or another, vary considerably.

As a rule the dominance of a warrior class over a peaceful multitude is attributed to a superposition of races, to the conquest of a relatively unwarlike group by an aggressive one. Sometimes that is actually the case—we have examples in India after the Aryan invasions, in the Roman Empire after the Germanic invasions and in Mexico after the Aztec conquest. But more often, under certain social conditions, we note the rise of a warlike ruling class in places where there is absolutely no trace of a foreign conquest. As long as a horde lives exclusively by the chase, all individuals can easily become warriors. There will of course be leaders who will rule over the tribe, but we will not find a warrior class rising to exploit, and at the same time to protect, another class that is devoted to peaceful pursuits. As the tribe emerges from the hunting stage and enters the agricultural and pastoral stage, then, along with an enormous increase in population and a greater stability in the means of exerting social influence, a more or less clean-cut division into two classes will take place, one class being devoted exclusively to agriculture, the other class to war. In this event, it is inevitable that the warrior class should little by little acquire such ascendancy over the other as to be able to oppress it with impunity. . . .

5. Everywhere—in Russia and Poland, in India and medieval Europe—the ruling warrior classes acquire almost exclusive ownership of the land. Land, as we have seen, is the chief source of production and wealth in countries that are not very far advanced in civilization. But as civilization progresses, revenue from land increases proportionately. With the growth of population there is, at least in certain periods, an increase in rent, in the Ricardian sense of the term, largely because great centers of consumption arise—such at all times have been the great capitals and other large cities, ancient and modern. Eventually, if other circumstances permit, a very important social transformation occurs. Wealth rather than military valor comes to be the characteristic feature of the dominant class: the people who rule are the rich rather than the brave.

The condition that in the main is required for this transformation is that social organization shall have concentrated and become perfected to such an extent that the protection offered by public authority is considerably more effective than the protection offered by private force. In other words, private property must be so well protected by the practical and real efficacy of the laws as to render the power of the proprietor himself superfluous. This comes about through a series of gradual alterations in the social structure whereby a type of political organization, which we shall call the “feudal state,” is transformed into an essentially different type, which we shall term the “bureaucratic state.” We are to discuss these types at some length hereafter, but we may say at once that the evolution here referred to is as a rule greatly facilitated by progress in pacific manners and customs and by certain moral habits which societies contract as civilization advances.

Once this transformation has taken place, wealth produces political power just as political power has been producing wealth. In a society already somewhat mature—where, therefore, individual power is curbed by the collective power—if the powerful are as a

rule the rich, to be rich is to become powerful. And, in truth, when fighting with the mailed fist is prohibited whereas fighting with pounds and pence is sanctioned, the better posts are inevitably won by those who are better supplied with pounds and pence.

There are, to be sure, states of a very high level of civilization which in theory are organized on the basis of moral principles of such a character that they seem to preclude this overbearing assertiveness on the part of wealth. But this is a case—and there are many such—where theoretical principles can have no more than a limited application in real life. In the United States all powers flow directly or indirectly from popular elections, and suffrage is equal for all men and women in all the states of the Union. What is more, democracy prevails not only in institutions but to a certain extent also in morals. The rich ordinarily feel a certain aversion to entering public life, and the poor a certain aversion to choosing the rich for elective office. But that does not prevent a rich man from being more influential than a poor man, since he can use pressure upon the politicians who control public administration. It does not prevent elections from being carried on to the music of clinking dollars. It does not prevent whole legislatures and considerable numbers of national congressmen from feeling the influence of powerful corporations and great financiers.² . . .

6. In societies in which religious beliefs are strong and ministers of the faith form a special class a priestly aristocracy almost always arises and gains possession of a more or less important share of the wealth and the political power. Conspicuous examples of that situation would be ancient Egypt (during certain periods), Brahman India and medieval Europe. Oftentimes the priests not only perform religious functions. They possess legal and scientific knowledge and constitute the class of highest intellectual culture. Consciously or unconsciously, priestly hierarchies often show a tendency to monopolize learning and hamper the dissemination of the methods and procedures

that make the acquisition of knowledge possible and easy. To that tendency may have been due, in part at least, the painfully slow diffusion of the demotic alphabet in ancient Egypt, though that alphabet was infinitely more simple than the hieroglyphic script. The Druids in Gaul were acquainted with the Greek alphabet but would not permit their rich store of sacred literature to be written down, requiring their pupils to commit it to memory at the cost of untold effort. To the same outlook may be attributed the stubborn and frequent use of dead languages that we find in ancient Chaldea, in India, and in medieval Europe. Sometimes, as was the case in India, lower classes have been explicitly forbidden to acquire knowledge of sacred books.

Specialized knowledge and really scientific culture, purged of any sacred or religious aura, become important political forces only in a highly advanced stage of civilization, and only then do they give access to membership in the ruling class to those who possess them. But in this case too, it is not so much learning in itself that has political value as the practical applications that may be made of learning to the profit of the public or the state. Sometimes all that is required is mere possession of the mechanical processes that are indispensable to the acquisition of a higher culture. This may be due to the fact that on such a basis it is easier to ascertain and measure the skill which a candidate has been able to acquire—it is easier to “mark” or grade him. So in certain periods in ancient Egypt the profession of scribe was a road to public office and power, perhaps because to have learned the hieroglyphic script was proof of long and patient study. In modern China, again, learning the numberless characters in Chinese script has formed the basis of the mandarin’s education.³ In present-day Europe and America the class that applies the findings of modern science to war, public administration, public works and public sanitation holds a fairly important position, both socially and politically, and in our western world, as in ancient Rome, an altogether priv-

ileged position is held by lawyers. They know the complicated legislation that arises in all peoples of long-standing civilization, and they become especially powerful if their knowledge of law is coupled with the type of eloquence that chances to have a strong appeal to the taste of their contemporaries.

There are examples in abundance where we see that longstanding practice in directing the military and civil organization of a community creates and develops in the higher reaches of the ruling class a real art of governing which is something better than crude empiricism and better than anything that mere individual experience could suggest. In such circumstances aristocracies of functionaries arise, such as the Roman senate, the Venetian nobility and to a certain extent the English aristocracy. Those bodies all stirred John Stuart Mill to admiration and certainly they all three developed governments that were distinguished for carefully considered policies and for great steadfastness and sagacity in carrying them out. This art of governing is not political science, though it has, at one time or another, anticipated applications of a number of the postulates of political science. However, even if the art of governing has now and again enjoyed prestige with certain classes of persons who have long held possession of political functions, knowledge of it has never served as an ordinary criterion for admitting to public offices persons who were barred from them by social station. The degree of mastery of the art of governing that a person possesses is, moreover, apart from exceptional cases, a very difficult thing to determine if the person has given no practical demonstration that he possesses it.

7. In some countries we find hereditary castes. In such cases the governing class is explicitly restricted to a given number of families, and birth is the one criterion that determines entry into the class or exclusion from it. Examples are exceedingly common. There is practically no country of long-standing civilization that has not had a hereditary aristocracy at one period or another in its his-

tory. We find hereditary nobilities during certain periods in China and ancient Egypt, in India, in Greece before the wars with the Medes, in ancient Rome, among the Slavs, among the Latins and Germans of the Middle Ages, in Mexico at the time of the Discovery and in Japan down to a few years ago.

In this connection two preliminary observations are in point. In the first place, all ruling classes tend to become hereditary in fact if not in law. All political forces seem to possess a quality that in physics used to be called the force of inertia. They have a tendency, that is, to remain at the point and in the state in which they find themselves. Wealth and military valor are easily maintained in certain families by moral tradition and by heredity. Qualification for important office—the habit of, and to an extent the capacity for, dealing with affairs of consequence—is much more readily acquired when one has had a certain familiarity with them from childhood. Even when academic degrees, scientific training, special aptitudes as tested by examinations and competitions, open the way to public office, there is no eliminating that special advantage in favor of certain individuals which the French call the advantage of *positions déjà prises*. In actual fact, though examinations and competitions may theoretically be open to all, the majority never have the resources for meeting the expense of long preparation, and many others are without the connections and kinships that set an individual promptly on the right road, enabling him to avoid the gropings and blunders that are inevitable when one enters an unfamiliar environment without any guidance or support.

The democratic principle of election by broadbased suffrage would seem at first glance to be in conflict with the tendency toward stability which, according to our theory, ruling classes show. But it must be noted that candidates who are successful in democratic elections are almost always the ones who possess the political forces above enumerated, which are very often hereditary. In the English, French and Italian parliaments

we frequently see the sons, grandsons, brothers, nephews and sons-in-law of members and deputies, ex-members and ex-deputies.

In the second place, when we see a hereditary caste established in a country and monopolizing political power, we may be sure that such a status de jure was preceded by a similar status de facto. Before proclaiming their exclusive and hereditary right to power the families or castes in question must have held the scepter of command in a firm grasp, completely monopolizing all the political forces of that country at that period. Otherwise such a claim on their part would only have aroused the bitterest protests and provoked the bitterest struggles.

Hereditary aristocracies often come to vaunt supernatural origins, or at least origins different from, and superior to, those of the governed classes. Such claims are explained by a highly significant social fact, namely that every governing class tends to justify its actual exercise of power by resting it on some universal moral principle. This same sort of claim has come forward in our time in scientific trappings. A number of writers, developing and amplifying Darwin's theories, contend that upper classes represent a higher level in social evolution and are therefore superior to lower classes by organic structure. Gumplowicz goes to the point of maintaining that the divisions of populations into trade groups and professional classes in modern civilized countries are based on ethnological heterogeneity.⁴

Now history very definitely shows the special abilities as well as the special defects—both very marked—which have been displayed by aristocracies that have either remained absolutely closed or have made entry into their circles difficult. The ancient Roman patriciate and the English and German nobilities of modern times give a ready idea of the type we refer to. Yet in dealing with this fact, and with the theories that tend to exaggerate its significance, we can always raise the same objection—that the individuals who belong to the aristocracies in question

owe their special qualities not so much to the blood that flows in their veins as to their very particular upbringing, which has brought out certain intellectual and moral tendencies in them in preference to others. . . .

8. Finally, if we were to keep to the idea of those who maintain the exclusive influence of the hereditary principle in the formation of ruling classes, we should be carried to a conclusion somewhat like the one to which we were carried by the evolutionary principle: The political history of mankind ought to be much simpler than it is. If the ruling class really belonged to a different race, or if the qualities that fit it for dominion were transmitted primarily by organic heredity, it is difficult to see how, once the class was formed, it could decline and lose its power. The peculiar qualities of a race are exceedingly tenacious. Keeping to the evolutionary theory, acquired capacities in the parents are inborn in their children and, as generation succeeds generation, are progressively accentuated. The descendants of rulers, therefore, ought to become better and better fitted to rule, and the other classes ought to see their chances of challenging or supplanting them become more and more remote. Now the most commonplace experience suffices to assure one that things do not go in that way at all.

What we see is that as soon as there is a shift in the balance of political forces—when, that is, a need is felt that capacities different from the old should assert themselves in the management of the state, when the old capacities, therefore, lose some of their importance or changes in their distribution occur—then the manner in which the ruling class is constituted changes also. If a new source of wealth develops in a society, if the practical importance of knowledge grows, if an old religion declines or a new one is born, if a new current of ideas spreads, then, simultaneously, far-reaching dislocations occur in the ruling class. One might say, indeed, that the whole history of civilized mankind comes down to a conflict between the tendency of dominant

elements to monopolize political power and transmit possession of it by inheritance, and the tendency toward a dislocation of old forces and an insurgence of new forces; and this conflict produces an unending ferment of endosmosis and exosmosis between the upper classes and certain portions of the lower. Ruling classes decline inevitably when they cease to find scope for the capacities through which they rose to power, when they can no longer render the social services which they once rendered, or when their talents and the services they render lose in importance in the social environment in which they live. So the Roman aristocracy declined when it was no longer the exclusive source of higher officers for the army, of administrators for the commonwealth, of governors for the provinces. So the Venetian aristocracy declined when its nobles ceased to command the galleys and no longer passed the greater part of their lives in sailing the seas and in trading and fighting.

In inorganic nature we have the example of our air, in which a tendency to immobility produced by the force of inertia is continuously in conflict with a tendency to shift about as the result of inequalities in the distribution of heat. The two tendencies, prevailing by turn in various regions on our planet, produce now calm, now wind and storm. In much the same way in human societies there prevails now the tendency that produces closed, stationary, crystallized ruling classes, now the tendency that results in a more or less rapid renovation of ruling classes.

NOTES

1. Mosca, *Teorica dei governi e governo parlamentare*, chap. I.

2. Jannet, *Le istituzioni politiche e sociali degli Stati Uniti d'America*, part II, chap. X.

3. This was true up to a few years ago, the examination of a mandarin covering only literary and historical studies—as the Chinese understood such studies, of course.

4. *Der Rassenkampf*. This notion transpires from Gumplowicz's whole volume. It is explicitly formulated in book II, chap. XXXIII.

29 ■ C. Wright Mills

The Power Elite

The powers of ordinary men are circumscribed by the everyday worlds in which they live, yet even in these rounds of job, family, and neighborhood they often seem driven by forces they can neither understand nor govern. 'Great changes' are beyond their control, but affect their conduct and outlook none the less. The very framework of modern society confines them to projects not their own, but from every side, such changes now press upon the men and women of the mass society, who accordingly feel that they are without purpose in an epoch in which they are without power.

But not all men are in this sense ordinary. As the means of information and of power are centralized, some men come to occupy positions in American society from which they can look down upon, so to speak, and by their decisions mightily affect, the everyday worlds of ordinary men and women. They are not made by their jobs; they set up and break down jobs for thousands of others; they are not confined by simple family responsibilities; they can escape. They may live in many hotels and houses, but they are bound by no one community. They need not merely 'meet the demands of the day and hour'; in some part, they create these demands, and cause others to meet them. Whether or not they profess their power, their technical and political experience of it far transcends that of the underlying population. What Jacob Burckhardt said of 'great men,' most Americans might well say of their elite: 'They are all that we are not.'¹

The power elite is composed of men whose positions enable them to transcend the ordinary environments of ordinary men and women; they are in positions to make decisions having major consequences. Whether they do or do not make such decisions is less important than the fact that they do occupy such pivotal positions: their failure to act, their failure to make decisions, is itself an act that is often of greater consequence than the decisions they do make. For they are in command of the major hierarchies and organizations of modern society. They rule the big corporations. They run the machinery of the state and claim its prerogatives. They direct the military establishment. They occupy the strategic command posts of the social structure, in which are now centered the effective means of the power and the wealth and the celebrity which they enjoy.

The power elite are not solitary rulers. Advisers and consultants, spokesmen and opinion-makers are often the captains of their higher thought and decision. Immediately below the elite are the professional politicians of the middle levels of power, in the Congress and in the pressure groups, as well as among the new and old upper classes of town and city and region. Mingling with them in curious ways are those professional celebrities who live by being continually displayed but are never, so long as they remain celebrities, displayed enough. If such celebrities are not at the head of any dominating hierarchy, they do often have the power to distract the attention of the public or afford

C. Wright Mills. *The Power Elite*, pp. 3–18, 20–23, 365–367. Copyright © 1956 by C. Wright Mills. Renewed 1984 by Yaraslave Mills. Reprinted by permission of Oxford University Press, Inc.

sensations to the masses, or, more directly, to gain the ear of those who do occupy positions of direct power. More or less unattached, as critics of morality and technicians of power, as spokesmen of God and creators of mass sensibility, such celebrities and consultants are part of the immediate scene in which the drama of the elite is enacted. But that drama itself is centered in the command posts of the major institutional hierarchies.

1

The truth about the nature and the power of the elite is not some secret which men of affairs know but will not tell. Such men hold quite various theories about their own roles in the sequence of event and decision. Often they are uncertain about their roles, and even more often they allow their fears and their hopes to affect their assessment of their own power. No matter how great their actual power, they tend to be less acutely aware of it than of the resistances of others to its use. Moreover, most American men of affairs have learned well the rhetoric of public relations, in some cases even to the point of using it when they are alone, and thus coming to believe it. The personal awareness of the actors is only one of the several sources one must examine in order to understand the higher circles. Yet many who believe that there is no elite, or at any rate none of any consequence, rest their argument upon what men of affairs believe about themselves, or at least assert in public.

There is, however, another view: those who feel, even if vaguely, that a compact and powerful elite of great importance does now prevail in America often base that feeling upon the historical trend of our time. They have felt, for example, the domination of the military event, and from this they infer that generals and admirals, as well as other men of decision influenced by them, must be enormously powerful. They hear that the Congress has again abdicated to a handful of

men decisions clearly related to the issue of war or peace. They know that the bomb was dropped over Japan in the name of the United States of America, although they were at no time consulted about the matter. They feel that they live in a time of big decisions; they know that they are not making any. Accordingly, as they consider the present as history, they infer that at its center, making decisions or failing to make them, there must be an elite of power.

On the one hand, those who share this feeling about big historical events assume that there is an elite and that its power is great. On the other hand, those who listen carefully to the reports of men apparently involved in the great decisions often do not believe that there is an elite whose powers are of decisive consequence.

Both views must be taken into account, but neither is adequate. The way to understand the power of the American elite lies neither solely in recognizing the historic scale of events nor in accepting the personal awareness reported by men of apparent decision. Behind such men and behind the events of history, linking the two, are the major institutions of modern society. These hierarchies of state and corporation and army constitute the means of power; as such they are now of a consequence not before equaled in human history—and at their summits, there are now those command posts of modern society which offer us the sociological key to an understanding of the role of the higher circles in America.

Within American society, major national power now resides in the economic, the political, and the military domains. Other institutions seem off to the side of modern history, and, on occasion, duly subordinated to these. No family is as directly powerful in national affairs as any major corporation; no church is as directly powerful in the external biographies of young men in America today as the military establishment; no college is as powerful in the shaping of momentous

events as the National Security Council. Religious, educational, and family institutions are not autonomous centers of national power; on the contrary, these decentralized areas are increasingly shaped by the big three, in which developments of decisive and immediate consequence now occur. . . .

Within each of the big three, the typical institutional unit has become enlarged, has become administrative, and, in the power of its decisions, has become centralized. Behind these developments there is a fabulous technology, for as institutions, they have incorporated this technology and guide it, even as it shapes and paces their developments.

The economy—once a great scatter of small productive units in autonomous balance—has become dominated by two or three hundred giant corporations, administratively and politically interrelated, which together hold the keys to economic decisions.

The political order, once a decentralized set of several dozen states with a weak spinal cord, has become a centralized, executive establishment which has taken up into itself many powers previously scattered, and now enters into each and every cranny of the social structure.

The military order, once a slim establishment in a context of distrust fed by state militia, has become the largest and most expensive feature of government, and, although well versed in smiling public relations, now has all the grim and clumsy efficiency of a sprawling bureaucratic domain.

In each of these institutional areas, the means of power at the disposal of decision-makers have increased enormously; their central executive powers have been enhanced; within each of them modern administrative routines have been elaborated and tightened up.

As each of these domains becomes enlarged and centralized, the consequences of its activities become greater, and its traffic with the others increases. The decisions of a handful of corporations bear upon military and political

as well as upon economic developments around the world. The decisions of the military establishment rest upon and grievously affect political life as well as the very level of economic activity. The decisions made within the political domain determine economic activities and military programs. There is no longer, on the one hand, an economy, and, on the other hand, a political order containing a military establishment unimportant to politics and to money-making. There is a political economy linked, in a thousand ways, with military institutions and decisions. On each side of the world-split running through central Europe and around the Asiatic rimlands, there is an ever-increasing interlocking of economic, military, and political structures.² If there is government intervention in the corporate economy, so is there corporate intervention in the governmental process. In the structural sense, this triangle of power is the source of the interlocking directorate that is most important for the historical structure of the present.

The fact of the interlocking is clearly revealed at each of the points of crisis of modern capitalist society—slump, war, and boom. In each, men of decision are led to an awareness of the interdependence of the major institutional orders. In the nineteenth century, when the scale of all institutions was smaller, their liberal integration was achieved in the automatic economy, by an autonomous play of market forces, and in the automatic political domain, by the bargain and the vote. It was then assumed that out of the imbalance and friction that followed the limited decisions then possible a new equilibrium would in due course emerge. That can no longer be assumed, and it is not assumed by the men at the top of each of the three dominant hierarchies.

For given the scope of their consequences, decisions—and indecisions—in any one of these ramify into the others, and hence top decisions tend either to become co-ordinated or to lead to a commanding indecision. It has

not always been like this. When numerous small entrepreneurs made up the economy, for example, many of them could fail and the consequences still remain local; political and military authorities did not intervene. But now, given political expectations and military commitments, can they afford to allow key units of the private corporate economy to break down in slump? Increasingly, they do intervene in economic affairs, and as they do so, the controlling decisions in each order are inspected by agents of the other two, and economic, military, and political structures are interlocked.

At the pinnacle of each of the three enlarged and centralized domains, there have arisen those higher circles which make up the economic, the political, and the military elites. At the top of the economy, among the corporate rich, there are the chief executives; at the top of the political order, the members of the political directorate; at the top of the military establishment, the elite of soldier-statesmen clustered in and around the Joint Chiefs of Staff and the upper echelon. As each of these domains has coincided with the others, as decisions tend to become total in their consequence, the leading men in each of the three domains of power—the warlords, the corporation chieftains, the political directorate—tend to come together, to form the power elite of America.

2

The higher circles in and around these command posts are often thought of in terms of what their members possess: they have a greater share than other people of the things and experiences that are most highly valued. From this point of view, the elite are simply those who have the most of what there is to have, which is generally held to include money, power, and prestige—as well as all the ways of life to which these lead.³ But the elite are not simply those who have the most, for they could not ‘have the most’ were it not for

their positions in the great institutions. For such institutions are the necessary bases of power, of wealth, and of prestige, and at the same time, the chief means of exercising power, of acquiring and retaining wealth, and of cashing in the higher claims for prestige.

By the powerful we mean, of course, those who are able to realize their will, even if others resist it. No one, accordingly, can be truly powerful unless he has access to the command of major institutions, for it is over these institutional means of power that the truly powerful are, in the first instance, powerful. Higher politicians and key officials of government command such institutional power; so do admirals and generals, and so do the major owners and executives of the larger corporations. Not all power, it is true, is anchored in and exercised by means of such institutions, but only within and through them can power be more or less continuous and important. . . .

If we took the one hundred most powerful men in America, the one hundred wealthiest, and the one hundred most celebrated away from the institutional positions they now occupy, away from their resources of men and women and money, away from the media of mass communication that are now focused upon them—then they would be powerless and poor and uncelebrated. For power is not of a man. Wealth does not center in the person of the wealthy. Celebrity is not inherent in any personality. To be celebrated, to be wealthy, to have power requires access to major institutions, for the institutional positions men occupy determine in large part their chances to have and to hold these valued experiences.

3

The people of the higher circles may also be conceived as members of a top social stratum, as a set of groups whose members know one another, see one another socially and at business, and so, in making decisions,

take one another into account. The elite, according to this conception, feel themselves to be, and are felt by others to be, the inner circle of 'the upper social classes.'⁴ They form a more or less compact social and psychological entity; they have become self-conscious members of a social class. People are either accepted into this class or they are not, and there is a qualitative split, rather than merely a numerical scale, separating them from those who are not elite. They are more or less aware of themselves as a social class and they behave toward one another differently from the way they do toward members of other classes. They accept one another, understand one another, marry one another, tend to work and to think if not together at least alike.

Now, we do not want by our definition to prejudge whether the elite of the command posts are conscious members of such a socially recognized class, or whether considerable proportions of the elite derive from such a clear and distinct class. These are matters to be investigated. Yet in order to be able to recognize what we intend to investigate, we must note something that all biographies and memoirs of the wealthy and the powerful and the eminent make clear: no matter what else they may be, the people of these higher circles are involved in a set of overlapping 'crowds' and intricately connected 'cliques.' There is a kind of mutual attraction among those who 'sit on the same terrace'—although this often becomes clear to them, as well as to others, only at the point at which they feel the need to draw the line; only when, in their common defense, they come to understand what they have in common, and so close their ranks against outsiders.

The idea of such ruling stratum implies that most of its members have similar social origins, that throughout their lives they maintain a network of informal connections, and that to some degree there is an interchangeability of position between the various hierarchies of money and power and

celebrity. We must, of course, note at once that if such an elite stratum does exist, its social visibility and its form, for very solid historical reasons, are quite different from those of the noble cousinhoods that once ruled various European nations.

That American society has never passed through a feudal epoch is of decisive importance to the nature of the American elite, as well as to American society as a historic whole. For it means that no nobility or aristocracy, established before the capitalist era, has stood in tense opposition to the higher bourgeoisie. It means that this bourgeoisie has monopolized not only wealth but prestige and power as well. It means that no set of noble families has commanded the top positions and monopolized the values that are generally held in high esteem; and certainly that no set has done so explicitly by inherited right. It means that no high church dignitaries or court nobilities, no entrenched landlords with honorific accouterments, no monopolists of high army posts have opposed the enriched bourgeoisie and in the name of birth and prerogative successfully resisted its self-making.

But this does *not* mean that there are no upper strata in the United States. That they emerged from a 'middle class' that had no recognized aristocratic superiors does not mean they remained middle class when enormous increases in wealth made their own superiority possible. Their origins and their newness may have made the upper strata less visible in America than elsewhere. But in America today there are in fact tiers and ranges of wealth and power of which people in the middle and lower ranks know very little and may not even dream. There are families who, in their well-being, are quite insulated from the economic jolts and lurches felt by the merely prosperous and those farther down the scale. There are also men of power who in quite small groups make decisions of enormous consequence for the underlying population.

The American elite entered modern history as a virtually unopposed bourgeoisie. No national bourgeoisie, before or since, has had such opportunities and advantages. Having no military neighbors, they easily occupied an isolated continent stocked with natural resources and immensely inviting to a willing labor force. A framework of power and an ideology for its justification were already at hand. Against mercantilist restriction, they inherited the principle of *laissez-faire*; against Southern planters, they imposed the principle of industrialism. The Revolutionary War put an end to colonial pretensions to nobility, as loyalists fled the country and many estates were broken up. The Jacksonian upheaval with its status revolution put an end to pretensions to monopoly of descent by the old New England families. The Civil War broke the power, and so in due course the prestige, of the antebellum South's claimants for the higher esteem. The tempo of the whole capitalist development made it impossible for an inherited nobility to develop and endure in America.

No fixed ruling class, anchored in agrarian life and coming to flower in military glory, could contain in America the historic thrust of commerce and industry, or subordinate to itself the capitalist elite—as capitalists were subordinated, for example, in Germany and Japan. Nor could such a ruling class anywhere in the world contain that of the United States when industrialized violence came to decide history. Witness the fate of Germany and Japan in the two world wars of the twentieth century; and indeed the fate of Britain herself and her model ruling class, as New York became the inevitable economic, and Washington the inevitable political capital of the western capitalist world.

4

The elite who occupy the command posts may be seen as the possessors of power and wealth and celebrity; they may be seen as

members of the upper stratum of a capitalistic society. They may also be defined in terms of psychological and moral criteria, as certain kinds of selected individuals. So defined, the elite, quite simply, are people of superior character and energy.

The humanist, for example, may conceive of the 'elite' not as a social level or category, but as a scatter of those individuals who attempt to transcend themselves, and accordingly, are more noble, more efficient, made out of better stuff. It does not matter whether they are poor or rich, whether they hold high position or low, whether they are acclaimed or despised; they are elite because of the kind of individuals they are. The rest of the population is mass, which, according to this conception, sluggishly relaxes into uncomfortable mediocrity.⁵

This is the sort of socially unlocated conception which some American writers with conservative yearnings have recently sought to develop. But most moral and psychological conceptions of the elite are much less sophisticated, concerning themselves not with individuals but with the stratum as a whole. Such ideas, in fact, always arise in a society in which some people possess more than do others of what there is to possess. People with advantages are loath to believe that they just happen to be people with advantages.

They come readily to define themselves as inherently worthy of what they possess; they come to believe themselves 'naturally' elite; and, in fact, to imagine their possessions and their privileges as natural extensions of their own elite selves. In this sense, the idea of the elite as composed of men and women having a finer moral character is an ideology of the elite as a privileged ruling stratum, and this is true whether the ideology is elite-made or made up for it by others.

In eras of equalitarian rhetoric, the more intelligent or the more articulate among the lower and middle classes, as well as guilty members of the upper, may come to entertain ideas of a counter-elite. In western soci-

ety, as a matter of fact, there is a long tradition and varied images of the poor, the exploited, and the oppressed as the truly virtuous, the wise, and the blessed. Stemming from Christian tradition, this moral idea of a counter-elite composed of essentially higher types condemned to a lowly station, may be and has been used by the underlying population to justify harsh criticism of ruling elites and to celebrate utopian images of a new elite to come.

The moral conception of the elite, however, is not always merely an ideology of the overprivileged or a counter-ideology of the underprivileged. It is often a fact: having controlled experiences and select privileges, many individuals of the upper stratum do come in due course to approximate the types of character they claim to embody. Even when we give up—as we must—the idea that the elite man or woman is born with an elite character, we need not dismiss the idea that their experiences and trainings develop in them characters of a specific type. . . .

5

These several notions of the elite, when appropriately understood, are intricately bound up with one another, and we shall use them all in this examination of American success. We shall study each of several higher circles as offering candidates for the elite, and we shall do so in terms of the major institutions making up the total society of America; within and between each of these institutions, we shall trace the interrelations of wealth and power and prestige. But our main concern is with the power of those who now occupy the command posts, and with the role which they are enacting in the history of our epoch.

Such an elite may be conceived as omnipotent, and its powers thought of as a great hidden design. Thus, in vulgar Marxism, events and trends are explained by reference to 'the will of the bourgeoisie'; in Nazism, by refer-

ence to 'the conspiracy of the Jews'; by the petty right in America today, by reference to 'the hidden force' of Communist spies. According to such notions of the omnipotent elite as historical cause, the elite is never an entirely visible agency. It is, in fact, a secular substitute for the will of God, being realized in a sort of providential design, except that usually non-elite men are thought capable of opposing it and eventually overcoming it.

The opposite view—of the elite as impotent—is now quite popular among liberal-minded observers. Far from being omnipotent, the elites are thought to be so scattered as to lack any coherence as a historical force. Their invisibility is not the invisibility of secrecy but the invisibility of the multitude. Those who occupy the formal places of authority are so check-mated—by other elites exerting pressure, or by the public as an electorate, or by constitutional codes—that, although there may be upper classes, there is no ruling class; although there may be men of power, there is no power elite; although there may be a system of stratification, it has no effective top. In the extreme, this view of the elite, as weakened by compromise and disunited to the point of nullity, is a substitute for impersonal collective fate; for, in this view, the decisions of the visible men of the higher circles do not count in history.

Internationally, the image of the omnipotent elite tends to prevail. All good events and pleasing happenings are quickly imputed by the opinion-makers to the leaders of their own nation; all bad events and unpleasant experiences are imputed to the enemy abroad. In both cases, the omnipotence of evil rulers or of virtuous leaders is assumed. Within the nation, the use of such rhetoric is rather more complicated: when men speak of the power of their own party or circle, they and their leaders are, of course, impotent; only 'the people' are omnipotent. But, when they speak of the power of their opponent's party or circle, they impute to them omnipotence; 'the people' are now powerlessly taken in.

More generally, American men of power tend, by convention, to deny that they are powerful. No American runs for office in order to rule or even govern, but only to serve; he does not become a bureaucrat or even an official, but a public servant. And nowadays, as I have already pointed out, such postures have become standard features of the public-relations programs of all men of power. So firm a part of the style of power-wielding have they become that conservative writers readily misinterpret them as indicating a trend toward an ‘amorphous power situation.’

But the ‘power situation’ of America today is less amorphous than is the perspective of those who see it as a romantic confusion. It is less a flat, momentary ‘situation’ than a graded, durable structure. And if those who occupy its top grades are not omnipotent, neither are they impotent. It is the form and the height of the gradation of power that we must examine if we would understand the degree of power held and exercised by the elite.

If the power to decide such national issues as are decided were shared in an absolutely equal way, there would be no power elite; in fact, there would be no *gradation* of power, but only a radical homogeneity. At the opposite extreme as well, if the power to decide issues were absolutely monopolized by one small group, there would be no gradation of power; there would simply be this small group in command, and below it, the undifferentiated, dominated masses. American society today represents neither the one nor the other of these extremes, but a conception of them is none the less useful: it makes us realize more clearly the question of the structure of power in the United States and the position of the power elite within it.

Within each of the most powerful institutional orders of modern society there is a gradation of power. The owner of a roadside fruit stand does not have as much power in any area of social or economic or political de-

cision as the head of a multi-million-dollar fruit corporation; no lieutenant on the line is as powerful as the Chief of Staff in the Pentagon; no deputy sheriff carries as much authority as the President of the United States. Accordingly, the problem of defining the power elite concerns the level at which we wish to draw the line. By lowering the line, we could define the elite out of existence; by raising it, we could make the elite a very small circle indeed. In a preliminary and minimum way, we draw the line crudely, in charcoal as it were: By the power elite, we refer to those political, economic, and military circles which as an intricate set of overlapping cliques share decisions having at least national consequences. In so far as national events are decided, the power elite are those who decide them. . . .

6

It is not my thesis that for all epochs of human history and in all nations, a creative minority, a ruling class, an omnipotent elite, shape all historical events. Such statements, upon careful examination, usually turn out to be mere tautologies,⁶ and even when they are not, they are so entirely general as to be useless in the attempt to understand the history of the present. The minimum definition of the power elite as those who decide whatever is decided of major consequence, does not imply that the members of this elite are always and necessarily the history-makers; neither does it imply that they never are. We must not confuse the conception of the elite, which we wish to define, with one theory about their role: that they are the history-makers of our time. To define the elite, for example, as ‘those who rule America’ is less to define a conception than to state one hypothesis about the role and power of that elite. No matter how we might define the elite, the extent of its members’ power is subject to historical variation. If, in a dogmatic way, we try to include that variation in our

generic definition, we foolishly limit the use of a needed conception. If we insist that the elite be defined as a strictly coordinated class that continually and absolutely rules, we are closing off from our view much to which the term more modestly defined might open to our observation. In short, our definition of the power elite cannot properly contain dogma concerning the degree and kind of power that ruling groups everywhere have. Much less should it permit us to smuggle into our discussion a theory of history.

During most of human history, historical change has not been visible to the people who were involved in it, or even to those enacting it. Ancient Egypt and Mesopotamia, for example, endured for some four hundred generations with but slight changes in their basic structure. That is six and a half times as long as the entire Christian era, which has only prevailed some sixty generations; it is about eighty times as long as the five generations of the United States' existence. But now the tempo of change is so rapid, and the means of observation so accessible, that the interplay of event and decision seems often to be quite historically visible, if we will only look carefully and from an adequate vantage point.

When knowledgeable journalists tell us that 'events, not men, shape the big decisions,' they are echoing the theory of history as Fortune, Chance, Fate, or the work of The Unseen Hand. For 'events' is merely a modern word for these older ideas, all of which separate men from history-making, because all of them lead us to believe that history goes on behind men's backs. History is drift with no mastery; within it there is action but no deed; history is mere happening and the event intended by no one.⁷

The course of events in our time depends more on a series of human decisions than on any inevitable fate. The sociological meaning of 'fate' is simply this: that, when the decisions are innumerable and each one is of small consequence, all of them add up in a way no man intended—to history as fate.

But not all epochs are equally fateful. As the circle of those who decide is narrowed, as the means of decision are centralized and the consequences of decisions become enormous, then the course of great events often rests upon the decisions of determinable circles. This does not necessarily mean that the same circle of men follow through from one event to another in such a way that all of history is merely their plot. The power of the elite does not necessarily mean that history is not also shaped by a series of small decisions, none of which are thought out. It does not mean that a hundred small arrangements and compromises and adaptations may not be built into the going policy and the living event. The idea of the power elite implies nothing about the process of decision-making as such: it is an attempt to delimit the social areas within which that process, whatever its character, goes on. It is a conception of who is involved in the process.

The degree of foresight and control of those who are involved in decisions that count may also vary. The idea of the power elite does not mean that the estimations and calculated risks upon which decisions are made are not often wrong and that the consequences are sometimes, indeed often, not those intended. Often those who make decisions are trapped by their own inadequacies and blinded by their own errors.

Yet in our time the pivotal moment does arise, and at that moment, small circles do decide or fail to decide. In either case, they are an elite of power. The dropping of the A-bombs over Japan was such a moment; the decision on Korea was such a moment; the confusion about Quemoy and Matsu, as well as before Dienbienphu were such moments; the sequence of maneuvers which involved the United States in World War II was such a 'moment.' Is it not true that much of the history of our times is composed of such moments? And is not that what is meant when it is said that we live in a time of big decisions, of decisively centralized power?

Most of us do not try to make sense of our age by believing in a Greek-like, eternal recurrence, nor by a Christian belief in a salvation to come, nor by any steady march of human progress. Even though we do not reflect upon such matters, the chances are we believe with Burckhardt that we live in a mere succession of events; that sheer continuity is the only principle of history. History is merely one thing after another; history is meaningless in that it is not the realization of any determinate plot. It is true, of course, that our sense of continuity, our feeling for the history of our time, is affected by crisis. But we seldom look beyond the immediate crisis or the crisis felt to be just ahead. We believe neither in fate nor providence; and we assume, without talking about it, that 'we'—as a nation—can decisively shape the future but that 'we' as individuals somehow cannot do so.

Any meaning history has, 'we' shall have to give to it by our actions. Yet the fact is that although we are all of us within history we do not all possess equal powers to make history. To pretend that we do is sociological nonsense and political irresponsibility. It is nonsense because any group or any individual is limited, first of all, by the technical and institutional means of power at its command; we do not all have equal access to the means of power that now exist, nor equal influence over their use. To pretend that 'we' are all history-makers is politically irresponsible because it obfuscates any attempt to locate responsibility for the consequential decisions of men who do have access to the means of power.

From even the most superficial examination of the history of the western society we learn that the power of decision-makers is first of all limited by the level of technique, by the *means* of power and violence and organization that prevail in a given society. In this connection we also learn that there is a fairly straight line running upward through the history of the West; that the means of op-

pression and exploitation, of violence and destruction, as well as the means of production and reconstruction, have been progressively enlarged and increasingly centralized.

As the institutional means of power and the means of communications that tie them together have become steadily more efficient, those now in command of them have come into command of instruments of rule quite unsurpassed in the history of mankind. And we are not yet at the climax of their development. We can no longer lean upon or take soft comfort from the historical ups and downs of ruling groups of previous epochs. In that sense, Hegel is correct: we learn from history that we cannot learn from it.

NOTES

1. Jacob Burckhardt, *Force and Freedom* (New York: Pantheon Books, 1943), pp. 303 ff.

2. Cf. Hans Gerth and C. Wright Mills, *Character and Social Structure* (New York: Harcourt, Brace, 1953), pp. 457 ff.

3. The statistical idea of choosing some value and calling those who have the most of it an elite derives, in modern times, from the Italian economist, Pareto, who puts the central point in this way: 'Let us assume that in every branch of human activity each individual is given an index which stands as a sign of his capacity, very much the way grades are given in the various subjects in examinations in school. The highest type of lawyer, for instance, will be given 10. The man who does not get a client will be given 1—reserving zero for the man who is an out-and-out idiot. To the man who has made his millions—honestly or dishonestly as the case may be—we will give 10. To the man who has earned his thousands we will give 6; to such as just manage to keep out of the poor-house, 1, keeping zero for those who get in. . . . So let us make a class of people who have the highest indices in their branch of activity, and to that class give the name of *elite*.' Vilfredo Pareto, *The Mind and Society* (New York: Harcourt, Brace, 1935), par. 2027 and 2031. Those who follow this approach end up not with one elite, but with a number corresponding to the number of values they select. Like many rather abstract ways of reasoning, this one is useful because it forces us to think in a clear-cut way. For a skillful use of this approach, see the work of Harold D. Lasswell, in particular, *Poli-*

tics: *Who Gets What, When, How* (New York: McGraw-Hill, 1936); and for a more systematic use, H. D. Lasswell and Abraham Kaplan, *Power and Society* (New Haven: Yale University Press, 1950).

4. The conception of the elite as members of a top social stratum, is, of course, in line with the prevailing common-sense view of stratification. Technically, it is closer to 'status group' than to 'class,' and has been very well stated by Joseph A. Schumpeter, 'Social Classes in an Ethically Homogeneous Environment,' *Imperialism and Social Classes* (New York: Augustus M. Kelley, Inc., 1951), pp. 133 ff., especially pp. 137–47. Cf. also his *Capitalism, Socialism and Democracy*, 3rd ed. (New York: Harper, 1950), Part II. For the distinction between class and status groups, see *From Max Weber: Essays in Sociology*, trans. and ed. by Gerth and Mills (New York: Oxford University Press, 1946). For an analysis of Pareto's conception of the elite compared with Marx's conception of classes, as well as data on

France, see Raymond Aron, 'Social Structure and Ruling Class,' *British Journal of Sociology*, vol. I, nos. 1 and 2 (1950).

5. The most popular essay in recent years which defines the elite and the mass in terms of a morally evaluated character-type is probably José Ortega y Gasset's, *The Revolt of the Masses*, 1932 (New York: New American Library, Mentor Edition, 1950), esp. pp. 91 ff.

6. As in the case, quite notably, of Gaetano Mosca, *The Ruling Class* (New York: McGraw-Hill, 1939). For a sharp analysis of Mosca, see Fritz Morstein Marx, 'The Bureaucratic State,' *Review of Politics*, vol. I, 1939, pp. 457 ff. Cf. also Mills, 'On Intellectual Craftsmanship,' April 1952, mimeographed, Columbia College, February 1955.

7. Cf. Karl Löwith, *Meaning in History* (Chicago: University of Chicago Press, 1949), pp. 125 ff. for concise and penetrating statements of several leading philosophies of history.

30 ■ Anthony Giddens

Elites and Power

It is certainly one of the most characteristic emphases of the Marxian perspective that, in capitalism especially (but also, in a general sense, in the prior types of class system), the realm of the 'political' is subordinate to that of the 'economic'. What remains relatively obscure in Marx is the specific form of this dependence, and how it is expressed concretely in the domination of the ruling class.¹ The importance of this is not confined to the analysis of the social structure of capitalism, but bears directly upon the question of the classless character of socialism. It relates, in addition, to issues brought to the forefront by the critique of the Marxian standpoint advanced by the

'elite theorists' of the turn of the century. The substance of this critique, in the writings of such as Pareto and Mosca, may be expressed as an attempt to transmute the Marxian concept of class, as founded in the relations of production, into an essentially *political* differentiation between those 'who rule' and those who 'are ruled'—a transmutation which was, indeed, in part made possible by Marx's failure to specify in a systematic fashion the modes whereby the economic hegemony of the capitalist class becomes 'translated' into the political domination of the *ruling* class. For if it is simply the case that economic control directly yields political power, the way is open for

Anthony Giddens. "Elites and Power," in *The Class Structure of the Advanced Societies*, pp. 118–124. Copyright © 1973 by Anthony Giddens. Reprinted by permission of HarperCollins, Inc.

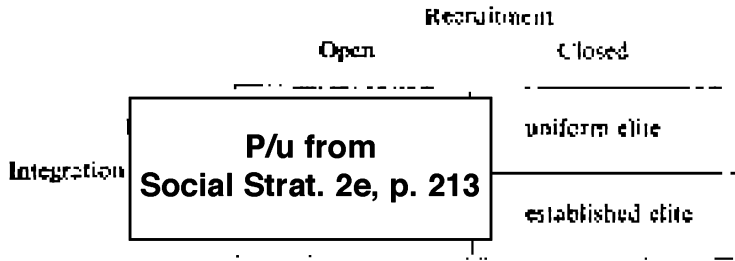
the assertion that, in socialism, as in capitalism (indeed as in any other conceivable type of complex society), whoever controls the means of production thereby achieves political domination as a ruling class. The movement of history from capitalism to socialism then becomes conceived of as a mere succession of 'ruling classes' ('elites'), as in classical 'elite theory', or more specifically as the emergence of the sort of 'managerial' or 'technocratic' ruling class described in Burnham's writings, and more recently in some of the variants of the theory of the 'technocratic society'.²

The points at issue between the Marxian standpoint and 'elite theory' have become further complicated in recent years by the use of concepts drawn from the latter, such as that of 'power elite', as if they were synonymous with that of 'ruling class'. It will be useful to clarify the usage of the terms 'ruling class', 'elite', 'power elite', 'governing class', etc., which involves, in part, looking more closely at the structuration of the upper class.

In the analysis which follows, I shall be interested primarily in developing a set of formulations which illuminate significant conceptual distinctions, rather than adhering to conventional terminological usage—if it can be said, in any case, that there is any conventional practice in a field in which there has been so much confusion.³ I shall suggest that, given the distinctions set out below, there can exist a 'governing class' without it necessarily being a 'ruling class'; that there can exist a 'power elite' without there necessarily being either a 'ruling' or a 'governing class'; that there can be what I shall call a system of 'leadership groups' which constitutes neither a 'ruling class', 'governing class', nor 'power elite'; and that *all* of these social formations are, in principle, compatible with the existence of a society which is 'capitalist' in its organisation. To begin with, a few elementary remarks are necessary about the notion of 'elite'. As it is sometimes

employed, 'elite' may refer to those who 'lead' in any given category of activity: to actors and sportsmen as well as to political or economic 'leaders'. There is evidently a difference, however, between the first and the second, in that the former 'lead' in terms of some sort of scale of 'fame' or 'achievement', whereas the second usage may be taken to refer to persons who are at the head of a specific social organisation which has an internal authority structure (the state, an economic enterprise, etc.). I shall use the term 'elite group' in this latter sense, to designate those individuals who occupy positions of formal authority at the head of a social organisation or institution; and 'elite' very generally, to refer either to an elite group or cluster of elite groups.

In these terms, it can be said that a major aspect of the structuration of the upper class concerns, first, the process of mobility into or recruitment to, elite positions and, second, the degree of social 'solidarity' within and between elite groups. Mediate structuration thus concerns how 'closed' the process of recruitment to elite positions is, in favour of those drawn from propertied backgrounds. Proximate structuration depends primarily upon the frequency and nature of the social contacts between the members of elite groups. These may take various forms, including the formation of marriage connections or the existence of other kin ties, the prevalence of personal ties of acquaintance or friendship, etc. If the extent of social 'integration' of elite groups is high, there is also likely to be a high degree of moral solidarity characterising the elite as a whole and, probably, a low incidence of either latent or manifest conflicts between them. There has never been any elite, however solidary, which has been free of conflicts and struggles; but the degree and intensity of overt conflict has varied widely, and thus it is reasonable to speak broadly of differentials in the solidarity of elite groups.

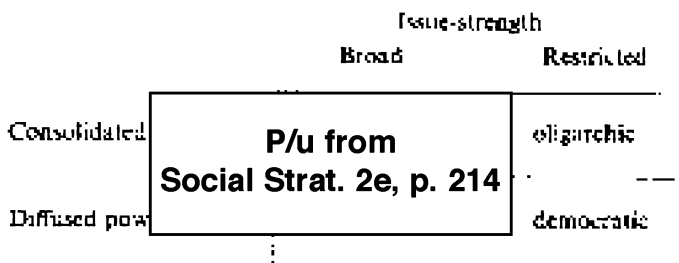


Combining these two aspects of structuration, we can establish a typology of elite formations [see diagram on this page].

A 'uniform' elite is one which shares the attributes of having a restricted pattern of recruitment and of forming a relatively tightly knit unity. It hardly needs emphasising that the classifications involved are not of an all-or-nothing character. The point has been made that even among traditional aristocracies there was never a completely 'closed' pattern of recruitment, something which has only been approached by the Indian caste system—all elites open their ranks, in some degree, to individuals from the lower orders, and may enhance their stability thereby. A relatively closed type of recruitment, however, is likely to supply the sort of coherent socialisation process producing a high level of solidarity between (and within) elite groups. But it is quite feasible to envisage the existence of instances which approximate more closely to the case of an 'established' elite, where there is a relatively closed pattern of recruitment, but only a low level of integration between elite groups. A 'solidary' elite, as defined in the classification, might also appear to involve an unlikely combination of elements, since it might seem difficult to attain a high degree of integration among elite groups whose members are drawn from diverse class backgrounds. But, while this type of social formation is probably rare in capitalist societies, at least some of the state socialist countries fit quite neatly into this category: the Communist Party is the main

channel of access to elite positions, and while it provides an avenue of mobility for individuals drawn in substantial proportions from quite lowly backgrounds, at the same time it ensures a high degree of solidarity among elite groups. An 'abstract' elite, involving both relatively open recruitment and a low level of elite solidarity, whatever its empirical reality, approximates closely to the picture of certain contemporary capitalist societies as these are portrayed in the writings of the theorists of so-called 'pluralist democracy'.

The distinguishing of different types of elite formation does not, in itself, enable us to conceptualise the phenomenon of power. As in the case of class structuration itself, we may distinguish two forms of the mediation of power relationships in society. The first I shall call the *institutional* mediation of power; the other, the mediation of power in terms of *control*. By the institutional mediation of power, I mean the general form of state and economy within which elite groups are recruited and structured. This concerns, among other things, the role of property in the overall organisation of economic life, the nature of the legal framework defining economic and political rights and obligations, and the institutional structure of the state itself. The mediation of control refers to the actual (effective) power of policy-formation and decision-making held by the members of particular elite groups: how far, for example, economic leaders are able to influence decisions taken by politicians, etc. To express it another way, we can say that power has two



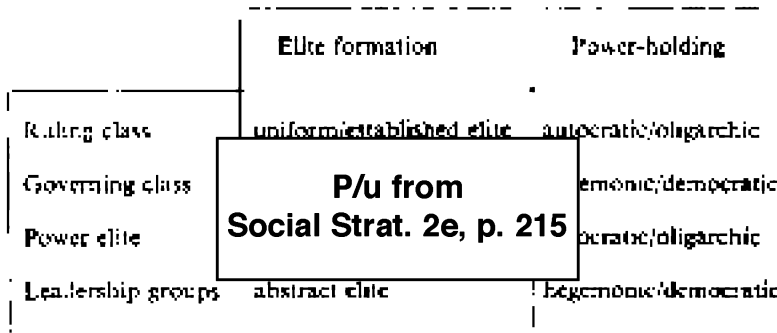
aspects: a ‘collective’ aspect, in the sense that the ‘parameters’ of any concrete set of power relationships are contingent upon the overall system of organisation of a society; and a ‘distributive’ aspect, in the sense that certain groups are able to exert their will at the expense of others.⁴ The mediation of control is thus expressed in terms of ‘effective’ power, manifest in terms of the capacity either to take or to influence the taking of decisions which affect the interests of two or more parties differentially.

We may conceptually separate two variable factors in analysing effective power (that is to say, power as differentiated from ‘formal authority’) in relation to types of elite formation. The first concerns how far such power is ‘consolidated’ in the hands of elite groups; the second refers to the ‘issue-strength’ of the power wielded by those in elite positions. While the former designates limitations upon effective power, deriving from constraints imposed from ‘below’, the latter concerns how far that power is limited *because it can only be exercised in relation to a range of restricted issues*. Thus it is often held to be characteristic of modern capitalist societies that there are quite narrowly defined limitations upon the issues over which elite groups are able to exercise control.⁵ By combining these two aspects of effective power as exercised by elite groups, we can establish a classification of forms of power-structure [see diagram on this page]. Like the previous typology, this sets out an abstract combination of possibilities; it goes almost without saying that this is no more than an elementary categorisation of a very complex set of

phenomena, and the labels applied here in no way exhaust the variety of characteristics which are frequently subsumed under these terms.

According to these definitions, the consolidation of effective power is greatest where it is not restricted to clearly defined limits in terms of its ‘lateral range’ (broad ‘issue-strength’), and where it is concentrated in the hands of the elite, or an elite group. Power-holding is ‘oligarchic’ rather than ‘autocratic’ where the degree of centralisation of power in the hands of elite groups is high, but where the issue-strength of that power is limited. In the case of ‘hegemonic’ control, those in elite positions wield power which, while it is not clearly defined in scope and limited to a restricted range of issues, is ‘shallow’. A ‘democratic’ order, in these terms, is one in which the effective power of elite groups is limited in both respects.

Finally, bringing together both classifications formulated above, we can set up an overall typology of elite formations and power within the class structure [see diagram on the next page]. This makes possible a clarification of the four concepts already mentioned—‘ruling class’, ‘governing class’, ‘power elite’ and ‘leadership groups’. It must be emphasised that these partially cross-cut some of the existing usages in the literature on class and elite theory. The Paretian term ‘governing class’ is here not, as in Pareto’s own writing, a replacement for the Marxian ‘ruling class’; in this scheme, a governing class is ‘one step down’, both in terms of elite formation and power-holding, from a ‘ruling class’.



In this scheme, the 'strongest' case of a ruling class is defined as that where a uniform elite wields 'autocratic' power; the weakest is where an established elite holds 'oligarchic' power. Where a relatively closed recruitment pattern is linked with the prevalence of defined restrictions upon the effective power of elite groups, a governing class exists, but not a ruling class. A governing class borders upon being a ruling class where a uniform elite possesses 'hegemonic' power; and comes closest to being a system of leadership groups where an established elite holds 'democratic' power. Where a governing class involves a combination of an established elite and 'hegemonic' power, it stands close to being a power elite. A power elite is distinguished from a ruling class in terms of pattern of recruitment, as is a governing class from a system of leadership groups. The latter exists where elite groups only hold limited power, and where, in addition, elite recruitment is relatively open in character.

In terms of the mediation of control, this classification leaves undefined the relative primacy of the power of any one elite group over others. This can be conceptually expressed as referring to the nature of the *hierarchy* which exists among elite groups. A hierarchy exists among elite groups in so far as one such group holds power of broader issue-strength than others, and is thereby able to exert a degree of control over decisions taken by those within them. Thus it may be that the economic elite, or certain

sectors of the economic elite, are able to significantly condition political decisions through the use of 'influence', 'inducement', or the 'direct' control of political positions—i.e., through the fact that members of the economic elite are also incumbents of political positions. We may refer to all of these modes of obtaining, or striving for, control as the *media of interchange* between elite groups. It is precisely one of the major tasks of the analysis of elite formations to examine the media of interchange which operate between elite groups in any given society in order to determine what kinds of elite hierarchy exist.

NOTES

1. Most subsequent Marxist authors have either been content with the most generalised assertions about this issue, or have wanted to have their cake and eat it by insisting that capitalism is dominated by a ruling class who do not actually 'rule'; cf. Nicos Poulantzas, *Pouvoir politique et classes sociales de l'état capitaliste* (Paris 1970), pp. 361ff.

2. James Burnham, *The Managerial Revolution* (New York 1941).

3. In this section of this chapter I have drawn upon part of my article 'Elites in the British class structure', *Sociological Review* 20, 1972.

4. cf. Talcott Parsons, 'On the concept of political power', *Proceedings of the American Philosophical Society* 107, 1963. The error in Parsons' analysis, however, is to take insufficient account of the fact that the 'collective' aspect of power is asymmetrical in its consequences for the different groupings in society.

5. As in Keller's 'strategic elites'. See Suzanne Keller, *Beyond the Ruling Class* (New York 1963).

■ CONTEMPORARY STATEMENTS ■

31 ■ G. William Domhoff

Who Rules America?

Power and Politics

Do corporations have far too much power in the United States? Does the federal government ignore the interests of everyday people? The great majority of Americans—70 to 75 percent in some surveys—answer “yes” to both questions.¹ This chapter explains why their answers are accurate even though there is freedom of speech, the possibility of full political participation, and increasing equality of opportunity due to the civil rights and women’s movements. In other words, it attempts to resolve a seeming paradox that has bedeviled social scientists and political activists for a long time: How is it possible to have such extreme corporate domination in a democratic country?

This paradox is made all the more striking because corporations do not have as much power in most other democratic countries. The wealth and income differences between people at the top and the bottom are not as great, and the safety net for those who are poor, ill, or elderly is stronger. Why does the richest nation in the world also have the most poverty compared to any other democratic country?

Using a wide range of systematic empirical findings, this chapter shows how the owners and top-level managers in large companies work together to maintain themselves as the

core of the dominant power group. Their corporations, banks, and agribusinesses form a *corporate community* that shapes the federal government on the policy issues of interest to it, issues that have a major impact on the income, job security, and well-being of most other Americans. At the same time, there is competition within the corporate community for profit opportunities, which can lead to highly visible policy conflicts among rival corporate leaders that are sometimes fought out in Congress. Yet the corporate community is cohesive on the policy issues that affect its general welfare, which is often at stake when political challenges are made by organized workers, liberals, or strong environmentalists. The chapter therefore deals with another seeming paradox: How can a highly competitive group of corporate leaders cooperate enough to work their common will in the political and policy arenas? . . .

Partly because the owners and high-level managers within the corporate community share great wealth and common economic interests, but also due to political opposition to their interests, they band together to develop their own social institutions—gated neighborhoods, private schools, exclusive social clubs, debutante balls, and secluded summer resorts. These social institutions

G. William Domhoff. *Who Rules America? Power and Politics*, Fourth Edition, pp. xi–xii, 45–51, 57, 67–68, 216–218. Copyright © 2002 by McGraw-Hill Companies.

create social cohesion and a sense of group belonging, a “we” feeling, and thereby mold wealthy people into a *social upper class*. In addition, the owners and managers supplement their small numbers by financing and directing a wide variety of nonprofit organizations—e.g., tax-free foundations, think tanks, and policy-discussion groups—to aid them in shaping public policy. The highest-ranking employees in these nonprofit organizations become part of a general leadership group for the corporate community and the upper class. . . .

The Corporate Community and the Upper Class

Why does it matter whether or not the corporate community and the upper class are intertwined? If the corporate community and the upper class are essentially one and the same in terms of people and objectives, then it is more likely that they have the wealth, social cohesion, and awareness of their common interests to organize themselves well enough to dominate government. If they are separate, the wealth and status of the upper class might form a rival power base and make the power structure more open.

The nationwide social upper class has its own exclusive social institutions and is based in the ownership of great wealth. The social cohesion that develops among members of the upper class . . . is based in the two types of relationships found in a membership network: common membership in specific social institutions and friendships based on social interactions within those institutions. Research on small groups in laboratory settings suggests that social cohesion is greatest when (1) the social groups are seen to be exclusive and of high status; and (2) when the interactions take place in relaxed and informal settings.² Many of the social institutions of the upper class fit these specifications very well. From the viewpoint of social psychology, the people who make up the upper class can be

seen as members of numerous small groups that meet at private schools, social clubs, retreats, and resorts. . . .

Prepping for Power

From infancy through young adulthood, members of the upper class receive a distinctive education. This education begins early in life in preschools that sometimes are attached to a neighborhood church of high social status. Schooling continues during the elementary years at a local private school called a day school. The adolescent years may see the student remain at day school, but there is a strong chance that at least one or two years will be spent away from home at a boarding school in a quiet rural setting. Higher education is obtained at one of a small number of prestigious private universities. Although some upper-class children may attend public high school if they live in a secluded suburban setting, or go to a state university if there is one of great esteem and tradition in their home state, the system of formal schooling is so insulated that many upper-class students never see the inside of a public school in all their years of education. This separate educational system is important evidence for the distinctiveness of the mentality and lifestyle that exists within the upper class, because schools play a large role in transmitting the class structure to their students.

The linchpins in the upper-class educational system are the dozens of boarding schools developed in the last half of the nineteenth and the early part of the twentieth centuries, coincident with the rise of a nationwide upper class whose members desired to insulate themselves from an inner city that was becoming populated by lower-class immigrants. They become surrogate families that play a major role in creating an upper-class subculture on a national scale in America.³ The role of boarding schools in providing connections to other upper-class social institutions is also important. . . .

Virtually all graduates of private secondary schools go on to college, and most do so at prestigious universities. Graduates of the New England boarding schools, for example, historically found themselves at three or four large Ivy League universities: Harvard, Yale, Princeton, and Columbia. However, that situation changed somewhat after World War II as the universities grew and provided more scholarships. An analysis of admission patterns for graduates of 14 prestigious boarding schools between 1953 and 1967 demonstrated this shift by showing that the percentage of their graduates attending Harvard, Yale, or Princeton gradually declined over those years from 52 to 25 percent. Information on the same 14 schools for the years 1969 to 1979 showed that the figure had bottomed out at 13 percent in 1973, 1975, and 1979, with some schools showing very little change from the late 1960s and others dropping even more dramatically.⁴ Now many upper-class students attend a select handful of smaller private colleges, most of which are in the East, but a few in the South and West as well.

Graduates of private schools outside of New England most frequently attend a prominent state university in their area, but a significant minority go to eastern Ivy League and top private universities in other parts of the country. For example, the Cate School, a boarding school near Santa Barbara, California, is modeled after its New England counterparts and draws most of its students from California and other western states. In the four years between 1993 and 1996, 35 percent of the 245 graduates went to one of fifteen prestigious Ivy League schools, with Middlebury (12), Harvard (10), and Brown (7) topping the list. The other leading destinations for Cate graduates were the University of California (27), Stanford (9), University of Colorado (9), Georgetown (8), Duke (7), Vanderbilt (6), and University of Chicago (5). Or, to take another example, St. John's in Houston is a lavishly endowed day

school built in the Gothic architecture typical of many universities. From 1992 through 1996, 22 percent of its 585 graduates went to the fifteen Ivy League schools used in the Cate analysis, with Princeton (27), the University of Pennsylvania (15), Cornell (13), Harvard (12), and Yale (12) the most frequent destinations. As might be expected, 105 graduates went to the University of Texas (18 percent), but Rice (49), Vanderbilt (33), and Stanford (15) were high on the list. Few graduates of either Cate or St. John's went to less prestigious state schools.⁵

Most private school graduates pursue careers in business, finance, or corporate law, which is the first evidence for the intertwining of the upper class and the corporate community. Their business-oriented preoccupations are demonstrated in the greatest detail in a study of all those who graduated from Hotchkiss between 1940 and 1950. Using the school's alumni files, the researcher followed the careers of 228 graduates from their date of graduation until 1970.⁶ Fifty-six percent of the sample are either bankers or business executives, with 80 of the 91 businessmen serving as president, vice-president, or partner in their firms. Another 10 percent of the sample are lawyers, mostly as partners in large firms closely affiliated with the corporate community.

The involvement of private school graduates on boards of directors is demonstrated in a study of all alumni over the age of 45 from one of the most prestigious eastern boarding schools, St. Paul's. Using *Poor's Register of Corporations, Directors and Executives*, and *Who's Who in America*, one finds that 303 of these several thousand men are serving as officers or directors in corporations in general, and that 102 are directors of 97 corporations in the *Fortune* 800. Their involvement is especially great in the financial sector. Most striking of all, 21 graduates of St. Paul's are either officers or directors at J. P. Morgan Bank, which for a time was one of the five largest banks in the country until it merged

with Chase Manhattan Bank in late 2000. This finding suggests that the alumni of particular schools may tend to cluster at specific banks or corporations.

Social Clubs

Private social clubs are a major point of orientation in the lives of upper-class adults. These clubs also have a role in differentiating members of the upper class from other members of society. The clubs of the upper class are many and varied, ranging from family-oriented country clubs and downtown men's and women's clubs to highly specialized clubs for yachtsmen, sportsmen, gardening enthusiasts, and fox hunters. Downtown men's clubs originally were places for having lunch and dinner, and occasionally for attending an evening performance or a weekend party. As upper-class families deserted the city for large suburban estates, a new kind of club, the country club, gradually took over some of these functions. The downtown club became almost entirely a luncheon club, a site to hold meetings, or a place to relax on a free afternoon. The country club, by contrast, became a haven for all members of the family. It offered social and sporting activities ranging from dances, parties, and banquets to golf, swimming, and tennis. Special group dinners were often arranged for all members on Thursday night, the traditional maid's night off across the United States.

Initiation fees, annual dues, and expenses vary from a few thousand dollars in downtown clubs to tens of thousands of dollars in some country clubs, but money is not the primary barrier in gaining membership to a club. Each club has a very rigorous screening process before accepting new members. Most require nomination by one or more active members, letters of recommendation from three to six members, and interviews with at least some members of the membership committee. Negative votes by two or three members of what is typically a 10- to 20-person

committee often are enough to deny admission to the candidate. The carefulness with which new members are selected extends to a guarding of club membership lists, which are usually available only to club members. Research on clubs therefore has to be based on out-of-date membership lists that have been given to historical libraries by members or their surviving spouses.

Men and women of the upper class often belong to clubs in several cities, creating a nationwide pattern of overlapping memberships. These overlaps provide evidence for social cohesion within the upper class. An indication of the nature and extent of this overlapping is revealed in a study of membership lists for 20 clubs in several major cities across the country, including the Links in New York, the Duquesne in Pittsburgh, the Chicago in Chicago, the Pacific Union in San Francisco, and the California in Los Angeles. There is sufficient overlap among 18 of the 20 clubs to form three regional groupings and a fourth group that provides a bridge between the two largest regional groups. The several dozen men who are in three or more of the clubs, most of them very wealthy people who also sit on several corporate boards, are especially important in creating the overall pattern.⁷

The overlap of this club network with corporate boards of directors provides important evidence for the intertwining of the upper class and corporate community. In one study, the club memberships of the chairpersons and outside directors of the 20 largest industrial corporations were counted. The overlaps with upper-class clubs in general are ubiquitous, but the concentration of directors in a few clubs is especially notable. At least one director from 12 of the 20 corporations is a member of the Links Club, which is the New York meeting grounds of the national corporate establishment. Seven of General Electric's directors are members, as are four from Chrysler, four from Westinghouse, and three from IBM. In addition to

the Links, several other clubs have directors from four or more corporations. Another study, using membership lists from 11 prestigious clubs in different parts of the country, confirms and extends these findings. A majority of the top 25 corporations in every major sector of the economy have directors in at least one of these clubs, and several have many more. For example, all of the 25 largest industrials have one or more directors in these 11 clubs. The Links in New York, with 79 connections to 21 industrial corporations, has the most.⁸ . . .

Continuity and Upward Mobility

Americans always have believed that anyone can rise from rags to riches if they try hard enough, but in fact a rise from the bottom to the top is very rare and often a matter of luck—being at the right place at the right time. In the late nineteenth century, a wealthy upper-class Bostonian with a Harvard education, Horatio Alger, became a best-selling author by writing short fictional books on young boys who had gone from penniless adversity to great wealth. In real life, the commentators of his day pointed to three or four actual examples. Subsequent research showed that most of the business leaders of that era did not fit the Horatio Alger myth. As one historian put it, Horatio Alger appeared more frequently in magazines and textbooks than they did in reality.⁹

Since 1982, the Horatio Alger story line has been taken up by *Forbes*, a business magazine that publishes an annual list of the allegedly 400 richest Americans. “Forget old money,” says the article that introduces the 1996 list. “Forget silver spoons. Great fortunes are being created almost monthly in the U.S. today by young entrepreneurs who hadn’t a dime when we created this list 14 years ago.”¹⁰ But the Horatio Alger story is no less rare today than it was in the 1890s. A study of all those on the *Forbes* lists for 1995 and 1996 showed that at least 56 percent of

them came from millionaire families and that another 14 percent came from the top 10 percent of the income ladder.¹¹ . . .

Class Awareness: A Capitalist Mentality

The institutions that establish the owners and high-level executives of corporations as a national upper class transcend the presence or absence of any given person or family. Families can rise and fall in the class structure, but the institutions of the upper class persist. Not everyone in this nationwide upper class knows everyone else, but everybody knows somebody who knows someone in other areas of the country, thanks to a common school experience, a summer at the same resort, membership in the same social club, or membership on the same board of directors. The upper class at any given historical moment consists of a complex network of overlapping social circles knit together by the members they have in common and by the numerous signs of equal social status that emerge from a similar lifestyle. Viewed from the standpoint of social psychology, the upper class is made up of innumerable face-to-face small groups that are constantly changing in their composition as people move from one social setting to another.

Involvement in these institutions usually instills a class awareness that includes feelings of superiority, pride, and justified privilege. Deep down, most members of the upper class think they are better than other people, and therefore fully deserving of their station in life. This class awareness is based in corporate ownership, but it is reinforced by the shared social identities and interpersonal ties created through participation in the social institutions of the upper class.

The fact that the upper class is based in the ownership and control of profit-producing investments in stocks, bonds, and real estate shows that it is a capitalist class as well as a social class. Its members are not concerned

simply with the interests of one corporation or business sector, but with such matters as the investment climate, the rate of profit, and the overall political climate. That is, they have a capitalist mentality.

NOTES

1. Aaron Bernstein, "Too Much Corporate Power?" *BusinessWeek*, 11 September 2000, 145–149.
2. Michael Hogg, *The Social Psychology of Group Cohesiveness* (New York: New York University Press, 1992).
3. Peter W. Cookson and Caroline Hodges Persell, *Preparing for Power: America's Elite Boarding Schools* (New York: Basic Books, 1985).
4. Michael Gordon, "Changing Patterns of Upper-Class Prep School College Placements," *Pacific Sociological Review* 12, no. 1 (1969): 23–26.
5. These figures were obtained from the admissions offices at Cate and St. John's in 1997.

6. Christopher Armstrong, "Privilege and Productivity: The Cases of Two Private Schools and Their Graduates," Ph.D. Dissertation, University of Pennsylvania, 1974.

7. G. William Domhoff, "Social Clubs, Policy-Planning Groups, and Corporations: A Network Study of Ruling-Class Cohesiveness," *Insurgent Sociologist* 5, no. 3 (1975): 173–184.

8. G. William Domhoff, "Social Clubs, Policy-Planning Groups, and Corporations: A Network Study of Ruling-Class Cohesiveness," *Insurgent Sociologist* 5, no. 3 (1975): 173–184; G. William Domhoff, *Who Rules America?*, 1st ed. (Englewood Cliffs, N.J.: Prentice-Hall, 1967).

9. William Miller, "American Historians and the Business Elite," *Journal of Economic History* 9 (1949): 184–208.

10. Ann Marsh, "The Forbes Four Hundred," *Forbes*, 14 October 1996, 100.

11. Chuck Collins, *Born on Third Base: The Sources of Wealth of the 1996 Forbes 400* (Boston: United for a Fair Economy, 1997).

32 ■ Alvin W. Gouldner

The Future of Intellectuals and the Rise of the New Class

In all countries that have in the twentieth century become part of the emerging world socioeconomic order, a New Class composed of intellectuals and technical intelligentsia—not the same—enters into contention with the groups already in control of the society's economy, whether these are businessmen or party leaders. A new contest of classes and a new class system is slowly arising in the third world of developing nations, in the second world of the USSR and its client states, and in the first world of late capitalism of North America, Western Europe, and Japan.

The early historical evolution of the New Class in Western Europe, its emergence into the public sphere as a structurally differentiated and (relatively) autonomous social stratum, may be defined in terms of certain critical episodes. What follows is only a synoptic inventory of some episodes decisive in the formation of the New Class.

1. A process of secularization in which most intelligentsia are no longer trained by, living within, and subject to close supervision by a churchly organization, and thus separated from the everyday life of society.¹

Alvin W. Gouldner. *The Future of Intellectuals and the Rise of the New Class* (New York: Continuum Publishing Service, 1979), pp. 1–8, 18–20, 28–29, 83–85, 102–104, 106, 110, 112–113. Reprinted by permission of Continuum International Publishing Group.

Secularization is important because it desacralizes authority-claims and facilitates challenges to definitions of social reality made by traditional authorities linked to the church. Secularization is important also because it is an infrastructure on which there develops the modern grammar of rationality, or culture of critical discourse, with its characteristic stress on self-groundedness—in Martin Heidegger’s sense of the “mathematical project.”²

2. A second episode in the emergence of the New Class is the rise of diverse vernacular languages, the corresponding decline of Latin as the language of intellectuals, and especially of their scholarly production. Latin becomes a ritual, rather than a technical language. This development further dissolves the membrane between everyday life and the intellectuals—whether clerical or secular.

3. There is a breakdown of the feudal and old regime system of personalized *patronage* relations between the old hegemonic elite and individual members of the New Class as cultural producers; and

4. A corresponding growth of an anonymous *market* for the products and services of the New Class, thus allowing them to make an independent living apart from close supervision and *personalized controls by patrons*. Along with secularization, this means that the residence and work of intellectuals are both now less closely supervised by others.

They may now more readily take personal initiatives in the public, political sphere, while also having a “private” life.

5. The character and development of the emerging New Class also depended importantly on the multi-national structure of European polities. That Europe was not a single empire with a central authority able to impose a single set of norms throughout its territory, but a system of competing and autonomous states with diverse cultures and religions, meant that dissenting intellectuals, scientists, and divines could and did protect their own intellectual innovations by migrat-

ing from their home country when conditions there grew insupportable and sojourning in foreign lands. Even the enforced travel of exiled intellectuals also enabled them to enter into a European-wide communication network. In an article (as yet unpublished), Robert Wuthnow has suggested that their often extensive travel led many intellectuals to share a cosmopolitan identity transcending national limits and enhancing their autonomy from local elites.

6. A sixth episode in the formation of the New Class is the waning of the extended, patriarchal family system and its replacement by the smaller, nuclear family. As middle class women become educated and emancipated, they may increasingly challenge paternal authority and side with their children in resisting it. With declining paternal authority and growing maternal influence, the autonomy strivings of children are now more difficult to repress; hostility and rebellion against paternal authority can become more overt. There is, correspondingly, increasing difficulty experienced by paternal authority in imposing and reproducing its social values and political ideologies in their children.

7. Following the French Revolution, there is in many parts of Europe, especially France and Germany, a profound reformation and extension of *public, non-church* controlled, (relatively more) *multi-class* education, at the lower levels as well as at the college, polytechnical, and university levels. On the one hand, higher education in the public school becomes the institutional basis for the *mass* production of the New Class of intelligentsia and intellectuals. On the other hand, the expansion of primary and secondary public school teachers greatly increases the jobs available to the New Class.

As teachers, intellectuals come to be defined, and to define themselves, as responsible for and “representative” of society as a *whole*,³ rather than as having allegiance to the class interests of their students or their parents. As teachers, they are not defined as

having an *obligation* to reproduce parental values in their children. Public teachers supersede private tutors.

8. The new structurally differentiated educational system is increasingly insulated from the family system, becoming an important source of values among students divergent from those of their families. The socialization of the young by their families is now mediated by a *semi*-autonomous group of teachers.

9. While growing public education limits family influence on education, it also increases the influence of the state on education. The public educational system thus becomes a major *cosmopolitanizing* influence on its students, with a corresponding distancing from *localistic* interests and values.

10. Again, the new school system becomes a major setting for the intensive linguistic conversion of students from casual to reflexive speech, or (in Basil Bernstein's terms) from "restricted" linguistic codes to "elaborated" linguistic codes,⁴ to a culture of discourse in which claims and assertions may *not* be justified by reference to the speaker's social status. This has the profound consequence of making all *authority-referring* claims potentially problematic.

11. This new culture of discourse often diverges from assumptions fundamental to everyday life, tending to put them into question even when they are linked to the upper classes. These school-inculcated modes of speech are, also, (relatively) situation-free language variants. Their situation-freeness is further heightened by the "communications revolution" in general, and by the development of printing technology, in particular. With the spread of printed materials, definitions of social reality available to intellectuals may now derive increasingly from *distant* persons, from groups geographically, culturally, and historically distant and even from dead persons, and may therefore diverge greatly from any local environment in which they are received. Definitions of social reality made by local elites may now be invidiously

contrasted (by intellectuals) with definitions made in other places and times.

12. With the spread of public schools, literacy spreads; humanistic intellectuals lose their exclusiveness and privileged market position, and now experience a status disparity between their "high" culture, as they see it, and their lower deference, repute, income and social power. The social position of humanistic intellectuals, *particularly in a technocratic and industrial society*, becomes more marginal and alienated than that of the technical intelligentsia. The New Class becomes internally differentiated.

13. Finally, a major episode in the emergence of the modern intelligentsia is the changing form of the revolutionary *organization*. Revolution itself becomes a technology to be pursued with "instrumental rationality." The revolutionary organization evolves from a ritualistic, oath-bound secret society into the modern "vanguard" party. When the *Communist Manifesto* remarks that Communists have nothing to hide,⁵ it is exactly a proposed emergence into *public* life which is implied. The *Communist Manifesto* was written by Marx and Engels for the "League of Communists," which was born of the "League of the Just" which, in turn, was descended from the "League of Outlaws." This latter group of German emigrants in Paris had a pyramidal structure, made a sharp distinction between upper and lower members, blindfolded members during initiation ceremonies, used recognition signs and passwords, and bound members by an oath.⁶ The vanguard organization, however, de-ritualizes participation and entails elements of both the "secret society" and of the public political party. In the vanguard organization, public refers to the public availability of the *doctrine* rather than the availability of the organization or its membership to public scrutiny. Here, to be "public" entails the organization's rejection of "secret doctrines" known only to an elite in the organization—as, for instance, Bakunin's doctrine of an elite dictatorship of anarchists.⁷ The

modern vanguard structure is first clearly encoded in Lenin's *What Is to Be Done?* Here it is plainly held that the proletariat cannot develop a *socialist* consciousness by itself, but must secure this from a scientific theory developed by the intelligentsia.⁸ The "vanguard" party expresses the *modernizing* and elite ambitions of the New Class as well as an effort to overcome its political limitations. Lenin's call for the development of "professional" revolutionaries, as the core of the vanguard, is a rhetoric carrying the tacit promise of a *career*-like life which invites young members of the New Class to "normalize" the revolutionary existence. . . .

There are several distinguishable conceptions of the New Class:

1. *New Class as Benign Technocrats*: Here the New Class is viewed as a new historical elite already entrenched in institutional influence which it uses in benign ways for society; it is more or less inevitable and trustworthy: e.g., Galbraith,⁹ Bell,¹⁰ Berle and Means.¹¹

(*Sed contra*: This obscures the manner in which the New Class egoistically pursues its own special vested interests. Moreover, the power of the New Class today is scarcely entrenched. This view also ignores the limits on the rationality of the New Class.)

2. *New Class as Master Class*: Here the New Class is seen as another moment in a long-continuing circulation of historical elites, as a socialist intelligentsia that brings little new to the world and continues to exploit the rest of society as the old class had, but now uses education rather than money to exploit others: Bakunin,¹² Machajski.¹³

(*Sed contra*: The New Class is more historically unique and discontinuous than this sees; while protecting its own special interests, it is not bound by the same *limits* as the old class and, at least transiently, contributes to collective needs.)

3. *New Class as Old Class Ally*: The New Class is here seen as a benign group of dedicated "professionals" who will uplift the old (moneyed) class from a venal group to a

collectivity-oriented elite and who, fusing with it, will forge a new, genteel elite continuous with but better than the past: Talcott Parsons.¹⁴

(*Sed contra*: Neither group is an especially morally bound agent; the old class is constrained to protect its profits, the New Class is cashing in on its education. Immersed in the present, this view misses the fact that each is ready to exploit the other, if need be, and shows little understanding of the profound [if different] limits imposed on the rationality and morality of each of these groups, and of the important tensions between them.)

4. *New Class as Servants of Power*: Here the New Class is viewed as subservient to the old (moneyed) class which is held to retain power much as it always did, and is simply using the New Class to maintain its domination of society: Noam Chomsky¹⁵ and Maurice Zeitlin.¹⁶

(*Sed contra*: This ignores the revolutionary history of the twentieth century in which radicalized elements of the New Class played a major leadership role in the key revolutions of our time. It greatly overemphasizes the common interests binding the New and old class, systematically missing the tensions between them; it ignores the fact that elimination of the old class is an historical option open to the New Class. This static conception underestimates the growth in the numbers and influence of the New Class. The view is also unexpectedly Marcusean in overstressing the prospects of old class continuity; it really sees the old class as having no effective opponents, either in the New Class or in the old adversary class, the proletariat. It thus ends as seeing even less social change in prospect than the Parsonian view [#3 above].)

5. *New Class as Flawed Universal Class (my own view)*: The New Class is elitist and self-seeking and uses its special knowledge to advance its own interests and power, and to control its own work situation. Yet the New Class may also be the best card that history

has presently given us to play. The power of the New Class is growing. It is substantially more powerful and independent than Chomsky suggests, while still much less powerful than is suggested by Galbraith who seems to conflate present reality with future possibility. The power of this morally ambiguous New Class is on the ascendent and it holds a mortgage on at least *one* historical future.

In my own left Hegelian sociology, the New Class bearers of knowledge are seen as an embryonic new “universal class”—as the prefigured embodiment of such future as the working class still has. It is that part of the working class which will survive cybernation. At the same time, a left Hegelian sociology also insists that the New Class is profoundly flawed as a universal class. Moreover, the New Class is not some unified subject or a seamless whole; it, too, has its own internal contradictions. It is a class internally divided with tensions between (technical) intelligentsia and (humanistic) intellectuals. No celebration, mine is a critique of the New Class which does not view its growing power as inevitable, which sees it as morally ambivalent, embodying the collective interest but partially and transiently, while simultaneously cultivating its own guild advantage. . . .

The New Class as a Cultural Bourgeoisie

1. The New Class and the old class are at first undifferentiated; the New Class commonly originates in classes with property advantages, that is, in the old class, or is sponsored by them. The New Class of intellectuals and intelligentsia are the relatively more *educated* counterpart—often the brothers, sisters, or children—of the old moneyed class. Thus the New Class contest sometimes has the character of a *civil war within the upper classes*. It is the differentiation of the old class into contentious factions. To understand the New Class contest it is vital to understand how the *privileged* and advantaged, not sim-

ply the suffering, come to be alienated from the very system that privileges them.

2. The “non-negotiable” objectives of the old moneyed class are to reproduce their capital, at a minimum, but, preferably, to make it accumulate and to appropriate profit: $M-C-M'$, as Marx said. This is done within a structure in which all of them must compete with one another. This unrelenting competition exerts pressure to rationalize their productive and administrative efforts and unceasingly to heighten efficiency. (Marx called it, “revolutionizing” production.) But this rationalization is dependent increasingly on the efforts of the New Class intelligentsia and its expert skills. It is inherent in its structural situation, then, that the old class must bring the New Class into existence.

3. Much of the New Class is at first trained under the direct control of the old class’s firms or enterprises. Soon, however, the old class is separated from the reproduction of the New Class by the emergence and development of a public system of education whose costs are “socialized.”¹⁷

4. The more that the New Class’s reproduction derives from specialized systems of public education, the more the New Class develops an ideology that stresses its *autonomy*, its separation from and presumable independence of “business” or political interests. This autonomy is said to be grounded in the specialized knowledge or cultural capital transmitted by the educational system, along with an emphasis on the obligation of educated persons to attend to the welfare of the collectivity. In other words, the *ideology* of “professionalism” emerges.

5. Professionalism is one of the public *ideologies* of the New Class, and is the genteel subversion of the old class by the new. Professionalism is a phase in the historical development of the “collective consciousness” of the New Class. While not overtly a critique of the old class, professionalism is a tacit claim by the New Class to *technical and moral superiority* over the old class, implying

that the latter lack technical credentials and are guided by motives of commercial venality. Professionalism silently installs the New Class as the paradigm of virtuous and legitimate authority, performing with technical skill and with dedicated concern for the society-at-large. Professionalism makes a focal claim for the legitimacy of the New Class which tacitly deauthorizes the old class.

On the one side, this is a bid for prestige *within* the established society; on the other, it tacitly presents the New Class as an *alternative* to the old. In asserting its own claims to authority, professionalism in effect *devalues the authority of the old class*.

6. The special privileges and powers of the New Class are grounded in their *individual* control of special cultures, languages, techniques, and of the skills resulting from these. The New Class is a cultural bourgeoisie who appropriates privately the advantages of a historically and collectively produced cultural capital. Let us be clear, then: the New Class is not just *like* the old class; its special culture is not just *like* capital. No metaphor is intended. The special culture of the New Class *is* a stock of capital that generates a stream of income (some of) which it appropriates privately.

7. The fundamental objectives of the New Class are: to increase its own share of the national product; to produce and reproduce the special social conditions enabling them to appropriate privately larger shares of the incomes produced by the special cultures they possess; to control their work and their work settings; and to increase their political power partly in order to achieve the foregoing. The struggle of the New Class is, therefore, to *institutionalize a wage system*, i.e., a social system with a distinct principle of distributive justice: “from each according to his ability, to each according to his work,” which is also the norm of “socialism.” Correspondingly, the New Class may oppose other social systems and their different systems of privilege, for example, systems that allocate privileges and incomes on the basis of controlling

stocks of money (i.e., old capital). The New Class, then, is prepared to be egalitarian so far as the privileges of the *old* class are concerned. That is, under certain conditions it is prepared to remove or restrict the special incomes of the old class: profits, rents, interest. The New Class is anti-egalitarian, however, in that it seeks special guild advantages—political powers and incomes—on the basis of its possession of cultural capital. . . .

The New Class as a Speech Community

1. The culture of critical discourse (CCD)¹⁸ is an historically evolved set of rules, a grammar of discourse, which (1) is concerned to *justify* its assertions, but (2) whose *mode* of justification does not proceed by invoking authorities, and (3) prefers to elicit the *voluntary* consent of those addressed solely on the basis of arguments adduced. CCD is centered on a specific speech act: justification. It is a culture of discourse in which there is nothing that speakers will on principle permanently refuse to discuss or make problematic; indeed, they are even willing to talk about the value of talk itself and its possible inferiority to silence or to practice. This grammar is the deep structure of the common ideology shared by the New Class. *The shared ideology of the intellectuals and intelligentsia is thus an ideology about discourse*. Apart from and underlying the various technical languages (or sociolects) spoken by specialized professions, intellectuals and intelligentsia are commonly committed to a culture of critical discourse (CCD). CCD is the latent but mobilizable infrastructure of modern “technical” languages.

2. The culture of critical discourse is characterized by speech that is *relatively* more *situation-free*, more context or field “independent.” This speech culture thus values expressly legislated meanings and devalues tacit, context-limited meanings. Its ideal is: “one word, one meaning,” for everyone and forever.

The New Class's special speech variant also stresses the importance of particular modes of *justification*, using especially explicit and articulate rules, rather than diffuse precedents or tacit features of the speech context. The culture of critical speech requires that the validity of claims be justified without reference to the speaker's *societal position or authority*. Here, good speech is speech that can make its own principles *explicit* and is oriented to conforming with them, rather than stressing context-sensitivity and context-variability. Good speech here thus has *theoreticity*.¹⁹

Being pattern- and principle-oriented, CCD implies that that which is said may *not* be correct, and may be *wrong*. It recognizes that "What Is" may be mistaken or inadequate and is therefore open to alternatives. CCD is also relatively more *reflexive*, self-monitoring, capable of more metacommunication, that is, of talk about talk; it is able to make its own speech problematic, and to edit it with respect to its lexical and grammatical features, as well as making problematic the validity of its assertions. CCD thus requires considerable "expressive discipline," not to speak of "instinctual renunciation."

3. Most importantly, the culture of critical speech forbids reliance upon the speaker's person, authority, or status in society to justify his claims. As a result, CCD de-authorizes all speech grounded in traditional societal authority, while it authorizes itself, the elaborated speech variant of the culture of critical discourse, as the standard of *all* "serious" speech. From now on, persons and their social positions must not be visible in their speech. Speech becomes impersonal. Speakers hide behind their speech. Speech seems to be dis-embodied, de-contextualized and self-grounded. (This is especially so for the speech of intellectuals and somewhat less so for technical intelligentsia who may not invoke CCD except when their paradigms break down.) The New Class becomes the guild masters of an invisible pedagogy.

4. The culture of critical discourse is the common ideology shared by the New Class, although technical intelligentsia sometimes keep it in latency. The skills and the social conditions required to reproduce it are among the common *interests* of the New Class. Correspondingly, it is in the common interest of the New Class to prevent or oppose all censorship of its speech variety and to install it as the standard of good speech. *The New Class thus has both a common ideology in CCD and common interests in its cultural capital. . . .*

The Flawed Universal Class

1. The New Class is the most progressive force in modern society and is a center of whatever human emancipation is possible in the foreseeable future. It has no motives to curtail the forces of production and no wish to develop them solely in terms of their profitability. The New Class possesses the scientific knowledge and technical skills on which the future of modern forces of production depend. At the same time, members of the New Class also manifest increasing sensitivity to the ecological "side effects" or distant diseconomies of continuing technical development. The New Class, further, is a center of opposition to almost all forms of censorship, thus embodying a universal societal interest in a kind of rationality broader than that invested in technology. Although the New Class is at the center of nationalist movements throughout the world, after that phase is secured, the New Class is also the most internationalist and most universalist of all social strata; it is the most cosmopolitan of all elites. Its control over ordinary "foreign" languages, as well as of technical sociolects, enables it to communicate with other nationalities and it is often a member of a technical guild of international scope.

2. For all that, however, the New Class is hardly the end of domination. While its ultimate significance is the end of the old moneyed class's domination, the New Class is also

the nucleus of a *new* hierarchy and the elite of a new form of cultural capital.

The historical limits of the New Class are inherent in both the nature of its own characteristic rationality, and in its ambitions as a cultural bourgeoisie. Its culture of critical discourse fosters a purely “theoretical” attitude toward the world. Speakers are held competent to the degree that they know and can *say* the rules, rather than just happening to follow them. The culture of critical discourse thus values the very theoreticity that the “common sense” long suspected was characteristic of intellectuals.

Intellectuals have long believed that those who know the rule, who know the theory by which they act, are superior because they lead an “examined” life. They thus exalt theory over practice, and are concerned less with the success of a practice than that the practice should have submitted itself to a reasonable rule. Since intellectuals and intelligentsia are concerned with doing things in the right way and for the right reason—in other words, since they value doctrinal conformity for its own sake—they (we) have a native tendency toward ritualism and *sectarianism*.

3. The culture of the New Class exacts still other costs: since its discourse emphasizes the importance of carefully edited speech, this has the vices of its virtues: in its *virtuous* aspect, self-editing implies a commendable circumspection, carefulness, self-discipline and “seriousness.” In its negative modality, however, self-editing also disposes toward an unhealthy self-consciousness, toward stilted convoluted speech, an inhibition of play, imagination and passion, and continual pressure for expressive discipline. The new rationality thus becomes the source of a new alienation.

Calling for watchfulness and self-discipline, CCD is productive of intellectual reflexivity *and* the loss of warmth and spontaneity. Moreover, that very reflexivity stresses the importance of adjusting action to some pattern of propriety. There is, therefore, a structured inflexibility when facing

changing situations; there is a certain disregard of the differences in situations, and an insistence on hewing to the required rule.

This inflexibility and insensitivity to the force of differing contexts, this inclination to impose one set of rules on different cases also goes by the ancient name of “dogmatism.” Set in the context of human relationships, the vulnerability of the New Class to dogmatism along with its very *task*-centeredness, imply a certain insensitivity to *persons*, to their feelings and reactions, and open the way to the disruption of human solidarity. Political brutality, then, finds a grounding in the culture of critical discourse; the new rationality may paradoxically allow a new darkness at noon.

4. The paradox of the New Class is that it is both emancipatory *and* elitist. It subverts all establishments, social limits, and privileges, including its own. The New Class bears a culture of critical and careful discourse which is an historically emancipatory rationality. The new discourse (CCD) is the grounding for a critique of established forms of domination and provides an escape from tradition, but it also bears the seeds of a new domination. Its discourse is a lumbering machinery of argumentation that can wither imagination, discourage play, and curb expressivity. The culture of discourse of the New Class seeks to *control* everything, its topic and itself, believing that such domination is the only road to truth. The New Class begins by monopolizing truth and by making itself its guardian. It thereby makes even the claims of the old class dependent on it. The New Class sets itself above others, holding that its speech is better than theirs; that the examined life (*their* examination) is better than the unexamined life which, it says, is sleep and no better than death. Even as it subverts old inequities, the New Class silently inaugurates a new hierarchy of the knowing, the knowledgeable, the reflexive and insightful. Those who talk well, it is held, excel those who talk poorly or not at all. It is now no longer enough simply to be good. Now, one

has to explain it. The New Class is the universal class in embryo, but badly flawed.

NOTES

1. It is not my intention to suggest that modern intellectuals are merely the secular counterpart of clericals. Indeed, my own stress (as distinct, say, from Edward Shils who does appear to view intellectuals as priests *manqués*) is on the discontinuity of the two.

2. For full development of this, see chapter 2, especially p. 42, of my *Dialectic of Ideology and Technology* (New York, 1976).

3. Doubtless some will insist this is a “false consciousness.” But this misses the point. My concern here is with their own definitions of their social role, precisely because these influence the manner in which they perform their roles. As W. I. Thomas and Florian Znaniecki long ago (and correctly) insisted, a thing defined as real is real in its consequences. Moreover, the state who employs most of these teachers is itself interested in having teachers consolidate the tie between students and it itself, rather than with the students’ parents.

4. See Basil Bernstein, *Class, Codes and Control*, vol. 1, *Theoretical Studies Towards a Sociology of Language* (London, 1971), vol. 2, *Applied Studies Towards a Sociology of Language* (London, 1973), vol. 3, *Towards a Theory of Educational Transmission* (London, 1975). Bernstein’s theory is used here in a critical appropriation facilitated by the work of Dell Hymes and William Labov. My own critique of Bernstein emerges, at least tacitly, in the discussion of [the “Flawed Universal Class”] in the text. It is developed explicitly in my *Dialectic of Ideology and Technology*, pp. 58–66. While Labov has sharply criticized Bernstein, he himself also stresses the general importance of self-monitored speech and of speech *reflexivity* in general (i.e., not only of careful pronunciation) thus converging with Bernstein’s focus on reflexivity as characterizing the elaborated linguistic variant and distinguishing it from the restricted variant. See William Labov, *Sociolinguistic Patterns* (Philadelphia, 1972), p. 208.

5. For example: “The Communists disdain to conceal their views and aims. They openly declare . . .” (*Communist Manifesto* [Chicago, 1888], authorized English edition edited by Engels, p. 58).

6. See E. Hobsbawm, *Primitive Rebels* (Manchester, 1959), p. 167 ff.

7. A secret doctrine is one which, because it is reserved only for the organization elite, can be made known only after persons join organizations and

reach a certain membership position in it. A secret doctrine thus is never one which can have been a *motive* for joining the organization in the first instance.

8. Lenin’s *What Is to Be Done?* was originally published in 1902.

9. *The New Industrial State* (Boston, 1967).

10. *The Coming of Post-Industrial Society* (New York, 1973).

11. *The Modern Corporation and Private Property* (New York, 1932).

12. “It stands to reason that the one who knows more will dominate the one who knows less,” M. Bakounine, *Oeuvres*, Vol. 5 (Paris, 1911), p. 106.

13. See V. F. Calverton, *The Making of Society* (New York, 1937).

14. Talcott Parsons, *The Social System* (Glencoe, 1951), chapter 10; *Essays in Sociological Theory* (Glencoe, 1954), chapter 18; “The Professions,” *International Encyclopedia of Social Sciences* (New York, 1968).

15. While Chomsky’s position is exhibited in various of his writings, I shall rely here on his most recent statement in his Huizinga lecture, “Intellectuals and the State,” delivered at Leiden, 9 October 1977. Citations will be from the manuscript copy. Cf. N. Chomsky, *American Power and the New Mandarins* (New York, 1969).

16. Maurice Zeitlin, “Corporate Ownership and Control: The Large Corporations and the Capitalist Class,” *American Journal of Sociology* (March 1974), pp. 1073–1119.

17. Cf. James O’Connor, *Corporations and the State* (New York, 1974), pp. 126–28 for the argument that government financing of R & D and advanced education constitute a socialization of part of the costs of production whose net surplus is privately appropriated.

18. This section is indebted to Basil Bernstein and is based on a critical appropriation of his “elaborated and restricted linguistic codes,” which have gone through various re-workings. That controversial classic was published in J. J. Gumperz and D. Hymes, *Directions in Sociolinguistics* (New York, 1972). A recent re-working is to be found in Bernstein’s “Social Class, Language, and Socialization,” in T. A. Sebeok, ed., *Current Trends in Linguistics* (The Hague, 1974). For full bibliographic and other details see note 4 above.

19. Cf. Peter McHugh, “A Common-Sense Perception of Deviance,” in H. P. Dreitzel, ed., *Recent Sociology*, Number 2 (London, 1970), pp. 165 ff. For good speech as “serious” speech see David Silverman, “Speaking Seriously,” *Theory and Society* (Spring, 1974).

33 ■ David Brooks

Bobos in Paradise**The New Upper Class and How They Got There****The Rise of the Educated Class**

I'm not sure I'd like to be one of the people featured on the *New York Times* wedding page, but I know I'd like to be the father of one of them. Imagine how happy Stanley J. Kogan must have been, for example, when his daughter Jamie was admitted to Yale. Then imagine his pride when Jamie made Phi Beta Kappa and graduated summa cum laude. Stanley himself is no slouch in the brains department: he's a pediatric urologist in Croton-on-Hudson, with teaching positions at the Cornell Medical Center and the New York Medical College. Still, he must have enjoyed a gloat or two when his daughter put on that cap and gown.

And things only got better. Jamie breezed through Stanford Law School. And then she met a man—Thomas Arena—who appeared to be exactly the sort of son-in-law that pediatric urologists dream about. He did his undergraduate work at Princeton, where he, too, made Phi Beta Kappa and graduated summa cum laude. And he, too, went to law school, at Yale. After school they both went to work as assistant U.S. attorneys for the mighty Southern District of New York.

These two awesome résumés collided at a wedding ceremony in Manhattan, and given all the school chums who must have attended, the combined tuition bills in that room must have been staggering. The rest of us got to read about it on the *New York*

Times weddings page. The page is a weekly obsession for hundreds of thousands of *Times* readers and aspiring Balzacs. Unabashedly elitist, secretive, and totally honest, the “mergers and acquisitions page” (as some of its devotees call it) has always provided an accurate look at at least a chunk of the American ruling class. And over the years it has reflected the changing ingredients of elite status.

When America had a pedigreed elite, the page emphasized noble birth and breeding. But in America today it's genius and geniality that enable you to join the elect. And when you look at the *Times* weddings page, you can almost feel the force of the mingling SAT scores. It's Dartmouth marries Berkeley, MBA weds Ph.D., Fulbright hitches with Rhodes, Lazard Frères joins with CBS, and summa cum laude embraces summa cum laude (you rarely see a summa settling for a magna—the tension in such a marriage would be too great). The *Times* emphasizes four things about a person—college degrees, graduate degrees, career path, and parents' profession—for these are the markers of upscale Americans today. . . .

The Fifties

The *Times* weddings page didn't always pulse with the accomplishments of the Résumé Gods. In the late 1950s, the page pro-

jected a calm and more stately ethos. The wedding accounts of that era didn't emphasize jobs or advanced degrees. The profession of the groom was only sometimes mentioned, while the profession of the bride was almost never listed (and on the rare occasions when the bride's profession was noted, it was in the past tense, as if the marriage would obviously end her career). Instead, the *Times* listed pedigree and connections. Ancestors were frequently mentioned. The ushers were listed, as were the bridesmaids. Prep schools were invariably mentioned, along with colleges. The *Times* was also careful to list the groom's clubs—the Union League, the Cosmopolitan Club. It also ran down the bride's debutante history, where she came out, and whatever women's clubs she might be a member of, such as the Junior League. In short, the page was a galaxy of restricted organizations. . . .

The section from the late fifties evokes an entire milieu that was then so powerful and is now so dated: the network of men's clubs, country clubs, white-shoe law firms, oak-paneled Wall Street firms, and WASP patriarchs. Everybody has his or her own mental images of the old Protestant Establishment: lockjaw accents, the Social Register, fraternity jocks passing through Ivy League schools, constant rounds of martinis and highballs, bankers' hours, starched old men like Averell Harriman, Dean Acheson, and John J. McCloy, the local bigwigs that appear in John Cheever and John O'Hara stories. . . .

It really was possible to talk about an aristocratic ruling class in the fifties and early sixties, a national elite populated by men who had gone to northeastern prep schools like Groton, Andover, Exeter, and St. Paul's and then ascended through old-line firms on Wall Street into the boardrooms of the Fortune 500 corporations and into the halls of Washington power. The WASPs didn't have total control of the country or anything like it, but they did have the hypnotic magic of

prestige. As Richard Rovere wrote in a famous 1962 essay entitled "The American Establishment," "It has very nearly unchallenged power in deciding what is and what is not respectable opinion in this country." . . .

This was the last great age of socially acceptable boozing. It was still an era when fox hunting and polo didn't seem antiquarian. But the two characteristics of that world that strike us forcefully today are its unabashed elitism and its segregation. Though this elite was nowhere near as restrictive as earlier elites—World War II had exerted its leveling influence—the 1950s establishment was still based on casual anti-Semitism, racism, sexism, and a thousand other silent barriers that blocked entry for those without the correct pedigree. Wealthy Jewish and Protestant boys who had been playing together from childhood were forced to endure "The Great Division" at age 17, when Jewish and Gentile society parted into two entirely separate orbits, with separate debutante seasons, dance schools, and social secretaries. A Protestant business executive may have spent his professional hours working intimately with his Jewish colleague, but he never would have dreamed of putting him up for membership in his club. When Senator Barry Goldwater attempted to play golf at the restricted Chevy Chase Club, he was told the club was restricted. "I'm only half Jewish, so can't I play nine holes?" he is said to have replied.

The WASP elite was also genially anti-intellectual. Its members often spoke of "eggheads" and "highbrows" with polite disdain. Instead, their status, as F. Scott Fitzgerald had pointed out a few decades before, derived from "animal magnetism and money." By contrast with today's ruling class, they had relatively uncomplicated attitudes about their wealth. They knew it was vulgar to be gaudy, they tended toward thriftiness, but they seem not to have seen their own money as an affront to American principles of equality. On the contrary, most took their elite status for granted, assuming that such position

was simply part of the natural and beneficent order of the universe. There was always going to be an aristocracy, and so for the people who happened to be born into it, the task was to accept the duties that came along with its privileges. . . .

The Hinge Years

Then came the change. By 1960 the average verbal SAT score for incoming freshmen at Harvard was 678, and the math score was 695—these are stratospheric scores. The average Harvard freshman in 1952 would have placed in the bottom 10 percent of the Harvard freshman class of 1960. Moreover, the 1960 class was drawn from a much wider socioeconomic pool. Smart kids from Queens or Iowa or California, who wouldn't have thought of applying to Harvard a decade earlier, were applying and getting accepted. Harvard had transformed itself from a school catering mostly to the northeastern social elite to a high-powered school reaching more of the brightest kids around the country. And this transformation was replicated in almost all elite schools. . . .

History, as Pareto once remarked, is the graveyard of aristocracies, and by the late fifties and early sixties the WASP Establishment had no faith in the code—and the social restrictions—that had sustained it. Maybe its members just lost the will to fight for their privileges. As the writer David Frum theorizes, it had been half a century since the last great age of fortune making. The great families were into at least their third genteel generation. Perhaps by then there wasn't much vigor left. Or perhaps it was the Holocaust that altered the landscape by discrediting the sort of racial restrictions that the Protestant Establishment was built on.

In any case, in 1964 Digby Baltzell astutely perceived the crucial trends. “What seems to be happening,” he wrote in *The Protestant Establishment*, “is that a scholarly

hierarchy of campus communities governed by the values of admissions committees is gradually supplanting the class hierarchies of local communities which are still governed by the values of parents. . . . Just as the hierarchy of the Church was the main avenue of advancement for the talented and ambitious youth from the lower orders during the medieval period, and just as the business enterprise was responsible for the nineteenth century rags-to-riches dream (when we were predominantly an Anglo-Saxon country), so the campus community has now become the principal guardian of our traditional opportunistic ideals.”

The campus gates were thus thrown open on the basis of brains rather than blood, and within a few short years the university landscape was transformed. Harvard, as we've seen, was changed from a school for the well-connected to a school for brainy strivers. The remaining top schools eliminated their Jewish quotas and eventually dropped their restrictions on women. Furthermore, the sheer numbers of educated Americans exploded. The portion of Americans going to college had been rising steadily throughout the 20th century, but between 1955 and 1974 the growth rate was off the charts. Many of the new students were women. Between 1950 and 1960 the number of female students increased by 47 percent. It then jumped by an additional 168 percent between 1960 and 1970. Over the following decades the student population kept growing and growing. In 1960 there were about 2,000 institutions of higher learning. By 1980 there were 3,200. In 1960 there were 235,000 professors in the United States. By 1980 there were 685,000. . . .

The Sixties

The educated-class rebellion we call “the sixties” was about many things, some of them important and related to the Civil Rights movement and Vietnam, some of them en-

tirely silly, and others, like the sexual revolution, overblown (actual sexual behavior was affected far more by the world wars than by the Woodstock era). But at its core the cultural radicalism of the sixties was a challenge to conventional notions of success. It was not only a political effort to dislodge the establishment from the seats of power. It was a cultural effort by the rising members of the privileged classes to destroy whatever prestige still attached to the WASP lifestyle and the WASP moral code, and to replace the old order with a new social code that would celebrate spiritual and intellectual ideals. The sixties radicals rejected the prevailing definition of accomplishment, the desire to keep up with the Joneses, the prevailing idea of social respectability, the idea that a successful life could be measured by income, manners, and possessions. . . .

And Then Comes Money

The hardest of the hard-core sixties radicals believed the only honest way out was to reject the notion of success altogether: drop out of the rat race, retreat to small communities where real human relationships would flourish. But that sort of utopianism was never going to be very popular, especially among college grads. Members of the educated class prize human relationships and social equality, but as for so many generations of Americans before them, achievement was really at the core of the sixties grads' value system. They were meritocrats, after all, and so tended to define themselves by their accomplishments. Most of them were never going to drop out or sit around in communes smelling flowers, raising pigs, and contemplating poetry. Moreover, as time went by, they discovered that the riches of the universe were lying at their feet.

At first, when the great hump of baby boom college graduates entered the workforce, having a college degree brought few fi-

nancial rewards or dramatic life changes. As late as 1976, the labor economist Richard Freeman could write a book called *The Overeducated American*, arguing that higher education didn't seem to be paying off in the marketplace. But the information age kicked in, and the rewards for education grew and grew. In 1980, according to labor market specialist Kevin Murphy of the University of Chicago, college graduates earned roughly 35 percent more than high school graduates. But by the mid-1990s, college graduates were earning 70 percent more than high school graduates, and those with graduate degrees were earning 90 percent more. The wage value of a college degree had doubled in 15 years.

The rewards for intellectual capital have increased while the rewards for physical capital have not. That means that even liberal arts majors can wake up one day and find themselves suddenly members of the top-income brackets. A full professor at Yale who renounced the capitalist rat race finds himself making, as of 1999, \$113,100, while a professor at Rutgers pulls in \$103,700 and superstar professors, who become the object of academic bidding wars, now can rake in more than \$300,000 a year. Congressional and presidential staffers top out at \$125,000 (before quintupling that when they enter the private sector), and the journalists at national publications can now count on six-figure salaries when they hit middle age, not including lecture fees. Philosophy and math majors head for Wall Street and can make tens of millions of dollars from their quantitative models. America has always had a lot of lawyers, and now the median income for that burgeoning group is \$72,500, while income for the big-city legal grinds can reach seven figures. And super-students still flood into medicine—three-quarters of private practitioners net more than \$100,000. Meanwhile, in Silicon Valley there are more millionaires than people. . . .

The Anxieties of Abundance

Those who want to win educated-class approval must confront the anxieties of abundance: how to show—not least to themselves—that even while climbing toward the top of the ladder they have not become all the things they still profess to hold in contempt. How to navigate the shoals between their affluence and their self-respect. How to reconcile their success with their spirituality, their elite status with their egalitarian ideals. Socially enlightened members of the educated elite tend to be disturbed by the widening gap between rich and poor and are therefore made somewhat uncomfortable by the fact that their own family income now tops \$80,000. Some of them dream of social justice yet went to a college where the tuition costs could feed an entire village in Rwanda for a year. Some once had “Question Authority” bumper stickers on their cars but now find themselves heading start-up software companies with 200 people reporting to them. The sociologists they read in college taught that consumerism is a disease, and yet now they find themselves shopping for \$3,000 refrigerators. They took to heart the lessons of *Death of a Salesman*, yet now find themselves directing a sales force. They laughed at the plastics scene in *The Graduate* but now they work for a company that manufactures . . . plastic. Suddenly they find themselves moving into a suburban house with a pool and uncomfortable about admitting it to their bohemian friends still living downtown. . . .

The Reconcilers

The grand achievement of the educated elites in the 1990s was to create a way of living that lets you be an affluent success and at the same time a free-spirit rebel. Founding design firms, they find a way to be an artist and still qualify for stock options. Building

gourmet companies like Ben & Jerry’s or Nantucket Nectars, they’ve found a way to be dippy hippies and multinational corporate fat cats. Using William S. Burroughs in ads for Nike sneakers and incorporating Rolling Stones anthems into their marketing campaigns, they’ve reconciled the antiestablishment style with the corporate imperative. Listening to management gurus who tell them to thrive on chaos and unleash their creative potential, they’ve reconciled the spirit of the imagination with service to the bottom line. Turning university towns like Princeton and Palo Alto into entrepreneurial centers, they have reconciled the highbrow with the high tax bracket. Dressing like Bill Gates in worn chinos on his way to a stockholders’ meeting, they’ve reconciled undergraduate fashion with upper-crust occupations. Going on eco-adventure vacations, they’ve reconciled aristocratic thrill-seeking with social concern. Shopping at Benetton or the Body Shop, they’ve brought together consciousness-raising and cost control.

When you are amidst the educated up-scalers, you can never be sure if you’re living in a world of hippies or stockbrokers. In reality you have entered the hybrid world in which everybody is a little of both.

Marx told us that classes inevitably conflict, but sometimes they just blur. The values of the bourgeois mainstream culture and the values of the 1960s counterculture have merged. That culture war has ended, at least within the educated class. In its place that class has created a third culture, which is a reconciliation between the previous two. The educated elites didn’t set out to create this reconciliation. It is the product of millions of individual efforts to have things both ways. But it is now the dominant tone of our age. In the resolution between the culture and the counterculture, it is impossible to tell who co-opted whom, because in reality the bohemians and the bourgeois co-opted each

other. They emerge from this process as bourgeois bohemians, or Bobos.

The New Establishment

Today the *New York Times* weddings section is huge once again. In the early 1970s the young rebels didn't want to appear there, but now that their own kids are in college and getting married, they are proud to see their offspring in the Sunday paper. For a fee the *Times* will send you a reproduction of your listing, suitable for framing.

And the young people, the second-generation Bobos, are willing to see their nuptials recorded. Look at the newlyweds on any given Sunday morning, beaming out at you from the pages of the *Times*. Their smiles seem so genuine. They all look so nice and approachable, not dignified or fearsome, the way some of the brides on the 1950s pages did. Things are different but somehow similar. . . .

Today's establishment is structured differently. It is not a small conspiracy of well-bred men with interlocking family and school ties who have enormous influence on the levers of power. Instead, this establishment is a large, amorphous group of meritocrats who share a consciousness and who unself-consciously reshape institutions to accord with their values. They are not confined to a few East Coast institutions. In 1962, Richard Rovere could write, "Nor has the Establishment ever made much headway in such fields as advertising, television or motion pictures." Today's establishment is everywhere. It exercises its power subtly, over ideas and concepts, and therefore pervasively. There are no sure-fire demographic markers to tell who is a member of this establishment. Members tend to have gone to competitive colleges, but not all have. They tend to live in upscale neighborhoods, such as Los Altos, California, and Bloomfield, Michigan, and Lincoln Park, Illinois, but

not all do. What unites them is their shared commitment to the Bobo reconciliation. People gain entry into the establishment by performing a series of delicate cultural tasks: they are prosperous without seeming greedy; they have pleased their elders without seeming conformist; they have risen toward the top without too obviously looking down on those below; they have achieved success without committing certain socially sanctioned affronts to the ideal of social equality; they have constructed a prosperous lifestyle while avoiding the old clichés of conspicuous consumption (it's OK to hew to the new clichés).

Obviously, none of this is to suggest that all members of the new Bobo establishment think alike, any more than it's true to say that all members of any establishment think alike. Some of the bourgeois bohemians are more on the bourgeois side; they are stockbrokers who happen to like artists' lofts. Some are on the bohemian side; they are art professors who dabble in the market. Nonetheless, if you look at some quintessential figures of the new establishment—such as Henry Louis Gates, Charlie Rose, Steven Jobs, Doris Kearns Goodwin, David Geffen, Tina Brown, Maureen Dowd, Jerry Seinfeld, Stephen Jay Gould, Lou Reed, Tim Russert, Steve Case, Ken Burns, Al Gore, Bill Bradley, John McCain, George W. Bush—you can begin to sense a common ethos that mingles 1960s rebellion with 1980s achievement. You can feel the Bobo ethos, too, in the old institutions that have been taken over by the new establishment, such as the *New Yorker*, Yale University, the American Academy of Arts and Letters (which now includes people like Toni Morrison, Jules Feiffer, and Kurt Vonnegut among its members), or the *New York Times* (which now runs editorials entitled "In Praise of the Counterculture"). You can sense the ethos with special force in the new establishment institutions that would have been alien to

the old elite: NPR, DreamWorks, Microsoft, AOL, Starbucks, Yahoo, Barnes & Noble, Amazon, and Borders.

And over the past few years, this new educated establishment has begun to assume the necessary role of an establishment. That is to say, it has begun to create a set of social codes that give coherent structure to national life. Today, America once again has a dominant class that defines the parameters of respectable opinion and taste—a class that determines conventional wisdom, that promulgates a code of good manners, that establishes a pecking order to give shape to society, that excludes those who violate its codes, that transmits its moral and etiquette codes down to its children, that imposes social discipline on the rest of society so as to improve the “quality of life,” to use the contemporary phrase. . . .

Class Rank

This has got to be one of the most anxious social elites ever. We Bobos are not anxious because there is an angry mob outside the gates threatening to send us to the guillotine. There isn't. The educated elite is anxious because its members are torn between their drive to succeed and their fear of turning into sellouts. Furthermore, we are anxious because we do not award ourselves status sinecures. Previous establishments erected social institutions that would give their members security. In the first part of the 20th century, once your family made it into the upper echelons of society, it was relatively easy to stay there. You were invited on the basis of your connections to the right affairs. You were admitted, nearly automatically, to the right schools and considered appropriate for the right spouses. The pertinent question in those circles was not what do you do, but who are you. Once you were established as a

Biddle or an Auchincloss or a Venderlip, your way was clear. But members of today's educated class can never be secure about their own future. A career crash could be just around the corner. In the educated class even social life is a series of aptitude tests; we all must perpetually perform in accordance with the shifting norms of propriety, ever advancing signals of cultivation. Reputations can be destroyed by a disgraceful sentence, a lewd act, a run of bad press, or a terrible speech at the financial summit at Davos.

And more important, members of the educated class can never be secure about their children's future. The kids have some domestic and educational advantages—all those tutors and developmental toys—but they still have to work through school and ace the SATs just to achieve the same social rank as their parents. Compared to past elites, little is guaranteed.

The irony is that all this status insecurity only makes the educated class stronger. Its members and their children must constantly be alert, working and achieving. Moreover, the educated class is in no danger of becoming a self-contained caste. Anybody with the right degree, job, and cultural competencies can join. Marx warned that “the more a ruling class is able to assimilate the most prominent men [or women] of the dominated classes, the more stable and dangerous its rule.” And in truth it is hard to see how the rule of the meritocrats could ever come to an end. The WASP Establishment fell pretty easily in the 1960s. It surrendered almost without a shot. But the meritocratic Bobo class is rich with the spirit of self-criticism. It is flexible and amorphous enough to co-opt that which it does not already command. The Bobo meritocracy will not be easily toppled, even if some group of people were to rise up and conclude that it should be.

34 ■ Gil Eyal, Iván Szelényi, and Eleanor Townsley

Post-Communist Managerialism

The unique feature of making capitalism from the ruins of state socialism in Central Europe is that it is happening without a propertied bourgeoisie. In all other historical sites where modern capitalism has developed, some form of private property and some class of private proprietors—no matter how embryonic, and no matter how different from modern capitalist entrepreneurs—already existed. In the classical case of transition, feudal landlords gradually converted their property into private ownership and began to be recruited into the new *grande bourgeoisie*. Urban artisans and merchants were busily accumulating capital, and were well positioned to transform themselves from the third estate of a feudal order into one of the fractions of the new dominant class in a capitalist mode of production. Post-communism is the first situation where the transition to private property from a collective form of ownership is being attempted. Moreover, this project is being led by the second *Bildungsbürgertum*—by an uneasy alliance between former communist apparatchiks, technocrats, managers and their former left-wing critics, the dissident intellectuals. In short, capitalism is being made by a coalition of propertyless agents, who only yesterday outbid each other in their anticapitalism.

In 1988 two leading sociologists of the region, Jadwiga Staniszkis and Elemér Hankiss, formulated a provocative hypothesis, which [may be termed] the theory of political capitalism.¹ According to Staniszkis and Hankiss, the former communist *nomenklatura* knew by

1988 that the destruction of the old communist order was inevitable. They therefore designed a scheme to convert political office into private wealth, and attempted to transform themselves into a new *grande bourgeoisie*. Indeed, many commentators on the Central European transformation think that this is what happened after 1989: communist officials used political office to convert public goods into private individual wealth—*de facto*, they stole state property and became a ‘kleptocracy’ . . .

The second theory we confront is Erzsébet Szalai’s theory of ‘technocratic revolution,’ which in some ways, can be viewed as a refinement of the general political capitalism thesis.² Szalai’s argument is that the late state-socialist *nomenklatura* was highly fragmented, and that the dynamics of social change should be understood as an intense struggle between the bureaucratic and technocratic fractions of the old ruling estate. In this view, 1989 was a successful revolution of the late state-socialist technocracy against the bureaucratic fraction of the communist ruling estate. . . . Rather than the *nomenklatura* as a whole grasping power, one of its fractions—the technocratic-managerial elite—appeared to have established itself as the new propertied class. . . .

In 1990, we launched a survey in six East and Central European countries—Russia, Poland, the Czech Republic, Slovakia, Hungary and Bulgaria—to assess the empirical support for [these] forecasts. In 1993, in each country, we interviewed 1,000 people who were members of the 1988 *nomenklatura*,

Gil Eyal, Iván Szelényi, and Eleanor Townsley. “Post-Communist Managerialism,” in *Making Capitalism Without Capitalists: The New Ruling Elites in Eastern Europe*, pp. 14–15, 113–115, 117, 119–121, 151–152, 230, 243. Copyright © 1998 by Verso Press. Used by permission of Verso Press.

TABLE 1
Occupational Destinations in 1993 of People Who Were in *Nomenklatura* Positions in 1988 by Country

| Occupation in 1993 | Czech Republic | Hungary | Poland |
|--|----------------|---------|--------|
| <i>All in position of authority</i> | 51.7 | 43.1 | 51.2 |
| High political office | 3.0 | 6.4 | 9.0 |
| High manager—public | 16.2 | 11.2 | 13.4 |
| High manager—private | 12.8 | 2.4 | 9.1 |
| High cultural office | 1.1 | 4.4 | 7.1 |
| Low-level managers | 12.6 | 13.0 | 8.6 |
| Entrepreneurs | 6.0 | 5.7 | 4.0 |
| <i>Professionals</i> | 12.2 | 19.9 | 13.9 |
| <i>Workers</i> | 12.6 | 5.5 | 9.5 |
| <i>Retired early (younger than 65)</i> | 15.4 | 19.1 | 17.2 |
| <i>Other retired and unemployed</i> | 8.1 | 12.6 | 8.2 |
| <i>All respondents</i> | 100% | 100% | 100% |
| <i>(n)</i> | (468) | (803) | (849) |

Note: Numbers may not add up to 100 percent due to rounding.

1,000 people who belonged to the new economic, political, and cultural elites at the time; we also conducted personal life history interviews of 5,000 adults randomly selected from the population. In this chapter we present data from three Central European countries: the Czech Republic, Hungary, and Poland. We ask: what happened to the old *nomenklatura*? What are the social origins of members of the new elite? Is there any evidence for the existence of a propertied bourgeoisie by 1993? If so, how is this new class of domestic proprietors constituted? How much, and what, do they own?

Whatever Happened to the *Nomenklatura*?

Political capitalism theory expects to find that the old communist elite has turned itself into the new propertied bourgeoisie of post-communist society. Its most general proposition is that people who were in *nomenklatura* positions prior to 1989 were able to retain their power and privilege through the post-communist transition by converting their political capital into private economic wealth. Our data cast doubt on these predictions.

Table 1 describes the 1993 post-communist occupational destinations of those individuals who occupied *nomenklatura* positions in

1988 in Hungary, the Czech Republic, and Poland.³ While there is some variation across the countries, the main finding in the table is one of massive downward mobility among *nomenklatura* members during the first five years of post-communism. Only half of those who occupied *nomenklatura* positions in 1988 were still in positions of authority in 1993, and this includes rather minor positions in low-level management. Indeed, the proportion of former *nomenklatura* members who occupied *any* authority position in 1993 was rather low. And—rather surprisingly—in Hungary, which boasted the most advanced market reform policies of the late communist period, the loss of authority positions among former *nomenklatura* members has been even more marked than it has been in the other countries: only 43.1 percent of the Hungarian *nomenklatura* retained jobs in which they have subordinates. . . .

If we dig deeper into the data, the problems with the political capitalism thesis become even more serious. In its original formulation, this theory stated that political office had been used for the accumulation of private wealth in Central Europe. In order to test the accuracy of this statement, we need to disaggregate the *nomenklatura* category further. There were very different kinds of *nomenklatura* positions in the communist system, and good reason to

TABLE 2
Occupational Destinations in 1993 of People Who Were in Economic, Political, and Cultural *Nomenklatura* Positions in 1988 in All Three Countries (Czech Republic, Hungary, and Poland)

| <i>Occupation in 1993</i> | <i>Economic Elite, 1988</i> | <i>Political Elite, 1988</i> | <i>Cultural Elite, 1988</i> |
|--|-----------------------------|------------------------------|-----------------------------|
| <i>All in position of authority</i> | 70.7 | 39.3 | 44.2 |
| High political office | 1.5 | 10.8 | 1.3 |
| High manager—public | 33.4 | 5.9 | 7.8 |
| High manager—private | 19.2 | 4.0 | 1.4 |
| High cultural office | 0.4 | 1.8 | 19.3 |
| Low-level managers | 10.8 | 11.0 | 11.8 |
| Entrepreneurs | 5.4 | 5.8 | 2.5 |
| <i>Professionals</i> | 4.7 | 16.9 | 27.4 |
| <i>Workers</i> | 2.2 | 13.4 | 3.3 |
| <i>Retired early (younger than 65)</i> | 14.7 | 20.9 | 11.1 |
| <i>Other retired and unemployed</i> | 7.6 | 9.4 | 14.1 |
| <i>All respondents (n)</i> | 100% (536) | 100% (1,186) | 100% (398) |

Note: Numbers may not add up to 100 percent due to rounding.

think that they were divided among themselves. For this reason, a fair test of political capitalism theory should investigate whether or not any particular component of the *nomenklatura* has successfully negotiated the post-communist transition and become private proprietors. After all, some of our *nomenklatura* members in 1988 were top managers of large firms—thus they were ‘technocrats,’ or what we term the ‘economic elite’ of the late communist period. In their case, it is not obvious that becoming a manager of a privately owned firm is a conversion of political capital into private wealth—managers and technocrats may simply have used their human capital and their managerial experience to maintain senior economic positions in the post-communist transition. Other former *nomenklatura* members belonged to the ‘cultural elite’—they were rectors of universities, Members of the Academies of Sciences, or editors of daily newspapers—and these jobs in the cultural sector were not positions from which one could easily generate vast amounts of private wealth. Thus, political capitalism theory would not necessarily expect former members of the communist cultural elite to be those most able to convert political office into private wealth. Arguably, however, the politi-

cal capitalism thesis should hold for members of the ‘political elite’. These were people who held positions in the Communist Party apparatus or in the civil service, and of all the members of the *nomenklatura* they were the best placed to use their ‘office’ to enrich themselves through the mechanism of ‘spontaneous privatization’.⁴ With this disaggregation of the *nomenklatura* into its economic, cultural, and political components, we are now in a position to offer a crucial test of Staniszkis’s and Hankiss’s version of the political capitalism thesis by asking: to what extent has the ‘bureaucratic’ fraction of the ruling estate—the ‘political elite’—benefited economically from the post-communist transition?

Table 2 documents the 1993 occupational destinations of former *nomenklatura* members for each component of the *nomenklatura*: the economic fraction, the cultural fraction, and the political fraction. It shows, first, that the political fraction of the *nomenklatura* was the least successful in weathering the post-communist transition. Only 39.3 percent of the political *nomenklatura* retained positions of authority between 1988 and 1993, compared to 44.2 percent of the cultural elite and 70.7 percent of the economic elite. Second, early retirement was also much more common among the

political fraction (20.9 percent) of the *nomenklatura* than among the economic and cultural fractions (14.7 and 11.1 percent respectively). This suggests that political capital was much less useful than cultural capital, and particularly cultural capital in the form of managerial expertise, for successfully navigating the pitfalls of post-communism. Third, Table 2 suggests that former Communist Party and state officials who comprised the political fraction of the *nomenklatura* in 1988 had less success in entering new private-sector positions than former communist managers. Among the former members of the economic *nomenklatura*, 24.6 percent (5.4 plus 19.2 percent) either owned or managed a private business in 1993, compared to only 9.8 percent (5.8 plus 4.0 percent) of members of the political fraction. These findings indicate that it was more advantageous to be a manager than a party or state official if one wanted to enter the new economic elite of post-communism. Moreover, these findings directly refute arguments by Staniszki and Hankiss that it was the political fraction of the former communist elite who were best placed to take advantage of post-communist market reforms; rather, we find that members of the economic *nomenklatura* were much bigger beneficiaries of the post-communist transition than members of the political *nomenklatura*. . . .

Diffuse Property Relations as the Context for Managerial Control

In one respect, however, our findings diverge from Szalai's predictions. Staniszki and Hankiss anticipated that former communists would use their power to become the new corporate owners of post-communism, while Szalai argued more specifically that the technocratic-managerial elite was the most likely candidate to achieve this aim. Our data suggest that both these predictions miss the mark. Managerial ownership, or the management buy-out of state-owned firms, is not the major story in post-communism. Indeed, the majority of corporate and industrial man-

agers have acquired no business property at all [for details, see Eyal, Szelényi and Townsley⁵]. Furthermore, fully half of those who own businesses possess stakes not in the firms they manage, but in small subcontracting firms. Finally, we find that those who own shares in the businesses they manage are likely to be managers of smaller firms, and typically own only a small fraction of the assets of these firms. In other words, the former communist technocracy do not hold ultimate economic decision-making power as owners, as Szalai predicted; rather, they exercise power as experts and managers.

While data available to us on ownership relations in large firms are sketchy and may not be sufficiently representative, the evidence at our disposal supports hypotheses put forward on the basis of ethnographic observations by David Stark and Larry King.⁶ Stark found that ownership in Central European corporations is 'recombinant', that is, it is neither private nor public. King found firms with 'recombinant' property, too, but he also identified a number of alternative strategies of privatization, most of which have not led to ownership by identifiable individuals. Our data on property also document diffuse patterns of ownership in post-communist Central Europe, and with the exception of foreign-owned firms (which are really significant only in Hungary), it is not easy to tell who the real owners are. Direct or indirect public ownership, institutional cross-ownership, ownership by banks that are owned by the government or state privatization agencies, and self-ownership (firms owning firms, which own them) are all typical. Together, this creates the material base for the substantial autonomy and power exercised by non-propertied technocrats and managers.

Finally, while the big winners of the post-communist transition are former communist technocrats, we find that they cannot rule by themselves. They have been forced to create a hegemonic power bloc together with the new politocracy and the opinion-making intellec-

tual elite, and these two groups are composed largely of former dissident intellectuals.⁷ Immediately following the fall of communism, the new politocracy and opinion-making intellectual elite made an attempt to squeeze the former communist technocracy out of power. They soon learned, however, that neither fraction of the intellectual elite could rule alone. During the second post-communist elections, many former dissidents were dropped from the politocracy, and the late communist pragmatists joined the new political elite. These are strange bedfellows indeed, who form the ‘unholy alliance’ of post-communism. . . .

An Outline of a Theory of Managerial Capitalism

On the basis of the analyses presented above [and elsewhere],⁸ we summarize our theory of managerial capitalism in the following six theses.

Thesis 1. Post-communist economies are characterized by *diffuse property relations*. At the present time it is impossible to identify individuals or groups of individuals with sufficient amounts of property who are able to exercise anything even remotely similar to owners’ control of economic decision-making.

Thesis 2. Ironically, it was precisely so-called ‘privatization’ which created these diffuse property relations. Privatization destroyed redistributive control over state firms, but it has not produced identifiable owners (yet).

Thesis 3. The dispersion of property rights is a *universal* phenomenon, but in market capitalist economies with an established propertied bourgeoisie it faces strict limits which do not exist in post-communism. Post-communist managers do not have to contend with a class of powerful capitalist proprietors; consequently, managerial power and decision-making are visible contribu-

tions to the prestige and ‘distinction’ of the new power bloc.

Thesis 4. Given the dispersion of property rights, the central representative of managerial power in Central Europe is not the manager of the industrial firm, but the finance manager. The most powerful people of the post-communist era are bank managers, managers of investment funds, experts at the Ministry of Finance, advisors at the IMF and the World Bank, and experts working for foreign and international financial agencies. In the absence of a class of big private proprietors, the power of finance managers is not a function of how many shares they own in the banks they manage, or in the firms their banks manage. Rather, their power is a form of ‘cultural capital’; it is a function of their capacity to appropriate the sacred knowledge of the workings of the world capitalist system.

Thesis 5. Even though Central European managers are not limited by the power of a propertied bourgeoisie, we would emphasize that they do not exercise power in a vacuum. Rather, they occupy a historically distinctive post-communist class context, formed by ongoing struggles over privatization and the formation of new class actors. These struggles take place among members of the power bloc—managers, technocrats, and intellectuals, in the first instance—and between this power bloc and former bureaucrats, in the second. As we have argued [elsewhere],⁹ no single class fraction has emerged as the decisive victor in these struggles, and as a consequence the power of managers and their capacity to control semi-public property comes to them by default.

Thesis 6. Managerial strategies reflect their knowledge that the current balance of class forces is precarious. They understand that they exercise power by default. In order to survive, therefore,

managers have developed a diverse range of strategies to navigate the political and economic uncertainties of post-communism. Probably the most prevalent managerial strategy was not managerial 'buy-out' but, rather, an attempt to stand 'on as many legs as possible'. During the process of privatization most managers acquired *some* property, but this was typically a relatively small stake, and not even necessarily in the firms they managed. Indeed, as early as the late 1980s some members of management teams were busy setting up small subcontracting firms, owned by themselves or by members of their families. They subcontracted the most lucrative activities of the state firms they managed to these companies, they even sold some of the more valuable assets of the parent firms to these subcontracting units at undervalued prices. Still, it is probably the exception to the rule that these managers retired from their main firm altogether, that they 'jumped the boat' to run the subcontracting firms they own. The reason for their reluctance to do so is clear: why should they swap a *major* managerial job for the position of owner-manager in a *minor* operation which employs only a handful of people? On the other hand, managers also have an interest in being more than *only* managers. In post-communist society, the managerial elite is closely intertwined with the polittocracy; hence post-communist managers are even more dependent upon politicians than are capitalist managers in the West. State bureaucracies in East Central Europe often have the power, through direct or indirect state ownership of firms, to appoint and dismiss managers. As long as their position can be threatened by the political elite, it seems to be wise for

managers to have their own small private firm in the background.

NOTES

1. Jadwiga Staniszkis (1991a) *The Dynamics of Breakthrough*. Berkeley: University of California Press; Elemér Hankiss (1990) *East European Alternatives*. Oxford: Clarendon Press.

2. Erzsébet Szalai (1989) 'The New Elite'. *Across Frontiers* 5(Fall-Winter): 25–31. [In Hungarian: *Beszélő* 26, 1989].

3. The population of *nomenklatura* members was defined as those individuals occupying the top 3,000–5,000 positions in these countries in 1988, positions for which appointment usually required the approval of some organ or official of the Central Committee of the Communist Party.

4. In both Hungary and Poland during 1988, the regimes launched programs which were referred to as 'spontaneous privatization'. Communist elites acknowledged the necessity of changing property rights, and privatizing publicly held assets, but they initiated this process in a rather unregulated way. Under spontaneous privatization, firms could initiate their own privatization and negotiate their own terms with state organizations (in 1988, this still meant the Communist Party). Indeed, the initial hypotheses about political capitalism put forward by Staniszkis and Hankiss were formulated in reaction to these spontaneous privatization plans.

5. Gil Eyal, Iván Szelényi and Eleanor Townsley (1998) *Making Capitalism without Capitalists: The New Ruling Elites in Eastern Europe*. London: Verso.

6. David Stark (1996) 'Recombinant Property in East European Capitalism'. *American Journal of Sociology* 101(4): 993–1027; Larry King (1997) *Pathways from Socialism: The Transformation of Firms in Hungary, the Czech Republic and Slovakia*. Ph.D dissertation, Department of Sociology, UCLA.

7. Erzsébet Szalai (1994) 'The Power Structure in Hungary after the Political Transition'. Pp. 120–43 in *The New Great Transformation*, ed. Christopher G.A. Bryant and Edmund Mokrzycki. London and New York: Routledge; Erzsébet Szalai (1994) *Utelagazás: Hatalom és értelmiség az államszocializmus után* (At the crossroads: power and intellectuals after state socialism). Budapest: Pesti Szalon Kiadó.

8. Eyal, Szelényi and Townsley, *Making Capitalism Without Capitalists*.

9. *Ibid.*

Poverty and the Underclass

■ THE EXPERIENCE OF POVERTY ■

35 ■ Barbara Ehrenreich

Nickel-and-Dimed

On (not) Getting by in America

At the beginning of June 1998 I leave behind everything that normally soothes the ego and sustains the body—home, career, companion, reputation, ATM card—for a plunge into the low-wage workforce. There, I become another, occupationally much diminished “Barbara Ehrenreich”—depicted on job-application forms as a divorced homemaker whose sole work experience consists of housekeeping in a few private homes. I am terrified, at the beginning, of being unmasked for what I am: a middle-class journalist setting out to explore the world that welfare mothers are entering, at the rate of approximately 50,000 a month, as welfare reform kicks in. Happily, though, my fears turn out to be entirely unwarranted: during a month of poverty and toil, my name goes unnoticed and for the most part unuttered. In this parallel universe where my father never got out of the mines and I never got through college, I am “baby,” “honey,” “blondie,” and, most commonly, “girl.”

My first task is to find a place to live. I figure that if I can earn \$7 an hour—which, from the want ads, seems doable—I can afford to spend \$500 on rent, or maybe, with severe economies, \$600. In the Key West area,

where I live, this pretty much confines me to flophouses and trailer homes—like the one, a pleasing fifteen-minute drive from town, that has no air-conditioning, no screens, no fans, no television, and, by way of diversion, only the challenge of evading the landlord’s Doberman pinscher. The big problem with this place, though, is the rent, which at \$675 a month is well beyond my reach. . . .

So I decide to make the common trade-off between affordability and convenience, and go for a \$500-a-month efficiency thirty miles up a two-lane highway from the employment opportunities of Key West, meaning forty-five minutes if there’s no road construction and I don’t get caught behind some sun-dazed Canadian tourists. . . .

I am not doing this for the anthropology. My aim is nothing so mistily subjective as to “experience poverty” or find out how it “really feels” to be a long-term low-wage worker. I’ve had enough unchosen encounters with poverty and the world of low-wage work to know it’s not a place you want to visit for touristic purposes; it just smells too much like fear. And with all my real-life assets—bank account, IRA, health insurance,

multiroom home—waiting indulgently in the background, I am, of course, thoroughly insulated from the terrors that afflict the genuinely poor.

No, this is a purely objective, scientific sort of mission. The humanitarian rationale for welfare reform—as opposed to the more punitive and stingy impulses that may actually have motivated it—is that work will lift poor women out of poverty while simultaneously inflating their self-esteem and hence their future value in the labor market. Thus, whatever the hassles involved in finding child care, transportation, etc., the transition from welfare to work will end happily, in greater prosperity for all. Now there are many problems with this comforting prediction, such as the fact that the economy will inevitably undergo a downturn, eliminating many jobs. Even without a downturn, the influx of a million former welfare recipients into the low-wage labor market could depress wages by as much as 11.9 percent, according to the Economic Policy Institute (EPI) in Washington, D.C.

But is it really possible to make a living on the kinds of jobs currently available to unskilled people? Mathematically, the answer is no, as can be shown by taking \$6 to \$7 an hour, perhaps subtracting a dollar or two an hour for child care, multiplying by 160 hours a month, and comparing the result to the prevailing rents. According to the National Coalition for the Homeless, for example, in 1998 it took, on average nationwide, an hourly wage of \$8.89 to afford a one-bedroom apartment, and the Preamble Center for Public Policy estimates that the odds against a typical welfare recipient's landing a job at such a "living wage" are about 97 to 1. If these numbers are right, low-wage work is not a solution to poverty and possibly not even to homelessness. . . .

On the morning of my first full day of job searching, I take a red pen to the want ads, which are auspiciously numerous. Everyone

in Key West's booming "hospitality industry" seems to be looking for someone like me—trainable, flexible, and with suitably humble expectations as to pay. I know I possess certain traits that might be advantageous—I'm white and, I like to think, well-spoken and poised—but I decide on two rules: One, I cannot use any skills derived from my education or usual work—not that there are a lot of want ads for satirical essayists anyway. Two, I have to take the best-paid job that is offered me and of course do my best to hold it; no Marxist rants or sneaking off to read novels in the ladies' room. . . .

Most of the big hotels run ads almost continually, just to build a supply of applicants to replace the current workers as they drift away or are fired, so finding a job is just a matter of being at the right place at the right time and flexible enough to take whatever is being offered that day. This finally happens to me at one of the big discount hotel chains, where I go for housekeeping and am sent, instead, to try out as a waitress at the attached "family restaurant," a dismal spot with a counter and about thirty tables that looks out on a parking garage and features such tempting fare as "Pollish [sic] sausage and BBQ sauce" on 95-degree days. Phillip, the dapper young West Indian who introduces himself as the manager, interviews me with about as much enthusiasm as if he were a clerk processing me for Medicare, the principal questions being what shifts can I work and when can I start. I mutter something about being woefully out of practice as a waitress, but he's already on to the uniform: I'm to show up tomorrow wearing black slacks and black shoes; he'll provide the rust-colored polo shirt with Hearthsides embroidered on it, though I might want to wear my own shirt to get to work, ha ha. At the word "tomorrow," something between fear and indignation rises in my chest. I want to say, "Thank you for your time, sir, but this is just an experiment, you know, not my actual life."

So begins my career at the Hearthside, I shall call it, one small profit center within a global discount hotel chain, where for two weeks I work from 2:00 till 10:00 P.M. for \$2.43 an hour plus tips. In some futile bid for gentility, the management has barred employees from using the front door, so my first day I enter through the kitchen, where a red-faced man with shoulder-length blond hair is throwing frozen steaks against the wall and yelling, “Fuck this shit!” “That’s just Jack,” explains Gail, the wiry middle-aged waitress who is assigned to train me. “He’s on the rag again”—a condition occasioned, in this instance, by the fact that the cook on the morning shift had forgotten to thaw out the steaks. For the next eight hours, I run after the agile Gail, absorbing bits of instruction along with fragments of personal tragedy. All food must be trayed, and the reason she’s so tired today is that she woke up in a cold sweat thinking of her boyfriend, who killed himself recently in an upstate prison. No refills on lemonade. And the reason he was in prison is that a few DUIs caught up with him, that’s all, could have happened to anyone. Carry the creamers to the table in a monkey bowl, never in your hand. And after he was gone she spent several months living in her truck, peeing in a plastic pee bottle and reading by candlelight at night, but you can’t live in a truck in the summer, since you need to have the windows down, which means anything can get in, from mosquitoes on up.

At least Gail puts to rest any fears I had of appearing overqualified. From the first day on, I find that of all the things I have left behind, such as home and identity, what I miss the most is competence. Not that I have ever felt utterly competent in the writing business, in which one day’s success augurs nothing at all for the next. But in my writing life, I at least have some notion of procedure: do the research, make the outline, rough out a draft, etc. As a server, though, I am beset by

requests like bees: more iced tea here, ketchup over there, a to-go box for table fourteen, and where are the high chairs, anyway? Of the twenty-seven tables, up to six are usually mine at any time, though on slow afternoons or if Gail is off, I sometimes have the whole place to myself. There is the touch-screen computer-ordering system to master, which is, I suppose, meant to minimize server-cook contact, but in practice requires constant verbal fine-tuning: “That’s gravy on the mashed, okay? None on the meatloaf,” and so forth—while the cook scowls as if I were inventing these refinements just to torment him. Plus, something I had forgotten in the years since I was eighteen: about a third of a server’s job is “side work” that’s invisible to customers—sweeping, scrubbing, slicing, refilling, and restocking. If it isn’t all done, every little bit of it, you’re going to face the 6:00 P.M. dinner rush defenseless and probably go down in flames. I screw up dozens of times at the beginning, sustained in my shame entirely by Gail’s support—“It’s okay, baby, everyone does that sometime”—because, to my total surprise and despite the scientific detachment I am doing my best to maintain, I care. . . .

Sometimes I play with the fantasy that I am a princess who, in penance for some tiny transgression, has undertaken to feed each of her subjects by hand. But the non-princesses working with me are just as indulgent, even when this means flouting management rules—concerning, for example, the number of croutons that can go on a salad (six). “Put on all you want,” Gail whispers, “as long as Stu isn’t looking.” She dips into her own tip money to buy biscuits and gravy for an out-of-work mechanic who’s used up all his money on dental surgery, inspiring me to pick up the tab for his milk and pie. . . .

Ten days into it, this is beginning to look like a livable lifestyle. I like Gail, who is “looking at fifty” but moves so fast she can alight in one place and then another without

apparently being anywhere between them. I clown around with Lionel, the teenage Haitian busboy, and catch a few fragments of conversation with Joan, the svelte fortyish hostess and militant feminist who is the only one of us who dares to tell Jack to shut the fuck up. I even warm up to Jack when, on a slow night and to make up for a particularly unwarranted attack on my abilities, or so I imagine, he tells me about his glory days as a young man at “coronary school”—or do you say “culinary”?—in Brooklyn, where he dated a knock-out Puerto Rican chick and learned everything there is to know about food. I finish up at 10:00 or 10:30, depending on how much side work I’ve been able to get done during the shift, and cruise home to the tapes I snatched up at random when I left my real home—Marianne Faithfull, Tracy Chapman, Enigma, King Sunny Ade, the Violent Femmes—just drained enough for the music to set my cranium resonating but hardly dead. Midnight snack is Wheat Thins and Monterey Jack, accompanied by cheap white wine on ice and whatever AMC has to offer. To bed by 1:30 or 2:00, up at 9:00 or 10:00, read for an hour while my uniform whirls around in the landlord’s washing machine, and then it’s another eight hours spent following Mao’s central instruction, as laid out in the Little Red Book, which was: Serve the people.

I could drift along like this, in some dreamy proletarian idyll, except for two things. One is management. If I have kept this subject on the margins thus far it is because I still flinch to think that I spent all those weeks under the surveillance of men (and later women) whose job it was to monitor my behavior for signs of sloth, theft, drug abuse, or worse. Not that managers and especially “assistant managers” in low-wage settings like this are exactly the class enemy. In the restaurant business, they are mostly former cooks or servers, still capable of pinch-hitting in the kitchen or on the floor,

just as in hotels they are likely to be former clerks, and paid a salary of only about \$400 a week. But everyone knows they have crossed over to the other side, which is, crudely put, corporate as opposed to human. Cooks want to prepare tasty meals; servers want to serve them graciously; but managers are there for only one reason—to make sure that money is made for some theoretical entity that exists far away in Chicago or New York, if a corporation can be said to have a physical existence at all. . . .

Managers can sit—for hours at a time if they want—but it’s their job to see that no one else ever does, even when there’s nothing to do, and this is why, for servers, slow times can be as exhausting as rushes. You start dragging out each little chore, because if the manager on duty catches you in an idle moment, he will give you something far nastier to do. So I wipe, I clean, I consolidate ketchup bottles and recheck the cheesecake supply, even tour the tables to make sure the customer evaluation forms are all standing perkily in their places—wondering all the time how many calories I burn in these strictly theatrical exercises. When, on a particularly dead afternoon, Stu finds me glancing at a *USA Today* a customer has left behind, he assigns me to vacuum the entire floor with the broken vacuum cleaner that has a handle only two feet long, and the only way to do that without incurring orthopedic damage is to proceed from spot to spot on your knees. . . .

The other problem, in addition to the less-than-nurturing management style, is that this job shows no sign of being financially viable. You might imagine, from a comfortable distance, that people who live, year in and year out, on \$6 to \$10 an hour have discovered some survival stratagems unknown to the middle class. But no. It’s not hard to get my co-workers to talk about their living situations, because housing, in almost every case, is the principal source of disruption in their

lives, the first thing they fill you in on when they arrive for their shifts. After a week, I have compiled the following survey:

- Gail is sharing a room in a well-known downtown flophouse for which she and a roommate pay about \$250 a week. Her roommate, a male friend, has begun hitting on her, driving her nuts, but the rent would be impossible alone.
- Claude, the Haitian cook, is desperate to get out of the two-room apartment he shares with his girlfriend and two other, unrelated, people. As far as I can determine, the other Haitian men (most of whom only speak Creole) live in similarly crowded situations.
- Annette, a twenty-year-old server who is six months pregnant and has been abandoned by her boyfriend, lives with her mother, a postal clerk.
- Marianne and her boyfriend are paying \$170 a week for a one-person trailer.
- Jack, who is, at \$10 an hour, the wealthiest of us, lives in the trailer he owns, paying only the \$400-a-month lot fee.
- The other white cook, Andy, lives on his dry-docked boat, which, as far as I can tell from his loving descriptions, can't be more than twenty feet long. He offers to take me out on it, once it's repaired, but the offer comes with inquiries as to my marital status, so I do not follow up on it.
- Tina and her husband are paying \$60 a night for a double room in a Days Inn. This is because they have no car and the Days Inn is within walking distance of the Hearthside. When Marianne, one of the breakfast servers, is tossed out of her trailer for subletting (which is against the trailer-park rules), she leaves her boyfriend and moves in with Tina and her husband.
- Joan, who had fooled me with her numerous and tasteful outfits (hostesses

wear their own clothes), lives in a van she parks behind a shopping center at night and showers in Tina's motel room. The clothes are from thrift shops.

It strikes me, in my middle-class solipsism, that there is gross improvidence in some of these arrangements. When Gail and I are wrapping silverware in napkins—the only task for which we are permitted to sit—she tells me she is thinking of escaping from her roommate by moving into the Days Inn herself. I am astounded: How can she even think of paying between \$40 and \$60 a day? But if I was afraid of sounding like a social worker, I come out just sounding like a fool. She squints at me in disbelief, “And where am I supposed to get a month's rent and a month's deposit for an apartment?” I'd been feeling pretty smug about my \$500 efficiency, but of course it was made possible only by the \$1,300 I had allotted myself for start-up costs when I began my low-wage life: \$1,000 for the first month's rent and deposit, \$100 for initial groceries and cash in my pocket, \$200 stuffed away for emergencies. In poverty, as in certain propositions in physics, starting conditions are everything.

There are no secret economies that nourish the poor; on the contrary, there are a host of special costs. If you can't put up the two months' rent you need to secure an apartment, you end up paying through the nose for a room by the week. If you have only a room, with a hot plate at best, you can't save by cooking up huge lentil stews that can be frozen for the week ahead. You eat fast food, or the hot dogs and Styrofoam cups of soup that can be microwaved in a convenience store. If you have no money for health insurance—and the Hearthside's niggardly plan kicks in only after three months—you go without routine care or prescription drugs and end up paying the price. Gail, for example, was fine until she ran out of money for estrogen pills. She is supposed to be on the

company plan by now, but they claim to have lost her application form and need to begin the paperwork all over again. So she spends \$9 per migraine pill to control the headaches she wouldn't have, she insists, if her estrogen supplements were covered. Similarly, Marianne's boyfriend lost his job as a roofer because he missed so much time after getting a cut on his foot for which he couldn't afford the prescribed antibiotic.

My own situation, when I sit down to assess it after two weeks of work, would not be much better if this were my actual life. The seductive thing about waitressing is that you don't have to wait for payday to feel a few bills in your pocket, and my tips usually cover meals and gas, plus something left over to stuff into the kitchen drawer I use as a bank. But as the tourist business slows in the summer heat, I sometimes leave work with only \$20 in tips (the gross is higher, but servers share about 15 percent of their tips with the busboys and bartenders). With wages included, this amounts to about the minimum wage of \$5.15 an hour. Although the sum in the drawer is piling up, at the present rate of accumulation it will be more than a hundred dollars short of my rent when the end of the month comes around. Nor can I see any expenses to cut. True, I haven't gone the lentil-stew route yet, but that's because I don't have a large cooking pot, pot holders, or a ladle to stir with (which cost about \$30 at Kmart, less at thrift stores), not to mention onions, carrots, and the indispensable bay leaf. I do make my lunch almost every day—usually some slow-burning, high-protein combo like frozen chicken patties with melted cheese on top and canned pinto beans on the side. Dinner is at the Hearthside, which offers its employees a choice of BLT, fish sandwich, or hamburger for only \$2. The burger lasts longest, especially if it's heaped with gut-puckering jalapeños, but by midnight my stomach is growling again.

So unless I want to start using my car as a residence, I have to find a second, or alternative, job. I call all the hotels where I filled out housekeeping applications weeks ago—the Hyatt, Holiday Inn, Econo Lodge, HoJo's, Best Western, plus a half dozen or so locally run guesthouses. Nothing. Then I start making the rounds again, wasting whole mornings waiting for some assistant manager to show up, even dipping into places so creepy that the front-desk clerk greets you from behind bulletproof glass and sells pints of liquor over the counter. But either someone has exposed my real-life housekeeping habits—which are, shall we say, mellow—or I am at the wrong end of some infallible ethnic equation: most, but by no means all, of the working housekeepers I see on my job searches are African Americans, Spanish-speaking, or immigrants from the Central European post-Communist world, whereas servers are almost invariably white and monolingually English-speaking. When I finally get a positive response, I have been identified once again as server material. Jerry's, which is part of a well-known national family restaurant chain and physically attached here to another budget hotel chain, is ready to use me at once. The prospect is both exciting and terrifying, because, with about the same number of tables and counter seats, Jerry's attracts three or four times the volume of customers as the gloomy old Hearthside. . . .

I start out with the beautiful, heroic idea of handling the two jobs at once, and for two days I almost do it: the breakfast/lunch shift at Jerry's, which goes till 2:00, arriving at the Hearthside at 2:10, and attempting to hold out until 10:00. In the ten minutes between jobs, I pick up a spicy chicken sandwich at the Wendy's drive-through window, gobble it down in the car, and change from khaki slacks to black, from Hawaiian to rust polo. There is a problem, though. When during the 3:00 to 4:00 P.M. dead time I finally sit

down to wrap silver, my flesh seems to bond to the seat. I try to refuel with a purloined cup of soup, as I've seen Gail and Joan do dozens of times, but a manager catches me and hisses "No eating!" though there's not a customer around to be offended by the sight of food making contact with a server's lips. So I tell Gail I'm going to quit, and she hugs me and says she might just follow me to Jerry's herself.

But the chances of this are minuscule. She has left the flophouse and her annoying roommate and is back to living in her beat-up old truck. But guess what? she reports to me excitedly later that evening: Phillip has given her permission to park overnight in the hotel parking lot, as long as she keeps out of sight, and the parking lot should be totally safe, since it's patrolled by a hotel security guard! With the Hearthiside offering benefits like that, how could anyone think of leaving? . . .

Management at Jerry's is generally calmer and more "professional" than at the Hearthiside, with two exceptions. One is Joy, a plump, blowsy woman in her early thirties, who once kindly devoted several minutes to instructing me in the correct one-handed method of carrying trays but whose moods change disconcertingly from shift to shift and even within one. Then there's B. J., a.k.a. B.J.-the-bitch, whose contribution is to stand by the kitchen counter and yell, "Nita, your order's up, move it!" or, "Barbara, didn't you see you've got another table out there? Come on, girl!" Among other things, she is hated for having replaced the whipped-cream squirt cans with big plastic whipped-cream-filled baggies that have to be squeezed with both hands—because, reportedly, she saw or thought she saw employees trying to inhale the propellant gas from the squirt cans, in the hope that it might be nitrous oxide. On my third night, she pulls me aside abruptly and brings her face so close that it looks as if she's planning to butt me with her forehead. But instead of saying, "You're fired," she says,

"You're doing fine." The only trouble is I'm spending time chatting with customers: "That's how they're getting you." Furthermore I am letting them "run me," which means harassment by sequential demands: you bring the ketchup and they decide they want extra Thousand Island; you bring that and they announce they now need a side of fries; and so on into distraction. Finally she tells me not to take her wrong. She tries to say things in a nice way, but you get into a mode, you know, because everything has to move so fast. . . .

I make friends, over time, with the other "girls" who work my shift: Nita, the tattooed twenty-something who taunts us by going around saying brightly, "Have we started making money yet?" Ellen, whose teenage son cooks on the graveyard shift and who once managed a restaurant in Massachusetts but won't try out for management here because she prefers being a "common worker" and not "ordering people around." Easy-going fiftyish Lucy, with the raucous laugh, who limps toward the end of the shift because of something that has gone wrong with her leg, the exact nature of which cannot be determined without health insurance. We talk about the usual girl things—men, children, and the sinister allure of Jerry's chocolate peanut-butter cream pie—though no one, I notice, ever brings up anything potentially expensive, like shopping or movies. As at the Hearthiside, the only recreation ever referred to is partying, which requires little more than some beer, a joint, and a few close friends. Still, no one here is homeless, or cops to it anyway, thanks usually to a working husband or boyfriend. All in all, we form a reliable mutual-support group: If one of us is feeling sick or overwhelmed, another one will "bev" a table or even carry trays for her. If one of us is off sneaking a cigarette or a pee, the others will do their best to conceal her absence from the enforcers of corporate rationality. . . .

I make the decision to move closer to Key West. First, because of the drive. Second and third, also because of the drive: gas is eating up \$4 to \$5 a day, and although Jerry's is as high-volume as you can get, the tips average only 10 percent, and not just for a newbie like me. Between the base pay of \$2.15 an hour and the obligation to share tips with the busboys and dishwashers, we're averaging only about \$7.50 an hour. Then there is the \$30 I had to spend on the regulation tan slacks worn by Jerry's servers—a setback it could take weeks to absorb. (I had combed the town's two downscale department stores hoping for something cheaper but decided in the end that these marked-down Dockers, originally \$49, were more likely to survive a daily washing.) Of my fellow servers, everyone who lacks a working husband or boyfriend seems to have a second job: Nita does something at a computer eight hours a day; another welds. Without the forty-five-minute commute, I can picture myself working two jobs and having the time to shower between them.

So I take the \$500 deposit I have coming from my landlord, the \$400 I have earned toward the next month's rent, plus the \$200 reserved for emergencies, and use the \$1,100 to pay the rent and deposit on trailer number 46 in the Overseas Trailer Park, a mile from the cluster of budget hotels that constitute Key West's version of an industrial park. Number 46 is about eight feet in width and shaped like a barbell inside, with a narrow region—because of the sink and the stove—separating the bedroom from what might optimistically be called the “living” area, with its two-person table and half-sized couch. The bathroom is so small my knees rub against the shower stall when I sit on the toilet, and you can't just leap out of the bed, you have to climb down to the foot of it in order to find a patch of floor space to stand on. Outside, I am within a few yards of a liquor store, a bar that advertises “free beer

tomorrow,” a convenience store, and a Burger King—but no supermarket or, alas, laundromat. By reputation, the Overseas park is a nest of crime and crack, and I am hoping at least for some vibrant, multicultural street life. But desolation rules night and day, except for a thin stream of pedestrian traffic heading for their jobs at the Sheraton or 7-Eleven. There are not exactly people here but what amounts to canned labor, being preserved from the heat between shifts. . . .

When my month-long plunge into poverty is almost over, I finally land my dream job—housekeeping. I do this by walking into the personnel office of the only place I figure I might have some credibility, the hotel attached to Jerry's, and confiding urgently that I have to have a second job if I am to pay my rent and, no, it couldn't be front-desk clerk. “All right,” the personnel lady fairly spits, “So it's housekeeping,” and she marches me back to meet Maria, the housekeeping manager, a tiny, frenetic Hispanic woman who greets me as “babe” and hands me a pamphlet emphasizing the need for a positive attitude. The hours are nine in the morning till whenever, the pay is \$6.10 an hour, and there's one week of vacation a year. I don't have to ask about health insurance once I meet Carlotta, the middle-aged African-American woman who will be training me. Carla, as she tells me to call her, is missing all of her top front teeth.

On that first day of housekeeping and last day of my entire project—although I don't yet know it's the last—Carla is in a foul mood. We have been given nineteen rooms to clean, most of them “checkouts,” as opposed to “stay-overs,” that require the whole enchilada of bed-stripping, vacuuming, and bathroom-scrubbing. When one of the rooms that had been listed as a stay-over turns out to be a checkout, Carla calls Maria to complain, but of course to no avail. “So make up the motherfucker,” Carla orders

me, and I do the beds while she sloshes around the bathroom. For four hours without a break I strip and remake beds, taking about four and a half minutes per queen-sized bed, which I could get down to three if there were any reason to. We try to avoid vacuuming by picking up the larger specks by hand, but often there is nothing to do but drag the monstrous vacuum cleaner—it weighs about thirty pounds—off our cart and try to wrestle it around the floor. Sometimes Carla hands me the squirt bottle of “BAM” (an acronym for something that begins, ominously, with “butyric”; the rest has been worn off the label) and lets me do the bathrooms. No service ethic challenges me here to new heights of performance. I just concentrate on removing the pubic hairs from the bathtubs, or at least the dark ones that I can see. . . .

When I request permission to leave at about 3:30, another housekeeper warns me that no one has so far succeeded in combining housekeeping at the hotel with serving at Jerry’s: “Some kid did it once for five days, and you’re no kid.” With that helpful information in mind, I rush back to number 46, down four Advils, shower, stooping to fit into the stall, and attempt to compose myself for the oncoming shift. So much for what Marx termed the “reproduction of labor power,” meaning the things a worker has to do just so she’ll be ready to work again. . . .

Then it comes, the perfect storm. Four of my tables fill up at once. Four tables is nothing for me now, but only so long as they are obligingly staggered. As I bev table 27, tables 25, 28, and 24 are watching enviously. As I bev 25, 24 glowers because their bevs haven’t even been ordered. Twenty-eight is four yuppie types, meaning everything on the side and agonizing instructions as to the chicken Caesars. Twenty-five is a middle-aged black couple, who complain, with some justice, that the iced tea isn’t fresh and the tabletop is sticky. But table 24 is the meteorological

event of the century: ten British tourists who seem to have made the decision to absorb the American experience entirely by mouth. Here everyone has at least two drinks—iced tea and milk shake, Michelob and water (with lemon slice, please)—and a huge promiscuous orgy of breakfast specials, mozz sticks, chicken strips, quesadillas, burgers with cheese and without, sides of hash browns with cheddar, with onions, with gravy, seasoned fries, plain fries, banana splits. Poor Jesus (the cook)! Poor me! Because when I arrive with their first tray of food—after three prior trips just to refill bevs—Princess Di refuses to eat her chicken strips with her pancake-and-sausage special, since, as she now reveals, the strips were meant to be an appetizer. Maybe the others would have accepted their meals, but Di, who is deep into her third Michelob, insists that everything else go back while they work on their “starters.” Meanwhile, the yuppies are waving me down for more decaf and the black couple looks ready to summon the NAACP.

Much of what happened next is lost in the fog of war. Jesus starts going under. The little printer on the counter in front of him is spewing out orders faster than he can rip them off, much less produce the meals. Even the invincible Ellen is ashen from stress. I bring table 24 their reheated main courses, which they immediately reject as either too cold or fossilized by the microwave. When I return to the kitchen with their trays (three trays in three trips), Joy confronts me with arms akimbo: “What is this?” She means the food—the plates of rejected pancakes, hash browns in assorted flavors, toasts, burgers, sausages, eggs. “Uh, scrambled with cheddar,” I try, “and that’s. . . .” “NO,” she screams in my face. “Is it a traditional, a super-scramble, an eye-opener?” I pretend to study my check for a clue, but entropy has been up to its tricks, not only on the plates but in my head, and I have to admit that the original order is beyond reconstruction. “You don’t know an

eye-opener from a traditional?” she demands in outrage. All I know, in fact, is that my legs have lost interest in the current venture and have announced their intention to fold. I am saved by a yuppie (mercifully not one of mine) who chooses this moment to charge into the kitchen to bellow that his food is twenty-five minutes late. Joy screams at him to get the hell out of her kitchen, please, and then turns on Jesus in a fury, hurling an empty tray across the room for emphasis.

I leave. I don't walk out, I just leave. I don't finish my side work or pick up my credit-card tips, if any, at the cash register or, of course, ask Joy's permission to go. And the surprising thing is that you *can* walk out without permission, that the door opens, that the thick tropical night air parts to let me pass, that my car is still parked where I left it. There is no vindication in this exit, no fuck-you surge of relief, just an overwhelming, dank sense of failure pressing down on me and the entire parking lot. I had gone into this venture in the spirit of science, to test a mathematical proposition, but somewhere along the line, in the tunnel vision imposed by long shifts and relentless concentration, it became a test of myself, and clearly I have failed. . . .

When I moved out of the trailer park, I gave the key to number 46 to Gail and arranged for my deposit to be transferred to her. She told me that Joan is still living in her van and that Stu had been fired from the Hearthside. . . .

In one month, I had earned approximately \$1,040 and spent \$517 on food, gas, toiletries, laundry, phone, and utilities. If I had remained in my \$500 efficiency, I would have been able to pay the rent and

have \$22 left over (which is \$78 less than the cash I had in my pocket at the start of the month). During this time I bought no clothing except for the required slacks and no prescription drugs or medical care (I did finally buy some vitamin B to compensate for the lack of vegetables in my diet). Perhaps I could have saved a little on food if I had gotten to a supermarket more often, instead of convenience stores, but it should be noted that I lost almost four pounds in four weeks, on a diet weighted heavily toward burgers and fries.

How former welfare recipients and single mothers will (and do) survive in the low-wage workforce, I cannot imagine. Maybe they will figure out how to condense their lives—including child-raising, laundry, romance, and meals—into the couple of hours between full-time jobs. Maybe they will take up residence in their vehicles, if they have one. All I know is that I couldn't hold two jobs and I couldn't make enough money to live on with one. And I had advantages unthinkable to many of the long-term poor—health, stamina, a working car, and no children to care for and support. . . .

The thinking behind welfare reform was that even the humblest jobs are morally uplifting and psychologically buoying. In reality they are likely to be fraught with insult and stress. But I did discover one redeeming feature of the most abject low-wage work—the camaraderie of people who are, in almost all cases, far too smart and funny and caring for the work they do and the wages they're paid. The hope, of course, is that someday these people will come to know what they're worth, and take appropriate action.

■ HOW MUCH POVERTY IS THERE? ■

36 ■ Timothy M. Smeeding

Poverty, Work, and Policy

The United States in Comparative Perspective

Introduction

The United States has a long tradition of measuring income poverty and weighing the effectiveness, successes, and failures of government policies aimed at poverty reduction. But for the most part, examinations of U.S. domestic antipoverty policy are inherently parochial, for they are based on the experiences of only our nation in isolation from the others. The estimation of cross-nationally equivalent measures of income and poverty provides a unique opportunity to compare poverty rates and the design and effectiveness of American social policy with the experiences of other nations. The Luxembourg Income Study (LIS) database, which undergirds this paper, contains the information needed to construct comparable poverty measures for more than 30 nations. It allows comparisons of the level and trend of poverty and inequality across several nations, along with considerable detail on the sources of market incomes, disposable incomes, and public policies that in large part shape these outcomes.

In this paper we use analyses made possible by the LIS to examine America's experiences in fighting poverty in the face of substantial

and rising economic inequality in a cross-national context. In so doing, we compare the effectiveness of U.S. antipoverty policies to those of similar nations elsewhere in the industrialized world. We attempt to answer the following questions:

- Do other countries have an “official” poverty line in the sense that the United States does, or do they define poverty in a de facto sense based on eligibility for various government programs and relative economic standing?
- How do poverty rates in other countries compare with U.S. poverty rates?
- What are the big drivers of poverty in the United States compared to other countries, with low wages, low-skill immigrants, and large numbers of single-parent families being the most prominent candidates?

We believe that there are lessons about antipoverty policy that can be learned from cross-national comparisons. While every nation has its own idiosyncratic institutions and policies, reflecting its values, culture, institutions, and

Commissioned for this book. The author would like to thank Gary Burtless for comments on an earlier draft and Karen Cimilluca and Mary Santy for their help with manuscript preparation. The research that is reported here was sponsored in part by the Ford and Russell Sage Foundations. The author assumes responsibility for all errors of commission and omission.

history, wide differences in success and failure are evident from the comparisons that follow. And, there is evidence that such policies are becoming internationalized in their spread and evaluation (Banks et al. 2005; Francesconi and van der Klaauw 2007).

We begin by reviewing international concepts and measures of poverty, as they relate to the main measures used in domestic U.S. academic and policy discourse. We follow with a discussion of the relationship between policy differences and outcome differences among the several countries and consider the implications of our analysis for antipoverty policy in the United States. While all nations value low poverty, high levels of economic self-reliance, and equality of opportunity for younger persons, they seem to differ dramatically in the extent to which they reach these goals. Most nations have remarkable similarities in the sources of national social concern: births outside of wedlock and lone parent families; older women living alone; high unemployment; immigration pressures; low wages; and the sustainability of social expenditures in the face of rapid population aging and rising medical care costs. But they also exhibit differences in the extent to which working-age adults mix economic self-reliance (earned incomes), family support, and government support to avoid poverty. Moreover, in such comparisons the United States does not always look very supportive of low-income families.

Cross-National Comparisons of Poverty and Inequality: Methodology and Measurement

Who Measures Poverty and How?

Differing national experiences in social transfer and antipoverty programs provide a rich source of information for evaluating the effectiveness of alternative social policies in fighting poverty. While most rich nations share a concern over low incomes, poverty measurement began as an Anglo-American

social indicator. In fact, “official” measures of poverty (or measures of “low-income” status) exist in very few nations. Only the United States and the United Kingdom have regular “official” poverty series.

In Northern Europe and Scandinavia the debate centers instead on the level of income at which minimum benefits for social programs should be set and on the issue of “social exclusion” (Atkinson, Cantillon, Marlier, and Nolan 2005). Northern European and Scandinavian nations do not calculate low income or poverty rates. Most recognize that their social programs already ensure a low poverty rate under any reasonable set of measurement standards (Björklund and Freeman 1997).

While there is no international consensus on guidelines for measuring poverty, international bodies such as the United Nations Children’s Fund (UNICEF), the United Nations Human Development Report (UNHDR), the Organization for Economic Cooperation and Development (OECD), the European Statistical Office (Eurostat), the International Labor Office (ILO), and the Luxembourg Income Study (LIS) have produced several cross-national studies of the incidence of poverty in recent years. A large subset of these studies is based on LIS data.¹

For purposes of international comparisons, poverty is almost always a relative concept. A majority of cross-national studies define the poverty threshold as one-half of national median income. In this study, we use the 50 percent of median income to establish our national poverty lines. We could have selected 30 or 40 percent of national median income as our relative poverty threshold because it is closer to the ratio of the official U.S. poverty line to median U.S. household (pretax) cash income. This ratio was only 27–28 percent in 2000, as compared to 50 percent in 1963 (Smeeding 2006, Appendix Table 1). Alternatively, the United Kingdom and the European Union have selected a poverty rate of 60 percent of the median income (Eurostat 2000, Atkinson et al. 2002). However, we

have decided to stay with the conventional 50 percent level in most of our analyses. Previous research suggests more or less the same results in terms of country rankings regardless of the measure chosen (Smeeding 2006).

While the United States likes to think of itself as using an “absolute” poverty measure, there is no one absolute poverty measure. All poverty measures are, in some sense, relative and are chosen to be appropriate for the context in which they are used. The World Bank and the United Nations Millennium Development movement define poverty in Africa and Latin America using an income threshold of \$1 or \$2 per person per day, and in Central and Eastern Europe a threshold of \$2 or \$3 per day. In contrast, the absolute U.S. poverty line is six to nine times higher than these standards and the European poverty line is almost double the United States line as a percent of median income. While we do not provide absolute poverty comparisons below, they also show the United States as having some of the highest levels of poverty among all rich nations (Smeeding, 2006; Rainwater and Smeeding, 2003).

Other Measurement Issues

Comparisons of poverty across nations with LIS are based on many choices. A poverty line, a measure of resources such as (market and disposable) incomes, and an equivalence scale to adjust for family size are all important precursors to accurate cross-national measurement of poverty status.

Poverty measurement is based on the broadest income definition that still preserves comparability across nations. The best current definition is disposable cash and near-cash income (DPI), which includes all types of money income, minus direct income and payroll taxes and including all cash and near-cash transfers, such as food stamps and cash housing allowances, and refundable tax credits such as the earned income tax credit (EITC).^{2,3} We use this income definition in the analyses that follow.

For international comparisons of poverty, the “household” is the only comparable income-sharing unit available for almost all nations. While the household is the unit used for aggregating income, the person is the unit of analysis. Household income is assumed to be equally shared among individuals within a household. Poverty rates are calculated as the percentage of all persons of each type who are members of households of each type with incomes below the poverty line. We calculate the poverty rate for all persons and for children (17 and under) using this same poverty line.

Equivalence scales are used to adjust household income for differences in needs related to household size and other factors, such as the ages of household members. In the U.S. poverty literature, a set of equivalence scales is implicit in the official poverty lines, but these are neither consistent nor robust (Citro and Michael 1995). For our cross-national analysis of relative poverty rates, however, we use a consistent scale, which is much more commonly used in international analyses. After adjusting household incomes to reflect differences in household size, we compare the resulting adjusted incomes to the 50 percent of median poverty line. The equivalence scale used for this purpose, as in many cross-national studies that include both children and elders, is a single parameter scale with a square-root-of-household-size scale factor.⁴

We do not address either the well-being of the poor in terms of hardships, or mobility in or out of poverty. Several recent cross-national poverty studies suggest that mobility in and out of poverty is lower in the United States than in almost every other rich country (Bradbury, Jenkins, and Micklewright 2001; Goodin et al. 1999).

Data

The data we use for this analysis are taken from the LIS database, which now contains almost 150 household income data files for 32 nations covering the period 1967 to 2002

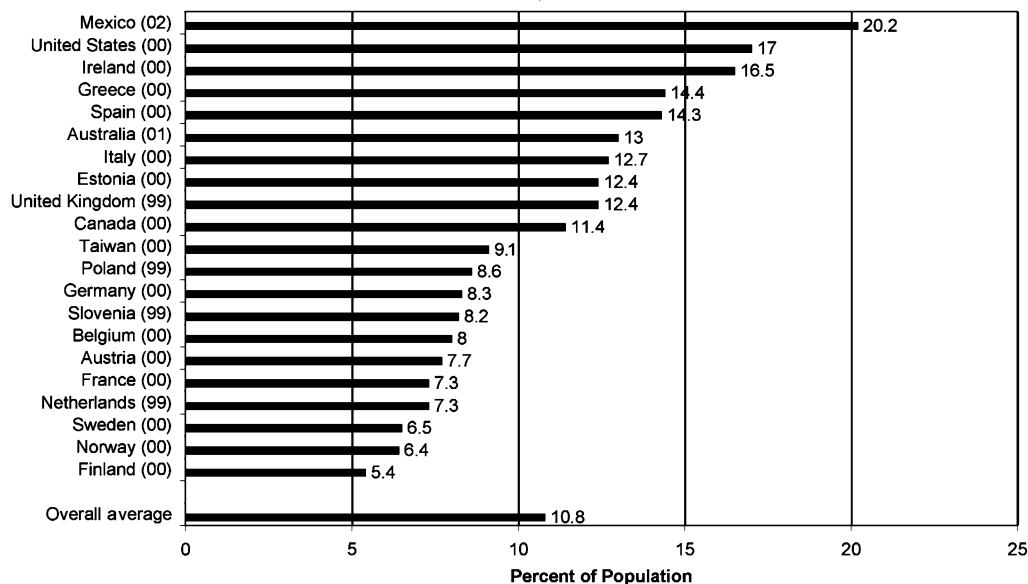


Figure 1. Relative Poverty Rates in Twenty-One Rich Nations at the Turn of the Century for All Persons (Percent of ALL Persons with Disposable Income Less than 50 Percent of Adjusted National Disposable Median Income)

(www.lisproject.org). Using this data one can analyze both the level and trend in poverty and low incomes for a considerable period across a wide range of nations. Because we are computing the level of relative poverty and real living standards for several major policy relevant groups, we have selected 12–21 nations for this paper, each with a recent 1999–2000 LIS database.

One can find relative poverty rates for all of the 32 LIS countries just by going to the LIS website for the “key figures” at: <http://www.lisproject.org/keyfigures/povertytable.htm>.

Results: Level of Overall and Child Poverty

Relative poverty rates in 21 nations are given in Figures 1 and 2 for all persons and for children. The overall poverty rate for all persons using the 50 percent poverty threshold varies from 5.4 percent in Finland to 20.2 percent in Mexico. The poverty rate is 17.0 percent in the United States, the second highest of all nations and the highest of all rich nations.

The average rate of poverty is 10.8 percent across the 21 countries (Figure 1). Higher overall poverty rates are found as one might expect, in Mexico, but also in Anglo-Saxon nations (United States, Australia, Canada, Ireland, the United Kingdom) and southern European nations (Greece, Spain, Italy) with a relatively high level of overall inequality. Still, Australian, Canadian, and British poverty are about 12–13 percent and are, therefore, below the U.S. levels.

The lowest poverty rates are more common in smaller, well-developed, and high-spending welfare states (Sweden, Finland), where they are about 5 or 6 percent. Middle level rates are found in major European countries, where social policies provide more generous support to single mothers and working women (through paid family leave, for example) and where social assistance minimums are high. For instance, the Netherlands, Austria, Belgium, and Germany have poverty rates that are in the 8 to 9 percent range, while France is at 7 percent. Even Taiwan and the former Soviet bloc nations of Estonia, Poland, and

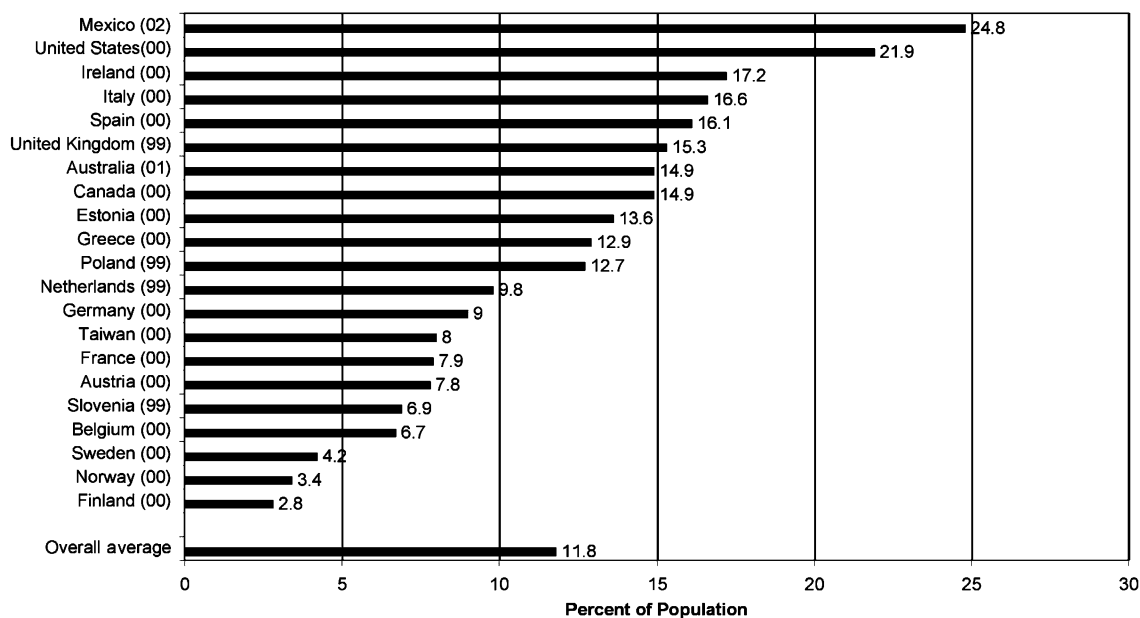


Figure 2. Relative Poverty Rates in Twenty-One Rich Nations at the Turn of the Century for Children* (Percent of CHILDREN with Disposable Income Less than 50 Percent of Adjusted National Disposable Median Income)

*Persons 17 or younger

Source: Author's calculations from Luxembourg Income Study.

Slovenia have much lower poverty rates than does the United States.

On average, child poverty is a slightly larger problem than is overall poverty in these nations, but the cross-national patterns are very similar (Figure 2). After Mexico, the U.S. child poverty rate is at 21.9 percent compared to the 11.8 percent average over these 21 nations. European child poverty rates are lower and Anglo-Saxon rates higher among these nations, but the United States is more than 4.0 percentage points higher than any other rich nation.

Moreover, note that the story is not one of poor immigrants, as two nations with substantially higher fractions of children born to foreigners, Canada and Australia, have child poverty rates that are both 14.9 percent, a full 7 percentage points less than the U.S. rate. We do not present trends in poverty rates here for any nations, but in many nations, though not all, child poverty has risen since

2000. This is most certainly the case in the United States but not in the United Kingdom (see “A Tale of Two Countries” below).

Cross-National Spending Patterns and Relation of Spending and Pay to Poverty

We have seen clearly different patterns of poverty in the United States relative to other nations. What explains these differences? The short answer is that these differences are related to two things: the amount of support we give to the poor, especially the working poor, and the level of wages paid in the United States compared to other nations. Redistributive social expenditures vary greatly across nations. The available evidence indicates that social expenditures (health, education, cash, and near-cash support) as a fraction of total government spending in OECD nations ranges from 0.67 in Australia to 0.90

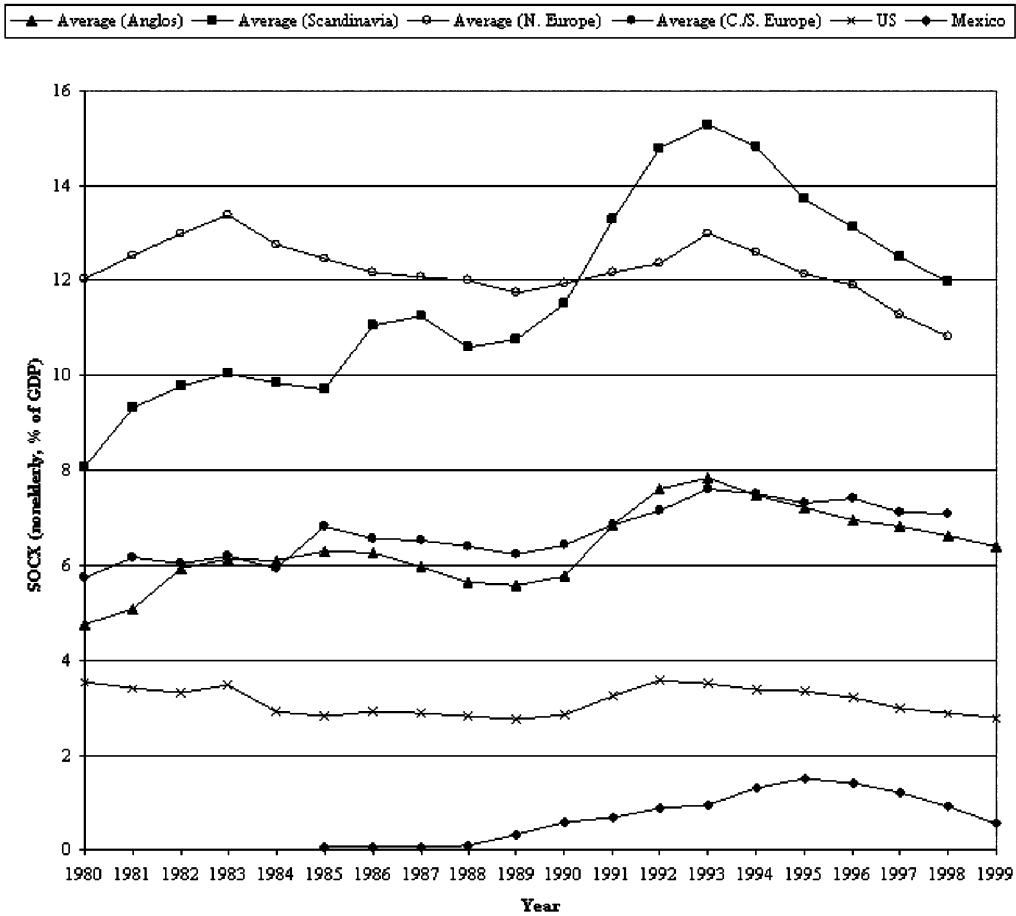


Figure 3. Nonelderly Social Expenditures in 6 Sets of 17 Nations*

*Total Nonelderly Social Expenditures (as percentage of GDP), including all cash plus near cash spending (e.g., food stamps) and public housing but excluding health care and education spending. OECD (2004). Anglos include Australia, UK, Canada; Scandinavia includes Finland, Norway, Sweden; Northern Europe includes Belgium, Denmark, Netherlands; Central/Southern Europe includes Austria, France, Germany, Italy, Greece, Luxembourg, Spain.

in Denmark and Sweden. That is, 67 to 90 percent of all government spending is made up of redistributive cash or in-kind (health, education) benefits (Osberg, Smeeding, and Schwabish 2004). Thus, the topic of social expenditure is about most of what most governments actually do.

We present the trend in nonelderly cash and near-cash (food, housing) benefits for OECD countries back over the past 20 years, using data from the OECD (2004) in comparable format in Figure 3. Here 17 OECD nations—all of the major nations except for the Central and Eastern Europeans—have

been grouped into six clusters: Scandinavia and Finland (Finland, Norway, Sweden); Northern Europe (Belgium, Denmark, Netherlands); Central and Southern Europe (Austria, France, Germany, Italy, Greece, Luxembourg, Spain); Anglo Saxons (Australia, United Kingdom and Canada); the United States; and Mexico.

The Scandinavian and Northern Europeans follow similar patterns—high levels of spending showing responsiveness to the recession of the early 1990s in Sweden and Finland and a tapering off after these events. The Central and Southern Europeans and

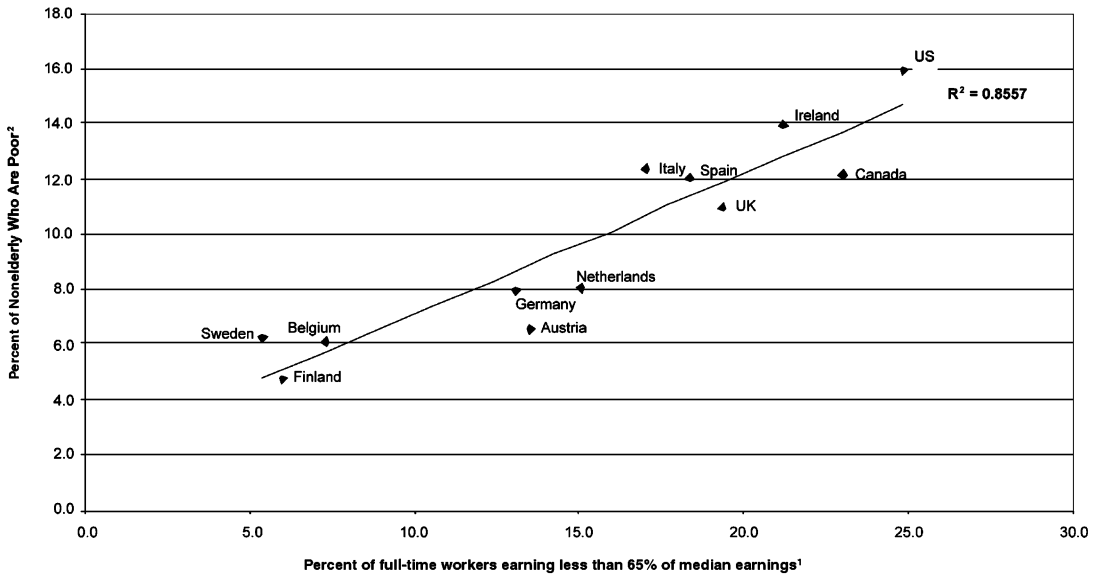


Figure 4. Relationship of Low Pay and Nonelderly Poverty Rates in Twelve Industrialized Countries, circa 2000

Source: OECD database on earnings (as reported in OECD Employment Outlook 2005) and author's tabulations of the LIS data files.

Notes: 1. Data refer to the most recent year for which data could be found (2000 for US, UK, Italy and Canada; 1998 for Germany, Sweden and the Netherlands; 1996 for Austria; 1995 for Belgium, Spain and Ireland). Data for Italy refer to net earnings. Data for Greece are not available.

2. Percentage of persons below 65 in poor households.

the Anglo-Saxon nations show remarkably similar spending patterns, again with expenditures rising in the early 1990s, but overall at a level distinctly below that of the other two groups. The United States is significantly below all these others and, by the late 1990s, is spending at a level closer, expressed as a fraction of GDP per capita, to Mexico than to the other richer OECD nations.

These figures illustrate the wide differences that one can find for both levels and trends in social spending, using figures that abstract from financing of health care, education, and retirement for the elderly. They also correspond very closely to the measures of money and near-money income poverty used in the analytic literature in this area, including that presented above.

A substantial fraction of the variance in nonelderly cross-national poverty rates ap-

pears to be accounted for by the cross-national variation in the incidence of low pay (Figure 4).

Because the United States has the highest proportion of workers in relatively poorly paid jobs,⁵ it also has the highest poverty rate, even among parents who work half time or more (Smeeding, Rainwater, and Burtless, 2001; Smeeding 2006). On the other hand, other countries that have a significantly lower incidence of low-paid employment also have significantly lower poverty rates than does the United States.

But, the prevalence of low-pay workers is, in fact, not the only reliable predictor of poverty rates. While low pay is a good predictor of U.S. poverty rates, and while poorly educated workers do not do well at keeping their families from poverty based on earnings alone, other factors, such as the antipoverty efforts of

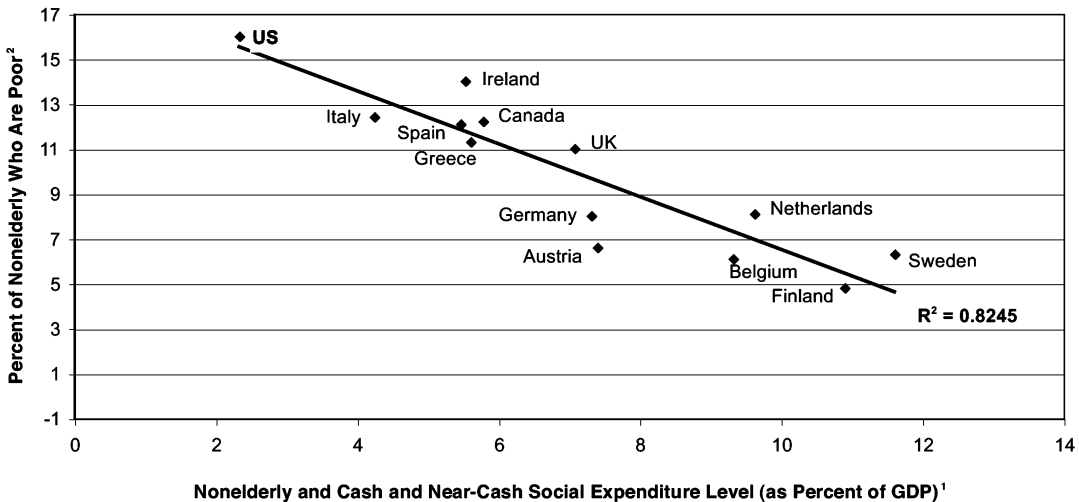


Figure 5. Relationship of Cash Social Expenditures and Nonelderly Poverty Rates in Thirteen Industrialized Countries, circa 2000

Source: OECD (2001) and author's tabulations of the LIS data files.

Notes: 1. Cash and non-cash social expenditures exclude health, education, and social services, but include all forms of cash benefits and near-cash housing subsidies, active labor market program subsidies and other contingent cash and other near-cash benefits. Nonelderly benefits include only those accruing to household head under age 65.

2. Percentage of persons below 65 in poor households.

the government, are also important predictors of the poverty rate (Figure 5). Here we see that higher social spending reduces poverty.

As a result of its low level of spending on cash and near-cash social transfers to the non-aged, the United States again has a very high poverty rate. Even though social spending in general has an inverse correlation with poverty rates, different patterns of social spending can produce different effects on national poverty rates. Antipoverty and social insurance programs are in most respects unique to each country. There is no one kind of program or set of programs that are conspicuously successful in all countries that use them. Social insurance, universal benefits (such as child allowances), and social assistance transfer programs targeted on low-income populations are mixed in different ways in different countries. So, too, are minimum wages, worker preparation and training programs, work-related benefits (such as child care and family leave), and other social benefits.

The United States differs from most nations that achieve lower poverty rates because of its emphasis on work and self-reliance for working-age adults, regardless of the wages workers must accept or the family situation of those workers. For over a decade, U.S. unemployment has been well below the OECD average, and until recently American job growth has been much faster than the OECD average. A strong economy coupled with a few specific antipoverty devices (like the expanded EITC) has produced most of the U.S. overall and child poverty reduction in recent years, though it is decidedly below the effects found in other nations (Smeeding 2006a; 2006b). Simply put, the United States does not spend enough to make up for low levels of pay, and so we end up with a relatively higher poverty rate than do other nations.

A Tale of Two Countries

While acknowledging that the United States has greater poverty than other industrialized

nations, many defenders of American economic and political institutions have argued that inequality plays a crucial role in creating incentives for people to improve their situations through saving, hard work, and investment in education and training. Without the powerful signals provided by big disparities in pay and incomes, the economy would operate less efficiently and average incomes would grow less rapidly. In the long run, poor people might enjoy higher absolute incomes in a society where wide income disparities are tolerated than in one where law and social convention keep income differentials small (Welch 1999). According to this line of argument, wide income disparities may be in the best long-term interest of the poor themselves.⁶ But, of course, there is no evidence that this is true (Burtless and Jencks 2003), and indeed there is some good historical evidence that higher social spending produces higher rates of economic growth and higher social well-being (Lindert 2004).

Our lower-income citizens' "real" incomes are at or below the incomes that most poor people receive in other rich countries that have less inequality (Smeeding 2006a; 2006b). The supposed efficiency advantages of high inequality have not accrued to low-income residents of the United States, at least so far. While the real incomes of families with children did rise in the latter 1990s (Blank and Schoeni 2003), they fell again after 2000 (Congressional Budget Office 2007), and most of the gains have been captured by Americans much further up the income scale, producing a conspicuously wide gap between the incomes of the nation's rich and poor children, elders, and adults.

In recent years, the economy of the United Kingdom and especially of the United States have performed better than many other economies where income disparities are smaller. Employment growth (even since 2001) has been relatively faster, joblessness lower, and economic growth higher than in many other OECD countries where public

policy and social convention have kept income disparities low. Figure 6 compares child poverty in the United States using the same "absolute" or "real" poverty standards—the U.S. official poverty line (about 38 percent of U.S. median income in 1997) with the U.K. poverty line set at 60 percent of U.K. median income in 1996–1997.

In the United States we show official Census Bureau poverty estimates that reflect the current official U.S. income definition. Because U.K. incomes are about 67 percent of U.S. incomes in 1996, this turns out to be just about the same "real" poverty standard.⁷

We noted earlier that these nations were very near the top-ranked nations in terms of child poverty (Figure 2). We also note that child poverty in both nations began to fall without the help of policy from the mid- to the late 1990s, owing mainly to the strong wage growth and tight labor markets in both countries (Figure 6). But beyond 2000, the patterns of child poverty differ completely.

Why so? In 1997, Prime Minister Blair announced his nation would rid itself of high child poverty, and he instituted a wide set of policies to reduce child poverty. In 1999, they began to be implemented. By 2000–2001, child poverty in the United Kingdom (15 percent) was just about the same as in the United States measured against this same "real" resource level. But as we entered the twenty-first century, and when both economies—and especially U.S. economic growth—turned sour, the United Kingdom continued to have policy-driven reductions in child poverty while the U.S. poverty decline stopped and even reversed. The poverty rate for U.K. children fell to 11 percent by 2004, while the official U.S. child poverty rate was 17.6 percent in 2005 according to the U.S. Census estimates. The 2005 estimates for the United Kingdom are not yet available, but projections show an even lower child poverty rate for 2005 once these figures are released.

Five years earlier, low-income U.K. kids were worse off than were U.S. kids in real

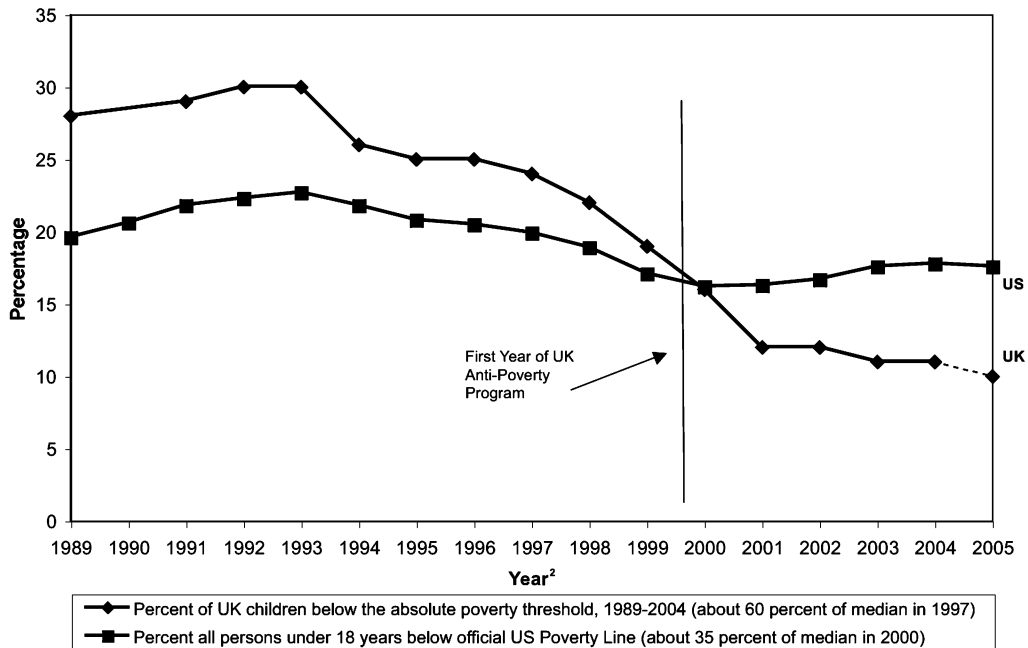


Figure 6. Trends in Absolute Child Poverty: U.K. vs. U.S., 1989–2005

Source: U.S. Census Bureau 2005; U.K. Office of National Statistics 2005.

terms (Rainwater and Smeeding 2003). The reason for the improvement in their situation is that they had a leader who set a national goal of improving living standards and eradicating child poverty in Britain over the next decade—a leader who matched his political rhetoric with some large measure of real and continuing fiscal effort that has already had important impacts not only on poverty, but on consumption and work as well (Gregg, Waldfogel, and Washbrook 2006; Francesconi and van der Klaauw 2007). In Britain, Prime Minister Blair spent an *extra* .9 percent of GDP for low-income families with children after 1999 (Hills 2003). Nine-tenths of a percent of U.S. GDP is about \$120 billion. This is substantially more than we now spend on the EITC, food stamps, child-care support, and Temporary Assistance for Needy Families (TANF) combined. The result of this spending in Britain is that child poverty rates in 2000 were 45 percent below their 1999 level, while real living standards for these children and employment of their mothers also rose

(U.K. Department of Work and Pensions 2005; Francesconi and van der Klaauw 2007). The real consumption levels of these children also increased dramatically over this period (Gregg, Waldfogel, and Washbrook 2006).

Summary and Conclusions

As long as the United States relies almost exclusively on the job market to generate incomes for working-age families, changes in the wage distribution that affect the earnings of less skilled workers will inevitably have a big negative effect on poverty among children and prime-age adults. Welfare reform has pushed many low-income women into the labor market and they have stayed there as TANF rolls continue to fall. Even with the \$25.4 billion spent on TANF today, less than \$10 billion is in the form of cash assistance. The rest is now in the form of child care, transportation assistance, training, and other services (Pear 2003). While the switch from cash to services has undoubtedly helped ac-

count for higher earnings among low-income parents, it has not helped move many of them from poverty. In fact, serious gaps still exist, especially in the child care arena and in family leave policy.

Labor markets alone cannot reduce poverty because not all of the poor can be expected to “earn” their way out of poverty. Single parents with young children, disabled workers, and the unskilled all face significant challenges earning an adequate income, no matter how much they work. The relationship between antipoverty spending and poverty rates is of course complicated, but the evidence discussed above is very suggestive. U.S. poverty rates, especially among children, are high when compared with those in other industrialized countries. Yet U.S. economic performance has also been good compared with that in most other rich countries. As the British have demonstrated, carefully crafted public policy can certainly reduce poverty if the policy effort is made.

Of course, the high direct and indirect costs of our child poverty are now widely recognized in public debate (Holzer 2007). The wisdom of expanding programs targeted at children and poor families depends on one’s values and subjective views about the economic, political, and moral tradeoffs of poverty alleviation. It is hard to argue that the United States cannot afford to do more to help the poor, particularly those who also help themselves via their work efforts. But it has not done so, so far (Shapiro and Parrott 2003; Holzer 2007). If the nation is to be successful in reducing poverty, it will need to do a better job of combining work and benefits targeted to low-wage workers in low-income families (e.g., see Ellwood 2000; Danziger et al. 2000). There is already evidence that such programs produce better outcomes for kids (Clark-Kauffman, Duncan, and Morris 2003; Francesconi and van der Klaauw 2007; Gregg, Waldfogel, and Washbrook 2006).

Given the political disposition of the American public, a 5 percent overall relative

poverty rate is not a plausible goal. A gradual reduction in the overall poverty rate from 17 percent overall and 21 percent for children, to a level of 10–12 percent using the 50 percent of median standard is certainly feasible, however. Although this rate would represent a considerable achievement by the standards of the United States, it is worth remembering that a 12 percent overall poverty rate is higher than the average overall and average child poverty rates in the 21 nations examined here, and would put us just below the poverty levels of our Irish, Australian, British, and Canadian counterparts.

NOTES

1. For UNICEF see UNICEF (2005) and Bradbury and Jäntti (2005); for the United Nations see United Nations Development Programme (1999); for the OECD, see Förster and Pellizzari (2000); for the European Union, see Eurostat (2000); and, for LIS, Jäntti and Danziger (2000), Smeeding (forthcoming, 2006), and Rainwater and Smeeding (2003).

2. See Atkinson, Rainwater, and Smeeding (1995) and Canberra Group (2001) for more on this income definition and its robustness across nations. Note that the use of this “LIS” disposable income concept is not unique to LIS alone. Eurostat and OECD have independently made comparisons of income poverty and inequality across nations using identical or very similar measures of net disposable income.

3. This income definition differs from the Census income definition used in most poverty studies. Still, the internationally comparable measure of income does not subtract work-related expenses or medical care spending. In particular, there is no account for provision of or costs of childcare. The EITC and similar refundable tax credits and near-cash benefits, such as food stamps and cash housing allowances, are included in this income measure, however, as are direct taxes paid.

4. Formally, adjusted disposable income (ADPI) is equal to *un*adjusted household income (DPI) divided by household size (*S*) raised to an exponential value (*e*), $ADPI = DPI/S^e$. We assume the value of *e* is 0.5. To determine whether a household is poor under the relative poverty measure, we compare its ADPI to 50 percent of the national median ADPI. National median ADPI is calculated by converting

all incomes into ADPI and then taking the median of this “adjusted” income distribution. The equivalence scale that we employ is robust, especially when comparing families of different size and structure (e.g., elders and children). See Atkinson, Rainwater, and Smeeding (1995) for detailed and exhaustive documentation of these sensitivities.

5. There are no figures for low pay in any other of the nations studied here, especially Mexico.

6. A lucid presentation and analysis of this viewpoint can be found in Okun (1975). See also Welch (1999).

7. Notice that these estimates are entirely consistent with those presented in Figure 2 earlier for the United Kingdom 1999 and United States 2000, using the LIS data. The difference is that we can go beyond the LIS to later years now using these comparable figures for these two nations alone.

REFERENCES

- Alesina, Alberto, Edward Glaeser, and Bruce Sacerdote. 2005. “Work and Leisure in the U.S. and Europe: Why So Different?” NBER Working Paper No. 11278. Cambridge, MA: National Bureau of Economic Research. April.
- Arjona, Roman, Maxime Ladaïque, and Mark Pearson. 2001. “Growth, Inequality, and Social Protection.” DEELSA/ELSA/WD(2001)4. JT00110378. Labour Market and Social Policy Occasional Paper No. 51. Paris: OECD. June 29.
- Atkinson, Anthony B., Bea Cantillon, Eric Marlier, and Brian Nolan. 2002. *Social Indicators: The EU and Social Inclusion*. Oxford: Oxford University Press.
- Atkinson, Anthony B., Bea Cantillon, Eric Marlier, and Brian Nolan. 2005. “Taking Forward the EU Social Inclusion Process.” European Union, 11-05-0180-E. Luxembourg: Council of the European Union. http://www.ceps.lu/eu2005_lu/report/final_report.pdf.
- Atkinson, Anthony B., Lee Rainwater, and Timothy M. Smeeding. 1995. “Income Distribution in OECD Countries: Evidence from the Luxembourg Income Study (LIS).” *Social Policy Studies* No. 18. Paris: OECD. October.
- Banks, James, Richard F. Disney, Alan Duncan, and John VanReenen. 2005. March. “The Internationalisation of Public Welfare.” *Economic Journal* 115 (502): C62–C81.
- Björklund, Anders, and Richard Freeman. 1997. “Generating Equality and Eliminating Poverty—The Swedish Way.” In *The Welfare State in Transition: Reforming the Swedish Model*, edited by Richard B. Freeman, Robert Topel, and Birgitta Swedenborg. Chicago: University of Chicago Press.
- Blank, Rebecca M., and Robert Schoeni. 2003. May. “Changes in the Distribution of Children’s Family Income over the 1990s.” *American Economic Review* 93 (2): 304–308.
- Bradbury, Bruce, and Markus Jantti. 2005. March. “Child Poverty, Labor Markets, and Public Policies across Industrialized Countries.” Social Policy Research Centre. University of New South Wales, Sydney, Australia.
- Bradbury, Bruce, Stephen P. Jenkins, and John Micklewright. 2001. *The Dynamics of Child Poverty in Industrialised Countries*. Cambridge: Cambridge University Press.
- Burtless, Gary, and Christopher Jencks. 2003. “American Inequality and Its Consequences.” In *Agenda for the Nation*, edited by Henry Aaron, James M. Lindsay, and Pietro Nivola. Washington, DC: Brookings Institution, 61–108.
- Canberra Group. 2001. *Expert Group on Household Income Statistics: Final Report and Recommendations*. Ottawa: Statistics Canada.
- Citro, Constance F., and Robert T. Michael. 1995. *Measuring Poverty: A New Approach*. Washington, DC: National Academies Press.
- Clark-Kaufmann, Elizabeth, Greg J. Duncan, and Pamela Morris. 2003. May. “How Welfare Policies Affect Child and Adolescent Achievement.” *American Economic Review* 93 (2): 299–303.
- Congressional Budget Office. 2007. May. “Changes in the Economic Resources of Low-Income Households with Children.” Publication 2602. Washington, DC.
- Danziger, Sheldon, Colleen M. Heflin, Mary E. Corcoran, Elizabeth Oltmans, and Hui-Chen Wang. 2002. “Does It Pay to Move from Welfare to Work?” *Journal of the Association for Policy Analysis and Management* 21 (Fall): 671–692.
- Ellwood, David T. 2000. August. “Anti-Poverty Policy for Families in the Next Century: From Welfare to Work—and Worries.” *Journal of Economic Perspectives* 14 (1): 187–198.
- Eurostat. 1998. “Recommendations of the Task Force on Social Exclusion and Poverty Statistics.” CPS 98/31/2. Thirty-first Meeting of the Statistical Programme Committee. Luxembourg: Statistical Office of the European Communities. November 26–27.
- Eurostat. 2000. *European Social Statistics: Income, Poverty and Social Exclusion*. KS-29-00-181-EN-I. Luxembourg: European Statistical Office.
- Förster, Michael F., and Michele Pellizzari. 2000. August. “Trends and Driving Factors in Income

- Distribution and Poverty in the OECD Area.” Labour Market and Social Policy—Occasional Paper No. 42. Paris: OECD.
- Francesconi, Marco, and Wilbert van der Klaauw. 2007. “The Socioeconomic Consequences of ‘In-Work’ Benefit Reform for British Lone Mothers.” *Journal of Human Resources* 42 (1): 1–31.
- Goodin, Robert E., Bruce Headey, Ruud Muffels, and Henk-Jan Dirven. 1999. *The Real Worlds of Welfare Capitalism*. Cambridge: Cambridge University Press.
- Gottschalk, Peter, and Timothy M. Smeeding. 2000. “Empirical Evidence on Income Inequality in Industrialized Countries.” In *Handbook of Income Distribution*, vol. 1, edited by Anthony B. Atkinson and François Bourguignon. Amsterdam: North-Holland, 261–308.
- Gregg, Paul, Jane Waldfogel, and Elizabeth Washbrook. 2006. “Family Expenditures Post Welfare Reform in the UK: Are Low Income Families Starting to Catch Up?” *Labour Economics* 13 (6): 721–746.
- Hills, John. 2003. “The Blair Government and Child Poverty: An Extra One Percent for Children in the United Kingdom.” In *One Percent for the Kids: New Policies, Brighter Futures for America’s Children*, edited by Isabel V. Sawhill. Washington, DC: Brookings Institution.
- Holzer, Harry J. 2007. “The Economic Costs of Child Poverty: Testimony Before the U.S. House Committee on Ways and Means.” January 24. <http://waysandmeans.house.gov/hearings.asp?formmode=detail&hearing=508>.
- Jäntti, Markus, and Sheldon Danziger. 2000. “Income Poverty in Advanced Countries.” In *Handbook of Income Distribution*, Vol. 1, edited by Anthony B. Atkinson and François Bourguignon. Amsterdam: North-Holland, 309–378.
- Lindert, Peter H. 2004. *Growing Public: Volume 1, The Story: Social Spending and Economic Growth Since the Eighteenth Century*. Cambridge: Cambridge University Press.
- OECD. 2004. “Social Expenditure Database (SOCX) 1980–2001.” Paris: OECD.
- OECD. 2005. “Employment Outlook.” http://www.oecd.org/document/1/0,2340,en_2649_201185_34855489_1_1_1_1,00.html.
- Okun, Arthur M. 1975. *Equality and Efficiency: The Big Tradeoff*. Washington, DC: Brookings Institution.
- Osberg, Lars, Timothy M. Smeeding, and Jonathan Schwabish. 2004. “Income Distribution and Public Social Expenditure: Theories, Effects, and Evidence.” In *Social Inequality*, edited by Kathryn Neckerman. New York: Russell Sage Foundation, 821–860.
- Pear, Robert. 2003. “Welfare Spending Shows Huge Shift from Checks to Service.” *New York Times*, October 12, A1.
- Rainwater, Lee, and Timothy M. Smeeding. 2003. *Poor Kids in a Rich Country: America’s Children in Comparative Perspective*. New York: Russell Sage Foundation.
- Shapiro, Isaac, and Sharon Parrott. 2003. August. “Are Policies That Assist Low-Income Workers Receiving Appropriate Priority?” Washington, DC: Center on Budget and Policy Priorities.
- Smeeding, Timothy. 2006a. “Government Programs and Social Outcomes: Comparison of the United States and Other Rich Nations.” In *Public Policy and the Income Distribution*, edited by Alan J. Auerbach, David Card, and John M. Quigley. New York: Russell Sage Foundation.
- _____. 2006b. “Poor People in a Rich Nation: The United States in Comparative Perspective.” *Journal of Economic Perspectives* 20 (1): 69–90.
- Smeeding, Timothy M., Lee Rainwater, and Gary Burtless. 2001. “United States Poverty in a Cross-National Context.” In *Understanding Poverty*, edited by Sheldon H. Danziger and Robert H. Haveman. New York and Cambridge, MA: Russell Sage Foundation and Harvard University Press, 162–189.
- UNICEF Innocenti Research Centre. 2005. “Child Poverty in Rich Countries 2005.” Innocenti Report Card No. 6. Florence: UNICEF. <http://www.unicef-icdc.org/pdf/poverty.pdf>.
- UK Department of Work and Pensions. 2004. “Opportunity for All: Sixth Annual Report 2004.” London: Department of Work and Pensions.
- United Nations Development Programme. 2005. “Human Development Report 2004: Cultural Liberty in Today’s Diverse World.” New York: UNDP. http://hdr.undp.org/reports/global/2004/pdf/hdr04_complete.pdf.
- United Nations Development Programme. 1999. *Human Development Report 1999: Globalization with a Human Face*. New York: UNDP. <http://hdr.undp.org/reports/global/1999/en/>.
- U.S. Census Bureau. 2006. August. “Money, Income, Poverty, and Health Insurance in the United States, 2005.” P-60. 278. Washington, DC.
- Welch, Finis. 1999. “In Defense of Inequality.” *American Economic Review* 89 (3): 1–17.

■ THE SOURCES OF POVERTY AND THE UNDERCLASS ■

37 ■ William Julius Wilson

Jobless Poverty

A New Form of Social Dislocation in the Inner-City Ghetto

In September 1996 my book, *When Work Disappears: The World of the New Urban Poor*, was published. It describes a new type of poverty in our nation's metropolises: poor, segregated neighborhoods in which a majority of adults are either unemployed or have dropped out of the labor force altogether. What is the effect of these "jobless ghettos" on individuals, families, and neighborhoods? What accounts for their existence? I suggest several factors and conclude with policy recommendations: a mix of public-and-private sector projects is more effective than relying on a strategy of employer subsidies.

The Research Studies

When Work Disappears was based mainly on three research studies conducted in Chicago between 1986 and 1993. The first of these three studies included a variety of data: a random survey of nearly 2,500 poor and non-poor African American, Latino, and white residents in Chicago's poor neighborhoods; a more focused survey of 175 participants who were reinterviewed and answered open-ended questions; a survey of 179 employers selected to reflect distribution of employment across

industry and firm size in the Chicago metropolitan areas; and comprehensive ethnographic research, including participant-observation research and life-history interviews by ten research assistants in a representative sample of inner-city neighborhoods.

The first of the two remaining projects also included extensive data: a survey of a representative sample of 546 black mothers and up to two of their adolescent children (aged eleven to sixteen—or 887 adolescents) in working-class, middle-class, and high-poverty neighborhoods; a survey of a representative sample of 500 respondents from two high-joblessness neighborhoods on the South Side of Chicago; and six focus-group discussions involving the residents and former residents of these neighborhoods.

Jobless Ghettos

The jobless poverty of today stands in sharp contrast to previous periods. In 1950, a substantial portion of the urban black population was poor but they were working. Urban poverty was quite extensive but people held jobs. However, as we entered the 1990s most adults in many inner-city ghetto neighbor-

William Julius Wilson. "Jobless Poverty: A New Form of Social Dislocation in the Inner-City Ghetto," in *A Nation Divided: Diversity, Inequality, and Community in American Society*, edited by Phyllis Moen, Donna Dempster-McClain, and Henry A. Walker, pp. 133–145, 149–150. Copyright © 1999 by Cornell University. Used by permission of the publisher.

hoods were not working. For example, in 1950 a significant majority of adults held jobs in a typical week in the three neighborhoods that represent the historic core of the Black Belt in Chicago—Douglas, Grand Boulevard, and Washington Park. But by 1990, only four in ten in Douglas worked in a typical week, one in three in Washington Park, and one in four in Grand Boulevard.¹ In 1950, 69 percent of all males aged fourteen and older who lived in these three neighborhoods worked in a typical week, and in 1960, 64 percent of this group were so employed. However, by 1990 only 37 percent of all males aged sixteen and over held jobs in a typical week in these three neighborhoods.

The disappearance of work has had negative effects not only on individuals and families, but on the social life of neighborhoods as well. Inner-city joblessness is a severe problem that is often overlooked or obscured when the focus is mainly on poverty and its consequences. Despite increases in the concentration of poverty since 1970, inner cities have always featured high levels of poverty. But the levels of inner-city joblessness reached during the first half of the 1990s were unprecedented.

Joblessness versus Informal Work Activity

I should note that when I speak of “joblessness” I am not solely referring to official unemployment. The unemployment rate represents only the percentage of workers in the *official* labor force—that is, those who are *actively* looking for work. It does not include those who are outside of or have dropped out of the labor market, including the nearly six million males aged twenty-five to sixty who appeared in the census statistics but were not recorded in the labor market statistics in 1990 (Thurow 1990).

These uncounted males in the labor market are disproportionately represented in the inner-city ghettos. Accordingly, in *When Work Disappears*, I use a more appropriate

measure of joblessness, a measure that takes into account both official unemployment and non-labor-force participation. That measure is the employment-to-population ratio, which corresponds to the percentage of adults aged sixteen and older who are working. Using the employment-to-population ratio we find, for example, that in 1990 only one in three adults aged sixteen and older held a job in the ghetto poverty areas of Chicago, areas representing roughly 425,000 men, women, and children. And in the ghetto tracts of the nation’s one hundred largest cities, for every ten adults who did not hold a job in a typical week in 1990 there were only six employed persons (Kasarda 1993).

The consequences of high neighborhood joblessness are more devastating than those of high neighborhood poverty. A neighborhood in which people are poor but employed is much different than a neighborhood in which people are poor and jobless. *When Work Disappears* shows that many of today’s problems in the inner-city ghetto neighborhoods—crime, family dissolution, welfare, low levels of social organization, and so on—are fundamentally a consequence of the disappearance of work.

It should be clear that when I speak of the disappearance of work, I am referring to the declining involvement in or lack of attachment to the formal labor market. It could be argued that, in the general sense of the term, “joblessness” does not necessarily mean “nonwork.” In other words, to be officially unemployed or officially outside the labor market does not mean that one is totally removed from all forms of work activity. Many people who are officially jobless are nonetheless involved in informal kinds of work activity, ranging from unpaid housework to work that draws income from the informal or illegal economies.

Housework is work, baby-sitting is work, even drug dealing is work. However, what contrasts work in the formal economy with

work activity in the informal and illegal economies is that work in the formal economy is characterized by, indeed calls for, greater regularity and consistency in schedules and hours. Work schedules and hours are formalized. The demands for discipline are greater. It is true that some work activities outside the formal economy also call for discipline and regular schedules. Several studies reveal that the social organization of the drug industry is driven by discipline and a work ethic, however perverse.² However, as a general rule, work in the informal and illegal economies is far less governed by norms or expectations that place a premium on discipline and regularity. For all these reasons, when I speak of the disappearance of work, I mean work in the formal economy, work that provides a framework for daily behavior because of the discipline, regularity, and stability that it imposes.

Effect of Joblessness on Routine and Discipline

In the absence of regular employment, a person lacks not only a place in which to work and the receipt of regular income but also a coherent organization of the present—that is, a system of concrete expectations and goals. Regular employment provides the anchor for the spatial and temporal aspects of daily life. It determines where you are going to be and when you are going to be there. In the absence of regular employment, life, including family life, becomes less coherent. Persistent unemployment and irregular employment hinder rational planning in daily life, a necessary condition of adaptation to an industrial economy (Bourdieu 1965).

Thus, a youngster who grows up in a family with a steady breadwinner and in a neighborhood in which most of the adults are employed will tend to develop some of the disciplined habits associated with stable or steady employment—habits that are reflected in the behavior of his or her parents and of other neighborhood adults. These might in-

clude attachment to a routine, a recognition of the hierarchy found in most work situations, a sense of personal efficacy attained through the routine management of financial affairs, endorsement of a system of personal and material rewards associated with dependability and responsibility, and so on. Accordingly, when this youngster enters the labor market, he or she has a distinct advantage over the youngsters who grow up in households without a steady breadwinner and in neighborhoods that are not organized around work—in other words, a milieu in which one is more exposed to the less disciplined habits associated with casual or infrequent work.

With the sharp recent rise of solo-parent families, black children who live in inner-city households are less likely to be socialized in a work environment for two main reasons. Their mothers, saddled with child-care responsibilities, can prevent a slide deeper into poverty by accepting welfare. Their fathers, removed from family responsibilities and obligations, are more likely to become idle as a response to restricted employment opportunities, which further weakens their influence in the household and attenuates their contact with the family. In short, the social and cultural responses to joblessness are reflected in the organization of family life and patterns of family formation; there they have implications for labor-force attachment as well.

Given the current policy debates that assign blame to the personal shortcomings of the jobless, we need to understand their behavior as responses and adaptations to chronic subordination, including behaviors that have evolved into cultural patterns. The social actions of the jobless—including their behavior, habits, skills, styles, orientations, attitudes—ought not to be analyzed as if they are unrelated to the broader structure of their opportunities and constraints that have evolved over time. This is not to argue that individuals and groups lack the freedom to make their own choices, engage in certain conduct, and develop certain styles and ori-

entations; but I maintain that their decisions and actions occur within a context of constraints and opportunities that are drastically different from those in middle-class society.

Explanations of the Growth of Jobless Ghettos

What accounts for the growing proportion of jobless adults in inner-city communities? An easy explanation would be racial segregation. However, a race-specific argument is not sufficient to explain recent changes in such neighborhoods. After all, these historical Black Belt neighborhoods were *just as segregated by skin color in 1950* as they are today, yet the level of employment was much higher then. One has to account for the ways in which racial segregation interacts with other changes in society to produce the recent escalating rates of joblessness. Several factors stand out: the decreasing demand for low-skilled labor, the suburbanization of jobs, the social deterioration of ghetto neighborhoods, and negative employer attitudes. I discuss each of these factors next.

Decreasing Demand for Low-Skilled Labor

The disappearance of work in many inner-city neighborhoods is in part related to the nationwide decline in the fortunes of low-skilled workers. The sharp decline in the relative demand for unskilled labor has had a more adverse effect on blacks than on whites because a substantially larger proportion of African Americans are unskilled. Although the number of skilled blacks (including managers, professionals, and technicians) has increased sharply in the last several years, the proportion of those who are unskilled remains large, because the black population, burdened by cumulative experiences of racial restrictions, was overwhelmingly unskilled just several decades ago (Schwartzman 1997).³

The factors involved in the decreased relative demand for unskilled labor include

changes in skill-based technology, the rapid growth in college enrollment that increased the supply and reduced the relative cost of skilled labor, and the growing internationalization of economic activity, including trade liberalization policies, which reduced the price of imports and raised the output of export industries (Schwartzman 1997). The increased output of export industries aids skilled workers, simply because they are heavily represented in export industries. But increasing imports, especially those from developing countries that compete with labor-intensive industries (for example, apparel, textile, toy, footwear, and some manufacturing industries), hurts unskilled labor (Schwartzman 1997).

Accordingly, inner-city blacks are experiencing a more extreme form of the economic marginality that has affected most unskilled workers in America since 1980. Unfortunately, there is a tendency among policy makers, black leaders, and scholars alike to separate the economic problems of the ghetto from the national and international trends affecting American families and neighborhoods. If the economic problems of the ghetto are defined solely in racial terms they can be isolated and viewed as only requiring race-based solutions as proposed by those on the left, or as only requiring narrow political solutions with subtle racial connotations (such as welfare reform), as strongly proposed by those on the right.

Overemphasis on Racial Factors

Race continues to be a factor that aggravates inner-city black employment problems as we shall soon see. But the tendency to overemphasize the racial factors obscures other more fundamental forces that have sharply increased inner-city black joblessness. As the late black economist Vivian Henderson put it several years ago, “[I]t is as if racism having put blacks in their economic place steps aside to watch changes in the economy destroy that place” (Henderson 1975, 54). To repeat,

the concentrated joblessness of the inner-city poor represents the most dramatic form of the growing economic dislocations among the unskilled stemming in large measure from changes in the organization of the economy, including the global economy.

Suburbanization of Jobs

But inner-city workers face an additional problem: the growing suburbanization of jobs. Most ghetto residents cannot afford an automobile and therefore have to rely on public transit systems that make the connection between inner-city neighborhoods and suburban job locations difficult and time consuming.

Although studies based on data collected before 1970 showed no consistent or convincing effects on black employment as a consequence of this spatial mismatch, the employment of inner-city blacks relative to suburban blacks has clearly deteriorated since then. Recent research (conducted mainly by urban labor economists) strongly shows that the decentralization of employment is continuing and that employment in manufacturing, most of which is already suburbanized, has decreased in central cities, particularly in the Northeast and Midwest (Holzer 1996).

Blacks living in central cities have less access to employment (as measured by the ratio of jobs to people and the average travel time to and from work) than do central-city whites. Moreover, unlike most other groups of workers across the urban-suburban divide, less-educated central-city blacks receive lower wages than suburban blacks who have similar levels of education. And the decline in earnings of central-city blacks is related to the decentralization of employment—that is, the movement of jobs from the cities to the suburbs—in metropolitan areas (Holzer 1996).

Social Deterioration of Ghetto Neighborhoods

Changes in the class, racial, and demographic composition of inner-city neighborhoods have also contributed to the high per-

centage of jobless adults in these neighborhoods. Because of the steady out-migration of more advantaged families, the proportion of nonpoor families and prime-age working adults has decreased sharply in the typical inner-city ghetto since 1970 (Wilson 1987). In the face of increasing and prolonged joblessness, the declining proportion of nonpoor families and the overall depopulation has made it increasingly difficult to sustain basic neighborhood institutions or to achieve adequate levels of social organization. The declining presence of working- and middle-class blacks has also deprived ghetto neighborhoods of key structural and cultural resources. Structural resources include residents with income high enough to sustain neighborhood services, and cultural resources include conventional role models for neighborhood children.

On the basis of our research in Chicago, it appears that what many high jobless neighborhoods have in common is a relatively high degree of social integration (high levels of local neighboring while being relatively isolated from contacts in the broader mainstream society) and low levels of informal social control (feelings that they have little control over their immediate environment, including the environment's negative influences on their children). In such areas, not only are children at risk because of the lack of informal social controls, they are also disadvantaged because the social interaction among neighbors tends to be confined to those whose skills, styles, orientations, and habits are not as conducive to promoting positive social outcomes (academic success, pro-social behavior, employment in the formal labor market, etc.) as those in more stable neighborhoods. Although the close interaction among neighbors in such areas may be useful in devising strategies, disseminating information, and developing styles of behavior that are helpful in a ghetto milieu (teaching children to avoid eye-to-eye contact with strangers and to develop a tough demeanor

in the public sphere for self-protection), they may be less effective in promoting the welfare of children in society at large.

Despite being socially integrated, the residents in Chicago's ghetto neighborhoods shared a feeling that they had little informal social control over the children in their environment. A primary reason is the absence of a strong organizational capacity or an institutional resource base that would provide an extra layer of social organization in their neighborhoods. It is easier for parents to control the behavior of the children in their neighborhoods when a strong institutional resource base exists and when the links between community institutions such as churches, schools, political organizations, businesses, and civic clubs are strong or secure. The higher the density and stability of formal organizations, the less illicit activities such as drug trafficking, crime, prostitution, and the formation of gangs can take root in the neighborhood.

Few Community Institutions

A weak institutional resource base is what distinguishes high jobless inner-city neighborhoods from stable middle-class and working-class areas. As one resident of a high jobless neighborhood on the South Side of Chicago put it, "Our children, you know, seems to be more at risk than any other children there is, because there's no library for them to go to. There's not a center they can go to, there's no field house that they can go into. There's nothing. There's nothing at all." Parents in high jobless neighborhoods have a much more difficult task controlling the behavior of their adolescents and preventing them from getting involved in activities detrimental to pro-social development. Given the lack of organizational capacity and a weak institutional base, some parents choose to protect their children by isolating them from activities in the neighborhood, including avoiding contact and interaction with neighborhood families. Wherever possible, and often with great

difficulty when one considers the problems of transportation and limited financial resources, they attempt to establish contacts and cultivate relations with individuals, families, and institutions, such as church groups, schools, and community recreation programs, outside their neighborhood. A note of caution is necessary, though. It is just as indefensible to treat inner-city residents as super heroes who overcome racist oppression as it is to view them as helpless victims. We should, however, appreciate the range of choices, including choices representing cultural influences that are available to inner-city residents who live under constraints that most people in the larger society do not experience.

Effect of Joblessness on Marriage and Family

It is within the context of labor-force attachment that the public policy discussion on welfare reform and family values should be couched. The research that we have conducted in Chicago suggests that as employment prospects recede, the foundation for stable relationships becomes weaker over time. More permanent relationships such as marriage give way to temporary liaisons that result in broken unions, out-of-wedlock pregnancies, and, to a lesser extent, separation and divorce. The changing norms concerning marriage in the larger society reinforce the movement toward temporary liaisons in the inner city, and therefore economic considerations in marital decisions take on even greater weight. Many inner-city residents have negative outlooks toward marriage, outlooks that are developed in and influenced by an environment featuring persistent joblessness.

The disrupting effect of joblessness on marriage and family causes poor inner-city blacks to be even more disconnected from the job market and discouraged about their role in the labor force. The economic marginality of the ghetto poor is cruelly reinforced, therefore, by conditions in the neighborhoods in which they live.

Negative Employer Attitudes

In the eyes of employers in metropolitan Chicago, the social conditions in the ghetto render inner-city blacks less desirable as workers, and therefore many are reluctant to hire them. One of the three studies that provided the empirical foundation for *When Work Disappears* included a representative sample of employers in the greater Chicago area who provided entry-level jobs. An overwhelming majority of these employers, both white and black, expressed negative views about inner-city ghetto workers, and many stated that they were reluctant to hire them. For example, a president of an inner-city manufacturing firm expressed a concern about employing residents from certain inner-city neighborhoods:

If somebody gave me their address, uh, Cabrini Green I might unavoidably have some concerns.

Interviewer: What would your concerns be?

That the poor guy probably would be frequently unable to get to work and . . . I probably would watch him more carefully even if it wasn't fair, than I would with somebody else. I know what I should do though is recognize that here's a guy that is trying to get out of his situation and probably will work harder than somebody else who's already out of there and he might be the best one around here. But I, I think I would have to struggle accepting that premise at the beginning. (Wilson 1996, field notes)

In addition to qualms about the neighborhood milieu of inner-city residents, the employers frequently mentioned concerns about applicants' language skills and educational training. An employer from a computer software firm in Chicago expressed the view "that in many businesses the ability to meet the public is paramount and you do not talk street talk to the buying public. Almost all your black welfare people talk street talk.

And who's going to sit them down and change their speech patterns?" (Wilson 1996, field notes) A Chicago real estate broker made a similar point:

A lot of times I will interview applicants who are black, who are sort of lower class. . . . They'll come to me and I cannot hire them because their language skills are so poor. Their speaking voice for one thing is poor . . . they have no verbal facility with the language . . . and these . . . you know, they just don't know how to speak and they'll say "salesmens" instead of "salesmen" and that's a problem. . . . They don't know punctuation, they don't know how to use correct grammar, and they cannot spell. And I can't hire them. And I feel bad about that and I think they're being very disadvantaged by the Chicago Public School system. (Wilson 1996, field notes)

Another respondent defended his method of screening out most job applicants on the telephone on the basis of their use of "grammar and English":

I have every right to say that that's a requirement for this job. I don't care if you're pink, black, green, yellow or orange, I demand someone who speaks well. You want to tell me that I'm a bigot, fine, call me a bigot. I know blacks, you don't even know they're black. (Wilson 1996, field notes)

Finally, an inner-city banker claimed that many blacks in the ghetto "simply cannot read. When you're talking our type of business, that disqualifies them immediately, we don't have a job here that doesn't require that somebody have minimum reading and writing skills" (Wilson 1996, field notes).

How should we interpret the negative attitudes and actions of employers? To what extent do they represent an aversion to blacks *per se* and to what degree do they reflect judgments based on the job-related skills and

training of inner-city blacks in a changing labor market? I should point out that the statements made by the African American employers concerning the qualifications of inner-city black workers did not differ significantly from those of the white employers. Whereas 74 percent of all the white employers who responded to the open-ended questions expressed negative views of the job-related traits of inner-city blacks, 80 percent of the black employers did so as well.

This raises a question about the meaning and significance of race in certain situations—in other words, how race intersects with other factors. A key hypothesis in this connection is that given the recent shifts in the economy, employers are looking for workers with a broad range of abilities: “hard” skills (literacy, numerical ability, basic mechanical ability, and other testable attributes) and “soft” skills (personalities suitable to the work environment, good grooming, group-oriented work behaviors, etc.). While hard skills are the product of education and training—benefits that are apparently in short supply in inner-city schools—soft skills are strongly tied to culture, and are therefore shaped by the harsh environment of the inner-city ghetto. For example, our research revealed that many parents in the inner-city ghetto neighborhoods of Chicago wanted their children not to make eye-to-eye contact with strangers and to develop a tough demeanor when interacting with people on the streets. While such behaviors are helpful for survival in the ghetto, they hinder successful interaction in mainstream society.

Statistical Discrimination

If employers are indeed reacting to the difference in skills between white and black applicants, it becomes increasingly difficult to discuss the motives of employers: are they rejecting inner-city black applicants out of overt racial discrimination or on the basis of qualifications?

Nonetheless, many of the selective recruitment practices do represent what economists

call “statistical discrimination”: employers make assumptions about the inner-city black workers *in general* and reach decisions based on those assumptions before they have had a chance to review systematically the qualifications of an individual applicant. The net effect is that many black inner-city applicants are never given the chance to prove their qualifications on an individual level because they are systematically screened out by the selective recruitment process.

Statistical discrimination, although representing elements of class bias against poor workers in the inner city, is clearly a matter of race both directly and indirectly. Directly, the selective recruitment patterns effectively screen out far more black workers from the inner city than Hispanic or white workers from the same types of backgrounds. But indirectly, race is also a factor, even in those decisions to deny employment to inner-city black workers on the basis of objective and thorough evaluations of their qualifications. The hard and soft skills among inner-city blacks that do not match the current needs of the labor market are products of racially segregated communities, communities that have historically featured widespread social constraints and restricted opportunities.

Thus the job prospects of inner-city workers have diminished not only because of the decreasing relative demand for low-skilled labor in the United States economy, the suburbanization of jobs, and the social deterioration of ghetto neighborhoods, but also because of negative employer attitudes. This combination of factors presents a real challenge to policy makers. Indeed, considering the narrow range of social policy options in the “balance-the-budget” political climate, how can we immediately alleviate the inner-city jobs problem—a problem which will undoubtedly grow when the new welfare reform bill takes full effect and creates a situation that will be even more harmful to inner-city children and adolescents?

Public Policy Dilemmas

What are the implications of these studies on public policy? A key issue is public-sector employment. If firms in the private sector cannot hire or refuse to hire low-skilled adults who are willing to take minimum-wage jobs, then policy makers should consider a policy of public-sector employment-of-last-resort. Indeed, until current changes in the labor market are reversed or until the skills of the next generation of workers can be upgraded before they enter the labor market, many workers, especially those who are not in the official labor force, will not be able to find jobs unless the government becomes an employer-of-last-resort (Danziger and Gottschalk 1995). This argument applies especially to low-skilled inner-city black workers. It is bad enough that they face the problem of shifts in labor-market demand shared by all low-skilled workers; it is even worse that they confront negative employer perceptions about their work-related skills and attitudes.

For all these reasons, the passage of the 1996 welfare reform bill, which did not include a program of job creation, could have very negative social consequences in the inner city. Unless something is done to enhance the employment opportunities of inner-city welfare recipients who reach the time limit for the receipt of welfare, they will flood a pool already filled with low-skilled, jobless workers. . . .

West Virginia, a state that has been plagued with a severe shortage of work opportunities, has provided community service jobs to recipients of welfare for several years. In Wisconsin, Governor Thompson's welfare reform plan envisions community service jobs for many parents in the more depressed areas of the state, and the New Hope program in Milwaukee provides community service jobs for those unable to find employment in the private sector (Center on Budget and Policy Priorities 1996). It is especially important that this mixed strategy include a plan to make *adequate* monies

available to localities or communities with high jobless and welfare dependency rates.

Obviously, as more people become employed and gain work experience, they will have a better chance of finding jobs in the private sector when jobs become available. The attitudes of employers toward inner-city workers could change, in part because they would be dealing with job applicants who have steady work experience and who could furnish references from their previous supervisors. Children are more likely to be socialized in a work-oriented environment and to develop the job readiness skills that are seen as important even for entry-level jobs.

Thus, given the recent welfare reform legislation, *adequate* strategies to enhance the employment opportunities of inner-city residents should be contemplated, strategies that would be adequately financed and designed to address the employment problems of low-skilled workers not only in periods of tight labor markets, but, even more important, in periods when the labor market is slack.

NOTES

1. The figures on adult employment are based on calculations from data provided by the 1990 U.S. Bureau of the Census (1993) and the *Local Community Fact Book for Chicago—1950* (1953) and the *Local Community Fact Book for Chicago—1960* (1963). The adult employment rates represent the number of employed individuals (aged fourteen and older in 1950 and sixteen and older in 1990) among the total number of adults in a given area. Those who are not employed include both the individuals who are members of the labor force but are not working and those who have dropped out or are not part of the labor force.

2. See, for example, Bourgois (1995) and Venkatesh (1996).

3. The economist David Schwartzman defines "unskilled workers to include operators, fabricators, and laborers, and those in service occupations, including private household workers, those working in protective service occupations, food service, and cleaning and building service." On the basis of this definition he estimates that 80 percent of all black workers and 38 percent of all white workers were unskilled in 1950. By 1990, 46 percent of black work-

ers and 27 percent of white workers were employed in unskilled occupations (Schwartzman 1997).

BIBLIOGRAPHY

- Bourdieu, Pierre. 1965. *Travail et travailleurs en Algérie*. Paris: Editions Mouton.
- Bourgois, Philippe. 1995. *In Search of Respect: Selling Crack in El Barrio*. New York: Cambridge University Press.
- Center on Budget and Policy Priorities. 1996. *The Administration's \$3 Billion Jobs Proposal*. Washington, DC: Center on Budget and Policy Priorities.
- Danziger, Sheldon H., and Peter Gottschalk. 1995. *America Unequal*. Cambridge, MA: Harvard University Press.
- Henderson, Vivian. 1975. "Race, Economics, and Public Policy." *Crisis* 83 (Fall):50–55.
- Holzer, Harry J. 1996. *What Employers Want: Job Prospects for Less-Educated Workers*. New York: Russell Sage.
- Kasarda, John D. 1993. "Inner-City Concentrated Poverty and Neighborhood Distress: 1970–1990." *Housing Policy Debate* 4(3): 253–302.
- Local Community Fact Book for Chicago—1950*. 1953. Chicago: Community Inventory, University of Chicago.
- Local Community Fact Book for Chicago—1960*. 1963. Chicago: Community Inventory, University of Chicago.
- Schwartzman, David. 1997. *Black Unemployment: Part of Unskilled Unemployment*. Westport, CT: Greenwood.
- Thurow, Lester. 1990. "The Crusade That's Killing Prosperity." *American Prospect* March/April:54–59.
- U.S. Bureau of the Census. 1993. *Census of Population: Detailed Characteristics of the Population*. Washington, DC: U.S. Government Printing Office.
- Venkatesh, Sudhir. 1996. "Private Lives, Public Housing: An Ethnography of the Robert Taylor Homes." Ph.D. dissertation, University of Chicago.
- Wilson, William Julius. 1987. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press.
- _____. 1996. *When Work Disappears: The World of the New Urban Poor*. New York: Alfred A. Knopf.

38 ■ Douglas S. Massey and Nancy A. Denton

American Apartheid

Segregation and the Making of the Underclass

It is quite simple. As soon as there is a group area then all your uncertainties are removed and that is, after all, the primary purpose of this Bill [requiring racial segregation in housing].

—Minister of the Interior, Union of South Africa
legislative debate on the Group Areas Act of 1950

During the 1970s and 1980s a word disappeared from the American vocabulary.¹ It was not in the speeches of politicians decrying the multiple ills besetting American cities. It was not spoken by government offi-

cial responsible for administering the nation's social programs. It was not mentioned by journalists reporting on the rising tide of homelessness, drugs, and violence in urban America. It was not discussed by foundation

Douglas S. Massey and Nancy A. Denton. *American Apartheid: Segregation and the Making of the Underclass*, pp. 1–9, 12–16, 239–241. Copyright © 1993 by the President and Fellows of Harvard University Press. Reprinted by permission of the publisher.

executives and think-tank experts proposing new programs for unemployed parents and unwed mothers. It was not articulated by civil rights leaders speaking out against the persistence of racial inequality; and it was nowhere to be found in the thousands of pages written by social scientists on the urban underclass. The word was segregation.

Most Americans vaguely realize that urban America is still a residentially segregated society, but few appreciate the depth of black segregation or the degree to which it is maintained by ongoing institutional arrangements and contemporary individual actions. They view segregation as an unfortunate holdover from a racist past, one that is fading progressively over time. If racial residential segregation persists, they reason, it is only because civil rights laws passed during the 1960s have not had enough time to work or because many blacks still prefer to live in black neighborhoods. The residential segregation of blacks is viewed charitably as a “natural” outcome of impersonal social and economic forces, the same forces that produced Italian and Polish neighborhoods in the past and that yield Mexican and Korean areas today.

But black segregation is not comparable to the limited and transient segregation experienced by other racial and ethnic groups, now or in the past. No group in the history of the United States has ever experienced the sustained high level of residential segregation that has been imposed on blacks in large American cities for the past fifty years. This extreme racial isolation did not just happen; it was manufactured by whites through a series of self-conscious actions and purposeful institutional arrangements that continue today. Not only is the depth of black segregation unprecedented and utterly unique compared with that of other groups, but it shows little sign of change with the passage of time or improvements in socioeconomic status.

If policymakers, scholars, and the public have been reluctant to acknowledge segre-

gation's persistence, they have likewise been blind to its consequences for American blacks. Residential segregation is not a neutral fact; it systematically undermines the social and economic well-being of blacks in the United States. Because of racial segregation, a significant share of black America is condemned to experience a social environment where poverty and joblessness are the norm, where a majority of children are born out of wedlock, where most families are on welfare, where educational failure prevails, and where social and physical deterioration abound. Through prolonged exposure to such an environment, black chances for social and economic success are drastically reduced.

Deleterious neighborhood conditions are built into the structure of the black community. They occur because segregation concentrates poverty to build a set of mutually reinforcing and self-feeding spirals of decline into black neighborhoods. When economic dislocations deprive a segregated group of employment and increase its rate of poverty, socioeconomic deprivation inevitably becomes more concentrated in neighborhoods where that group lives. The damaging social consequences that follow from increased poverty are spatially concentrated as well, creating uniquely disadvantaged environments that become progressively isolated—geographically, socially, and economically—from the rest of society.

The effect of segregation on black well-being is structural, not individual. Residential segregation lies beyond the ability of any individual to change; it constrains black life chances irrespective of personal traits, individual motivations, or private achievements. For the past twenty years this fundamental fact has been swept under the rug by policymakers, scholars, and theorists of the urban underclass. Segregation is the missing link in prior attempts to understand the plight of the urban poor. As long as blacks continue to

be segregated in American cities, the United States cannot be called a race-blind society.

The Forgotten Factor

The present myopia regarding segregation is all the more startling because it once figured prominently in theories of racial inequality. Indeed, the ghetto was once seen as central to black subjugation in the United States. In 1944 Gunnar Myrdal wrote in *An American Dilemma* that residential segregation “is basic in a mechanical sense. It exerts its influence in an indirect and impersonal way: because Negro people do not live near white people, they cannot . . . associate with each other in the many activities founded on common neighborhood. Residential segregation . . . becomes reflected in uni-racial schools, hospitals, and other institutions” and creates “an artificial city . . . that permits any prejudice on the part of public officials to be freely vented on Negroes without hurting whites.”²

Kenneth B. Clark, who worked with Gunnar Myrdal as a student and later applied his research skills in the landmark *Brown v. Topeka* school integration case, placed residential segregation at the heart of the U.S. system of racial oppression. In *Dark Ghetto*, written in 1965, he argued that “the dark ghetto’s invisible walls have been erected by the white society, by those who have power, both to confine those who have *no* power and to perpetuate their powerlessness. The dark ghettos are social, political, educational, and—above all—economic colonies. Their inhabitants are subject peoples, victims of the greed, cruelty, insensitivity, guilt, and fear of their masters.”³

Public recognition of segregation’s role in perpetuating racial inequality was galvanized in the late 1960s by the riots that erupted in the nation’s ghettos. In their aftermath, President Lyndon B. Johnson appointed a commission chaired by Governor Otto Kerner of Illinois to identify the causes of the violence

and to propose policies to prevent its recurrence. The Kerner Commission released its report in March 1968 with the shocking admonition that the United States was “moving toward two societies, one black, one white—separate and unequal.”⁴ Prominent among the causes that the commission identified for this growing racial inequality was residential segregation.

In stark, blunt language, the Kerner Commission informed white Americans that “discrimination and segregation have long permeated much of American life; they now threaten the future of every American.”⁵ “Segregation and poverty have created in the racial ghetto a destructive environment totally unknown to most white Americans. What white Americans have never fully understood—but what the Negro can never forget—is that white society is deeply implicated in the ghetto. White institutions created it, white institutions maintain it, and white society condones it.”⁶

The report argued that to continue present policies was “to make permanent the division of our country into two societies; one, largely Negro and poor, located in the central cities; the other, predominantly white and affluent, located in the suburbs.”⁷ Commission members rejected a strategy of ghetto enrichment coupled with abandonment of efforts to integrate, an approach they saw “as another way of choosing a permanently divided country.”⁸ Rather, they insisted that the only reasonable choice for America was “a policy which combines ghetto enrichment with programs designed to encourage integration of substantial numbers of Negroes into the society outside the ghetto.”⁹

America chose differently. Following the passage of the Fair Housing Act in 1968, the problem of housing discrimination was declared solved, and residential segregation dropped off the national agenda. Civil rights leaders stopped pressing for the enforcement of open housing, political leaders

increasingly debated employment and educational policies rather than housing integration, and academicians focused their theoretical scrutiny on everything from culture to family structure, to institutional racism, to federal welfare systems. Few people spoke of racial segregation as a problem or acknowledged its persisting consequences. By the end of the 1970s residential segregation became the forgotten factor in American race relations.¹⁰

While public discourse on race and poverty became more acrimonious and more focused on divisive issues such as school busing, racial quotas, welfare, and affirmative action, conditions in the nation's ghettos steadily deteriorated.¹¹ By the end of the 1970s, the image of poor minority families mired in an endless cycle of unemployment, unwed childbearing, illiteracy, and dependency had coalesced into a compelling and powerful concept: the urban underclass.¹² In the view of many middle-class whites, inner cities had come to house a large population of poorly educated single mothers and jobless men—mostly black and Puerto Rican—who were unlikely to exit poverty and become self-sufficient. In the ensuing national debate on the causes for this persistent poverty, four theoretical explanations gradually emerged: culture, racism, economics, and welfare.

Cultural explanations for the underclass can be traced to the work of Oscar Lewis, who identified a "culture of poverty" that he felt promoted patterns of behavior inconsistent with socioeconomic advancement.¹³ According to Lewis, this culture originated in endemic unemployment and chronic social immobility, and provided an ideology that allowed poor people to cope with feelings of hopelessness and despair that arose because their chances for socioeconomic success were remote. In individuals, this culture was typified by a lack of impulse control, a strong present-time orientation, and little

ability to defer gratification. Among families, it yielded an absence of childhood, an early initiation into sex, a prevalence of free marital unions, and a high incidence of abandonment of mothers and children.

Although Lewis explicitly connected the emergence of these cultural patterns to structural conditions in society, he argued that once the culture of poverty was established, it became an independent cause of persistent poverty. This idea was further elaborated in 1965 by the Harvard sociologist and then Assistant Secretary of Labor Daniel Patrick Moynihan, who in a confidential report to the President focused on the relationship between male unemployment, family instability, and the intergenerational transmission of poverty, a process he labeled a "tangle of pathology."¹⁴ He warned that because of the structural absence of employment in the ghetto, the black family was disintegrating in a way that threatened the fabric of community life.

When these ideas were transmitted through the press, both popular and scholarly, the connection between culture and economic structure was somehow lost, and the argument was popularly perceived to be that "people were poor because they had a defective culture." This position was later explicitly adopted by the conservative theorist Edward Banfield, who argued that lower-class culture—with its limited time horizon, impulsive need for gratification, and psychological self-doubt—was primarily responsible for persistent urban poverty.¹⁵ He believed that these cultural traits were largely imported, arising primarily because cities attracted lower-class migrants.

The culture-of-poverty argument was strongly criticized by liberal theorists as a self-serving ideology that "blamed the victim."¹⁶ In the ensuing wave of reaction, black families were viewed not as weak but, on the contrary, as resilient and well-adapted survivors in an oppressive and racially prej-

udiced society.¹⁷ Black disadvantages were attributed not to a defective culture but to the persistence of institutional racism in the United States. According to theorists of the underclass such as Douglas Glasgow and Alphonso Pinkney, the black urban underclass came about because deeply imbedded racist practices within American institutions—particularly schools and the economy—effectively kept blacks poor and dependent.¹⁸

As the debate on culture versus racism ground to a halt during the late 1970s, conservative theorists increasingly captured public attention by focusing on a third possible cause of poverty: government welfare policy. According to Charles Murray, the creation of the underclass was rooted in the liberal welfare state.¹⁹ Federal antipoverty programs altered the incentives governing the behavior of poor men and women, reducing the desirability of marriage, increasing the benefits of unwed childbearing, lowering the attractiveness of menial labor, and ultimately resulted in greater poverty.

A slightly different attack on the welfare state was launched by Lawrence Mead, who argued that it was not the generosity but the permissiveness of the U.S. welfare system that was at fault.²⁰ Jobless men and unwed mothers should be required to display “good citizenship” before being supported by the state. By not requiring anything of the poor, Mead argued, the welfare state undermined their independence and competence, thereby perpetuating their poverty.

This conservative reasoning was subsequently attacked by liberal social scientists, led principally by the sociologist William Julius Wilson, who had long been arguing for the increasing importance of class over race in understanding the social and economic problems facing blacks.²¹ In his 1987 book *The Truly Disadvantaged*, Wilson argued that persistent urban poverty stemmed primarily from the structural transformation of the inner-city economy.²² The decline of

manufacturing, the suburbanization of employment, and the rise of a low-wage service sector dramatically reduced the number of city jobs that paid wages sufficient to support a family, which led to high rates of joblessness among minorities and a shrinking pool of “marriageable” men (those financially able to support a family). Marriage thus became less attractive to poor women, unwed childbearing increased, and female-headed families proliferated. Blacks suffered disproportionately from these trends because, owing to past discrimination, they were concentrated in locations and occupations particularly affected by economic restructuring.

Wilson argued that these economic changes were accompanied by an increase in the spatial concentration of poverty within black neighborhoods. This new geography of poverty, he felt, was enabled by the civil rights revolution of the 1960s, which provided middle-class blacks with new opportunities outside the ghetto.²³ The outmigration of middle-class families from ghetto areas left behind a destitute community lacking the institutions, resources, and values necessary for success in post-industrial society. The urban underclass thus arose from a complex interplay of civil rights policy, economic restructuring, and a historical legacy of discrimination.

Theoretical concepts such as the culture of poverty, institutional racism, welfare disincentives, and structural economic change have all been widely debated. None of these explanations, however, considers residential segregation to be an important contributing cause of urban poverty and the underclass. In their principal works, Murray and Mead do not mention segregation at all;²⁴ and Wilson refers to racial segregation only as a historical legacy from the past, not as an outcome that is institutionally supported and actively created today.²⁵ Although Lewis mentions segregation sporadically in his writings, it is not assigned a central role in the set of structural

factors responsible for the culture of poverty, and Banfield ignores it entirely. Glasgow, Pinkney, and other theorists of institutional racism mention the ghetto frequently, but generally call not for residential desegregation but for race-specific policies to combat the effects of discrimination in the schools and labor markets. In general, then, contemporary theorists of urban poverty do not see high levels of black-white segregation as particularly relevant to understanding the underclass or alleviating urban poverty.²⁶

The purpose of this book is to redirect the focus of public debate back to issues of race and racial segregation, and to suggest that they should be fundamental to thinking about the status of black Americans and the origins of the urban underclass. Our quarrel is less with any of the prevailing theories of urban poverty than with their systematic failure to consider the important role that segregation has played in mediating, exacerbating, and ultimately amplifying the harmful social and economic processes they treat.

We join earlier scholars in rejecting the view that poor urban blacks have an autonomous “culture of poverty” that explains their failure to achieve socioeconomic success in American society. We argue instead that residential segregation has been instrumental in creating a structural niche within which a deleterious set of attitudes and behaviors—a culture of segregation—has arisen and flourished. Segregation created the structural conditions for the emergence of an oppositional culture that devalues work, schooling, and marriage and that stresses attitudes and behaviors that are antithetical and often hostile to success in the larger economy. Although poor black neighborhoods still contain many people who lead conventional, productive lives, their example has been overshadowed in recent years by a growing concentration of poor, welfare-dependent families that is an inevitable result of residential segregation.

We readily agree with Douglas, Pinkney, and others that racial discrimination is widespread and may even be institutionalized within large sectors of American society, including the labor market, the educational system, and the welfare bureaucracy. We argue, however, that this view of black subjugation is incomplete without understanding the special role that residential segregation plays in enabling all other forms of racial oppression. Residential segregation is the institutional apparatus that supports other racially discriminatory processes and binds them together into a coherent and uniquely effective system of racial subordination. Until the black ghetto is dismantled as a basic institution of American urban life, progress ameliorating racial inequality in other arenas will be slow, fitful, and incomplete.

We also agree with William Wilson’s basic argument that the structural transformation of the urban economy undermined economic supports for the black community during the 1970s and 1980s.²⁷ We argue, however, that in the absence of segregation, these structural changes would not have produced the disastrous social and economic outcomes observed in inner cities during these decades. Although rates of black poverty were driven up by the economic dislocations Wilson identifies, it was segregation that confined the increased deprivation to a small number of densely settled, tightly packed, and geographically isolated areas.

Wilson also argues that concentrated poverty arose because the civil rights revolution allowed middle-class blacks to move out of the ghetto. Although we remain open to the possibility that class-selective migration did occur,²⁸ we argue that concentrated poverty would have happened during the 1970s with or without black middle-class migration. Our principal objection to Wilson’s focus on middle-class out-migration is not that it did not occur,

but that it is misdirected: focusing on the flight of the black middle class deflects attention from the real issue, which is the limitation of black residential options through segregation.

Middle-class households—whether they are black, Mexican, Italian, Jewish, or Polish—always try to escape the poor. But only blacks must attempt their escape within a highly segregated, racially segmented housing market. Because of segregation, middle-class blacks are less able to escape than other groups, and as a result are exposed to more poverty. At the same time, because of segregation no one will move into a poor black neighborhood except other poor blacks. Thus both middle-class blacks and poor blacks lose compared with the poor and middle class of other groups: poor blacks live under unrivaled concentrations of poverty and affluent blacks live in neighborhoods that are far less advantageous than those experienced by the middle class of other groups.

Finally, we concede Murray's general point that federal welfare policies are linked to the rise of the urban underclass, but we disagree with his specific hypothesis that generous welfare payments, by themselves, discouraged employment, encouraged unwed child-bearing, undermined the strength of the family, and thereby caused persistent poverty.²⁹ We argue instead that welfare payments were only harmful to the socioeconomic well-being of groups that were residentially segregated. As poverty rates rose among blacks in response to the economic dislocations of the 1970s and 1980s, so did the use of welfare programs. Because of racial segregation, however, the higher levels of welfare receipt were confined to a small number of isolated, all-black neighborhoods. By promoting the spatial concentration of welfare use, therefore, segregation created a residential environment within which welfare dependency was the norm, leading to

the intergenerational transmission and broader perpetuation of urban poverty.

Coming to Terms with American Apartheid

Our fundamental argument is that racial segregation—and its characteristic institutional form, the black ghetto—are the key structural factors responsible for the perpetuation of black poverty in the United States. Residential segregation is the principal organizational feature of American society that is responsible for the creation of the urban underclass. . . . It can be shown that any increase in the poverty rate of a residentially segregated group leads to an immediate and automatic increase in the geographic concentration of poverty. When the rate of minority poverty is increased under conditions of high segregation, all of the increase is absorbed by a small number of neighborhoods. When the same increase in poverty occurs in an integrated group, the added poverty is spread evenly throughout the urban area, and the neighborhood environment that group members face does not change much.

During the 1970s and 1980s, therefore, when urban economic restructuring and inflation drove up rates of black and Hispanic poverty in many urban areas, underclass communities were created only where increased minority poverty coincided with a high degree of segregation—principally in older metropolitan areas of the northeast and the midwest. Among Hispanics, only Puerto Ricans developed underclass communities, because only they were highly segregated; and this high degree of segregation is directly attributable to the fact that a large proportion of Puerto Ricans are of African origin.

The interaction of intense segregation and high poverty leaves black neighborhoods extremely vulnerable to fluctuations in the urban economy, because any dislocation that causes an upward shift in black poverty rates

will also produce a rapid change in the concentration of poverty and, hence, a dramatic shift in the social and economic composition of black neighborhoods. The concentration of poverty, for example, is associated with the wholesale withdrawal of commercial institutions and the deterioration or elimination of goods and services distributed through the market.

Neighborhoods, of course, are dynamic and constantly changing, and given the high rates of residential turnover characteristic of contemporary American cities, their well-being depends to a great extent on the characteristics and actions of their residents. Decisions taken by one actor affect the subsequent decisions of others in the neighborhood. In this way isolated actions affect the well-being of the community and alter the stability of the neighborhood.

Because of this feedback between individual and collective behavior, neighborhood stability is characterized by a series of thresholds, beyond which various self-perpetuating processes of decay take hold. Above these thresholds, each actor who makes a decision that undermines neighborhood well-being makes it increasingly likely that other actors will do the same. Each property owner who decides not to invest in upkeep and maintenance, for example, lowers the incentive for others to maintain their properties. Likewise, each new crime promotes psychological and physical withdrawal from public life, which reduces vigilance within the neighborhood and undermines the capacity for collective organization, making additional criminal activity more likely.

Segregation increases the susceptibility of neighborhoods to these spirals of decline. During periods of economic dislocation, a rising concentration of black poverty is associated with the simultaneous concentration of other negative social and economic conditions. Given the high levels of racial segregation characteristic of American urban areas, increases in

black poverty such as those observed during the 1970s can only lead to a concentration of housing abandonment, crime, and social disorder, pushing poor black neighborhoods beyond the threshold of stability.

By building physical decay, crime, and social disorder into the residential structure of black communities, segregation creates a harsh and extremely disadvantaged environment to which ghetto blacks must adapt. In concentrating poverty, moreover, segregation also concentrates conditions such as drug use, joblessness, welfare dependency, teenage childbearing, and unwed parenthood, producing a social context where these conditions are not only common but the norm. By adapting to this social environment, ghetto dwellers evolve a set of behaviors, attitudes, and expectations that are sharply at variance with those common in the rest of American society.

As a direct result of the high degree of racial and class isolation created by segregation, for example, Black English has become progressively more distant from Standard American English, and its speakers are at a clear disadvantage in U.S. schools and labor markets. Moreover, the isolation and intense poverty of the ghetto provides a supportive structural niche for the emergence of an "oppositional culture" that inverts the values of middle-class society. Anthropologists have found that young people in the ghetto experience strong peer pressure not to succeed in school, which severely limits their prospects for social mobility in the larger society. Quantitative research shows that growing up in a ghetto neighborhood increases the likelihood of dropping out of high school, reduces the probability of attending college, lowers the likelihood of employment, reduces income earned as an adult, and increases the risk of teenage childbearing and unwed pregnancy.

Segregation also has profound political consequences for blacks, because it so isolates them geographically that they are the only

ones who benefit from public expenditures in their neighborhoods. The relative integration of most ethnic groups means that jobs or services allocated to them will generally benefit several other groups at the same time. Integration thus creates a basis for political coalitions and pluralist politics, and most ethnic groups that seek public resources are able to find coalition partners because other groups can anticipate sharing the benefits. That blacks are the only ones to benefit from resources allocated to the ghetto—and are the only ones harmed when resources are removed—makes it difficult for them to find partners for political coalitions. Although segregation paradoxically makes it easier for blacks to elect representatives, it limits their political influence and marginalizes them within the American polity. Segregation prevents blacks from participating in pluralist politics based on mutual self-interest.

Because of the close connection between social and spatial mobility, segregation also perpetuates poverty. One of the primary means by which individuals improve their life chances—and those of their children—is by moving to neighborhoods with higher home values, safer streets, higher-quality schools, and better services. As groups move up the socioeconomic ladder, they typically move up the residential hierarchy as well, and in doing so they not only improve their standard of living but also enhance their chances for future success. Barriers to spatial mobility are barriers to social mobility, and by confining blacks to a small set of relatively disadvantaged neighborhoods, segregation constitutes a very powerful impediment to black socioeconomic progress.

Despite the obvious deleterious consequences of black spatial isolation, policymakers have not paid much attention to segregation as a contributing cause of urban poverty and have not taken effective steps to dismantle the ghetto. Indeed, for most of the past two decades public policies tolerated and

even supported the perpetuation of segregation in American urban areas. Although many political initiatives were launched to combat discrimination and prejudice in the housing and banking industries, each legislative or judicial act was fought tenaciously by a powerful array of people who believed in or benefited from the status quo.

Although a comprehensive open housing bill finally passed Congress under unusual circumstances in 1968, it was stripped of its enforcement provisions as its price of enactment, yielding a Fair Housing Act that was structurally flawed and all but doomed to fail. As documentation of the law's defects accumulated in multiple congressional hearings, government reports, and scholarly studies, little was done to repair the situation until 1988, when a series of scandals and political errors by the Reagan administration finally enabled a significant strengthening of federal antidiscrimination law.

Yet even more must be done to prevent the permanent bifurcation of the United States into black and white societies that are separate and unequal. As of 1990, levels of racial segregation were still extraordinarily high in the nation's large urban areas, particularly those of the north. Segregation has remained high because fair housing enforcement relies too heavily on the private efforts of individual victims of discrimination. Whereas the processes that perpetuate segregation are entrenched and institutionalized, fair housing enforcement is individual, sporadic, and confined to a small number of isolated cases.

As long as the Fair Housing Act is enforced individually rather than systemically, it is unlikely to be effective in overcoming the structural arrangements that support segregation and sustain the ghetto. Until the government throws its considerable institutional weight behind efforts to dismantle the ghetto, racial segregation will persist. . . .

Ultimately, however, dismantling the ghetto and ending the long reign of racial

segregation will require more than specific bureaucratic reforms; it requires a moral commitment that white America has historically lacked. The segregation of American blacks was no historical accident; it was brought about by actions and practices that had the passive acceptance, if not the active support, of most whites in the United States. Although America's apartheid may not be rooted in the legal strictures of its South African relative, it is no less effective in perpetuating racial inequality, and whites are no less culpable for the socioeconomic deprivation that results.

As in South Africa, residential segregation in the United States provides a firm basis for a broader system of racial injustice. The geographic isolation of Africans within a narrowly circumscribed portion of the urban environment—whether African townships or American ghettos—forces blacks to live under extraordinarily harsh conditions and to endure a social world where poverty is endemic, infrastructure is inadequate, education is lacking, families are fragmented, and crime and violence are rampant.³⁰ Moreover, segregation confines these unpleasant by-products of racial oppression to an isolated portion of the urban geography far removed from the experience of most whites. Resting on a foundation of segregation, apartheid not only denies blacks their rights as citizens but forces them to bear the social costs of their own victimization.

Although Americans have been quick to criticize the apartheid system of South Africa, they have been reluctant to acknowledge the consequences of their own institutionalized system of racial separation. The topic of segregation has virtually disappeared from public policy debates; it has vanished from the list of issues on the civil rights agenda; and it has been ignored by social scientists spinning endless theories of the underclass. Residential segregation has become the forgotten factor of American race relations, a minor footnote in the ongoing de-

bate on the urban underclass. Until policy-makers, social scientists, and private citizens recognize the crucial role of America's own apartheid in perpetuating urban poverty and racial injustice, the United States will remain a deeply divided and very troubled society.³¹

NOTES

1. Epigraph from Edgar H. Brookes, *Apartheid: A Documentary Study of Modern South Africa* (London: Routledge and Kegan Paul, 1968), p. 142.

2. Gunnar Myrdal, *An American Dilemma*, vol. 1 (New York: Harper and Brothers, 1944), p. 618; see also Walter A. Jackson, *Gunnar Myrdal and America's Conscience* (Chapel Hill: University of North Carolina Press, 1990), pp. 88–271.

3. Kenneth B. Clark, *Dark Ghetto: Dilemmas of Social Power* (New York: Harper and Row, 1965), p. 11.

4. U.S. National Advisory Commission on Civil Disorders, *The Kerner Report* (New York: Pantheon Books, 1988), p. 1.

5. *Ibid.*

6. *Ibid.*, p. 2.

7. *Ibid.*, p. 22.

8. *Ibid.*

9. *Ibid.*

10. A few scholars attempted to keep the Kerner Commission's call for desegregation alive, but their voices have largely been unheeded in the ongoing debate. Thomas Pettigrew has continued to assert the central importance of residential segregation, calling it the "linchpin" of American race relations; see "Racial Change and Social Policy," *Annals of the American Academy of Political and Social Science* 441 (1979):114–31. Gary Orfield has repeatedly pointed out segregation's deleterious effects on black prospects for education, employment, and socioeconomic mobility; see "Separate Societies: Have the Kerner Warnings Come True?" in Fred R. Harris and Roger W. Wilkins, eds., *Quiet Riots: Race and Poverty in the United States* (New York: Pantheon Books, 1988), pp. 100–122; and "Ghettoization and Its Alternatives," in Paul E. Peterson, ed., *The New Urban Reality* (Washington, D.C.: Brookings Institution, 1985), pp. 161–96.

11. See Thomas B. Edsall and Mary D. Edsall, *Chain Reaction: The Impact of Race, Rights, and Taxes on American Politics* (New York: Norton, 1991).

12. For an informative history of the evolution of the concept of the underclass, see Michael B. Katz, *The Undeserving Poor: From the War on Poverty to the*

War on Welfare (New York: Pantheon, 1989), pp. 185–235.

13. Oscar Lewis, *La Vida: A Puerto Rican Family in the Culture of Poverty—San Juan and New York* (New York: Random House, 1965); “The Culture of Poverty,” *Scientific American* 215 (1966): 19–25; “The Culture of Poverty,” in Daniel P. Moynihan, ed., *On Understanding Poverty: Perspectives from the Social Sciences* (New York: Basic Books, 1968), pp. 187–220.

14. The complete text of this report is reprinted in Lee Rainwater and William L. Yancey, *The Moynihan Report and the Politics of Controversy* (Cambridge: MIT Press, 1967), pp. 39–125.

15. Edward C. Banfield, *The Unheavenly City* (Boston: Little, Brown, 1970).

16. William Ryan, *Blaming the Victim* (New York: Random House, 1971).

17. Carol Stack, *All Our Kin: Strategies of Survival in a Black Community* (New York: Harper and Row, 1974).

18. Douglas C. Glasgow, *The Black Underclass: Poverty, Unemployment, and Entrapment of Ghetto Youth* (New York: Vintage, 1981), p. 11; Alphonso Pinkney, *The Myth of Black Progress* (Cambridge: Cambridge University Press, 1984), pp. 78–80.

19. Charles Murray, *Losing Ground: American Social Policy, 1950–1980* (New York: Basic Books, 1984).

20. Lawrence M. Mead, *Beyond Entitlement: The Social Obligations of Citizenship* (New York: Free Press, 1986).

21. William Julius Wilson, *The Declining Significance of Race: Blacks and Changing American Institutions* (Chicago: University of Chicago Press, 1978).

22. William Julius Wilson, *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy* (Chicago: University of Chicago Press, 1987), pp. 1–108.

23. *Ibid.*, pp. 49–62.

24. The subject indices of *Losing Ground* and *Beyond Entitlement* contain no references at all to residential segregation.

25. The subject index of *The Truly Disadvantaged* contains two references to pre-1960s Jim Crow segregation.

26. Again with the exception of Thomas Pettigrew and Gary Orfield.

27. We have published several studies documenting how the decline of manufacturing, the subur-

banization of jobs, and the rise of low-wage service employment eliminated high-paying jobs for manual workers, drove up rates of black male unemployment, and reduced the attractiveness of marriage to black women, thereby contributing to a proliferation of female-headed families and persistent poverty. See Mitchell L. Eggers and Douglas S. Massey, “The Structural Determinants of Urban Poverty,” *Social Science Research* 20 (1991): 217–55; Mitchell L. Eggers and Douglas S. Massey, “A Longitudinal Analysis of Urban Poverty: Blacks in U.S. Metropolitan Areas between 1970 and 1980,” *Social Science Research* 21 (1992): 175–203.

28. The evidence on the extent of middle-class out-migration from ghetto areas is inconclusive. Because racial segregation does not decline with rising socioeconomic status, out-movement from poor black neighborhoods certainly has not been to white areas. When Kathryn P. Nelson measured rates of black out-migration from local “zones” within forty metropolitan areas, however, she found higher rates of out-movement for middle- and upper-class blacks compared with poor blacks; but her “zones” contained more than 100,000 inhabitants, making them considerably larger than neighborhoods (see “Racial Segregation, Mobility, and Poverty Concentration,” paper presented at the annual meetings of the Population Association of America, Washington, D.C., March 19–23, 1991). In contrast, Edward Gramlich and Deborah Laren found that poor and middle-class blacks displayed about the same likelihood of out-migration from poor census tracts (see “Geographic Mobility and Persistent Poverty,” Department of Economics, University of Michigan, Ann Arbor, 1990).

29. See Eggers and Massey, “A Longitudinal Analysis of Urban Poverty.”

30. See International Defense and Aid Fund for Southern Africa, *Apartheid: The Facts* (London: United Nations Centre against Apartheid, 1983), pp. 15–26.

31. We are not the first to notice the striking parallel between the institutionalized system of racial segregation in U.S. cities and the organized, state-sponsored system of racial repression in South Africa. See John H. Denton, *Apartheid American Style* (Berkeley, Calif.: Diablo Press, 1967); James A. Kushner, “Apartheid in America: An Historical and Legal Analysis of Contemporary Racial Residential Segregation in the United States,” *Howard Law Journal* 22 (1979): 547–60.

39 ■ Anne R. Pebley and Narayan Sastry

Neighborhoods, Poverty, and Children's Well-Being

Research on neighborhood effects suggests that neighborhood characteristics such as poverty, crime, and residential turnover influence several interrelated aspects of the neighborhood environment that, in turn, affect families and children. These mechanisms can be summarized in four categories: child and family-related institutions; social organization and interaction; normative environment; and labor and marriage markets. We will briefly describe each of these.

Child and family-related institutions include schools, child care providers, public libraries, recreational programs and activities (such as music lessons, youth organizations, sports activities, arts and theater activities, and mentoring programs), parks, religious institutions, and social service providers. These institutions play a vital role in the general process of socialization, but many also impart important skills and provide specific services. While the availability and quality of these institutions may be affected directly by public policy (for example, by school improvement programs in poorer neighborhoods), they are also likely to be determined by neighborhood socioeconomic characteristics (Jencks and Mayer 1990; Aber et al. 1997). For example, child care centers and after-school programs may be more readily available, hire better staff, and provide better service in more affluent or well-educated

neighborhoods, because residents demand it and can afford to pay for it. As described later, more socially organized neighborhoods may also be able to demand better institutions through collective action and the political process, even if income and educational levels are low. Poorer neighborhoods may be worse off than others not only because they have weaker institutions but also because the greater needs of families are likely to overtax the existing institutions (Aber et al. 1997).

Neighborhood *social organization and interaction* has recently received considerable attention in research on neighborhood effects. Social disorganization theory suggests that some neighborhood characteristics (such as poverty, ethnic heterogeneity, high residential turnover rates, low homeownership rates, and concentration of recent immigrants) make it harder for residents to establish social ties and agree on the values needed to exercise social control and work together on common goals. As a result, socially disorganized neighborhoods are more difficult, dangerous, and stressful places to live. Parents and children in these neighborhoods are both more likely to participate in deviant behavior (delinquency, crime, violence, substance abuse) and to suffer the consequences of this behavior in others (Shaw and McKay 1969; Sampson, Morenoff, and Gannon-Rowley 2002). Sampson and his colleagues (Sampson, Morenoff, and Earls 1999;

Sampson, Morenoff, and Gannon-Rowley 2002) argue that neighborhood collective efficacy—that is, shared expectations and the involvement of neighborhood residents in active support and social control of children—is key to a positive neighborhood environment for children. In neighborhoods with higher collective efficacy, residents are more likely to monitor and, when necessary, correct children's behavior. They are also more likely to work together on neighborhood problems and to build and maintain strong local institutions.

Two other theoretical perspectives—William Julius Wilson's (1987, 1996) collective socialization model and James Coleman's (1988) social capital theory—suggest related ways in which neighborhood social interaction may be important for children. Collective socialization models posit that neighborhood adults play an important role by monitoring children's behavior (as Sampson and his colleagues emphasize) and by providing role models. For example, Wilson argues that the selective out-migration of middle-class professionals from African American inner-city neighborhoods has resulted in fewer positive role models for the children in those neighborhoods. Social capital models suggest that the key elements are the dense and overlapping social ties among adults and children. In neighborhoods with more social capital, children know that they will be held accountable for their actions and that they can rely on neighborhood adults for support. However, as Wilson (1996) and Sampson, Morenoff, and Earls (1999) note, high levels of social capital can facilitate the enforcement of *both* negative and positive norms and behavior. Wilson (1996, 62) points out, for example, that in neighborhoods “characterized by high levels of individual and family involvement in aberrant behavior,” a high degree of social integration among adults can in fact help to create and reinforce problem behavior among children.

Connections with the world outside the neighborhood may also be important. Es-

pecially in disadvantaged neighborhoods, extralocal social ties can provide access to information about, or assistance with, opportunities, services, or normative feedback from those who move in other social circles (Coleman 1988; Stack 1974; Edin 1991; Tigges, Browne, and Green 1998). Melvin Oliver's (1988) study of social networks in urban African American communities in Los Angeles shows that extralocal social ties vary considerably among neighborhoods. He concludes that in poor neighborhoods the lack of outside social ties may be a significant disadvantage. In recent decades urban sociologists have argued that despatialized social networks have displaced the role of neighborhoods in urban life; neighborhoods are increasingly *unimportant*, they argue, in individuals' lives (South 2001; Fischer 1984; Wellman 1999). However, Barry Wellman (1999, 27), a proponent of this view, admits that “communities have not totally lost their domestic roots. . . . Local relationships are necessary for domestic safety, controlling actual land use, and quickly getting goods and services.” He shows, for example, that much of Toronto residents' telephone contact is with neighbors rather than with extralocal ties (Wellman 1996). As Sampson and his colleagues (2002) point out, social ties among neighbors do not need to be strong or close in order to be effective. In fact, social disorganization theory suggests that neighborhood environments depend on weak and limited ties among neighbors who share a minimum level of trust, agreement on basic standards, and willingness to live by and enforce those standards. Nonetheless, the relative importance for children's development of urban neighborhood environments versus social networks is an empirical question for which we do not yet have complete answers.

The economic models suggest that *labor and marriage markets* are key elements in neighborhood effects on families and children (Duncan and Hoffman 1990; Haveman and Wolfe 1994). Local labor markets,

marriage markets, and, in some neighborhoods, the illicit economy provide constraints and opportunities for neighborhood residents. Market conditions affect adults and adolescents most directly. However, by affecting their *parents'* probabilities of employment and marriage, local markets may have indirect effects on younger children. For example, in neighborhoods with poor labor markets, higher rates of parental unemployment may affect children by increasing stress on parents, depressing household income, and creating a more stressful home environment.

Neighborhood characteristics, such as high levels of marriage, are thought to affect children's well-being both directly, by providing a positive normative environment, strong institutions, effective monitoring and social control, and a supportive climate for children, and indirectly, through effects on parents and the home environment. As noted earlier, neighborhood labor market and marriage market conditions can affect parents' income, family structure, and the home environment. Neighborhoods may also directly affect parenting behavior and family dynamics (Aneshensel and Sucoff 1996; Klebanov et al. 1997; Coulton 1996; Korbin and Coulton 1997). For example, parents in extremely disadvantaged neighborhoods are more likely to exhibit more punitive, authoritarian, and coercive parenting styles and to use corporal punishment (McLoyd 1990; Sampson and Laub 1994) as well as to withdraw emotionally from their children (Klebanov, Brooks-Gunn, and Duncan 1994). These responses are likely to have detrimental consequences for children's emotional, cognitive, and social development, and those consequences may later be reinforced by other negative aspects of neighborhood life.

Several theorists emphasize the key role of the *normative environment* itself in linking neighborhood compositional characteristics (such as poverty or high turnover rates) and child outcomes. Neighborhood norms may

be a consequence of the characteristics of the people who live in the neighborhood—their income level, ethnic background, education, or immigrant experience. Those norms may also be affected by the social organization and interaction and by marriage and labor markets, as described earlier. The central idea in this literature is that the greater the concentration of like-minded people, the stronger the normative climate and the greater the exposure of neighborhood residents to these norms. For example, black children in poor inner-city neighborhoods may be more likely to be exposed to social problems, because the extreme concentration of poverty in inner-city African American neighborhoods since 1970 has created negative normative environments in which behavior considered negative by the middle class is reinforced and valued (Massey, Gross, and Eggers 1991; Massey and Denton 1993; Wilson, 1987, 1996; Fordham and Ogbu 1986). However, this process is not necessarily limited to concentrated-poverty neighborhoods or to negative outcomes—for example, some observers have argued that concentrated immigrant communities in Los Angeles can provide supportive climates for social mobility (Waldinger 1996). The “epidemic” hypothesis (Crane 1991; Case and Katz 1991) is a specific version of theories about normative environments. Jonathan Crane (1991) argues that concentrated-poverty neighborhoods dramatically increase adolescents' exposure to problem behavior and negative norms through contacts with peers. Epidemics of social problems can occur once neighborhoods reach a critically high level of negative social behaviors. . . .

Neighborhood Selection

A serious problem in studying neighborhood effects on children's well-being is the potential endogeneity of neighborhood (and school) characteristics. Endogenous characteristics are independent variables that may

be correlated with unobserved factors not included in the model. In the case of neighborhood selection, the problem is that parents can choose the neighborhood in which they live but they can also affect children's development in many other ways. Thus, parents' attitudes about child development may affect both the type of neighborhood their children grow up in and other factors that affect their children's development, such as parenting and the home environment. Parents who move to help their children escape the influences of gang activity, drug use, teenage pregnancy, or crime, or who choose a neighborhood for the quality of its schools, may also be better parents in other ways. Other neighborhood attributes that may influence both parents' choice of where to live and children's behavior and development include the strength of neighborhood ties and characteristics of other families and children in the neighborhood. To the extent that neighborhood of residence is a choice, all neighborhood characteristics should be treated as endogenous. It is important, however, to understand the source of the endogeneity in order to identify appropriate analytic strategies.

Neighborhood attributes may be endogenous because place of residence is a choice variable and is determined in part by factors that also influence children's behavior and development. Thus, a common set of parent and family characteristics determines both children's behavior and development *and* neighborhood choice. Some of these characteristics, such as household income and parents' education, are measurable and can be controlled in models of children's behavior and development. However, some are unobserved and hence are picked up in the random component of statistical models, where correlation with included regressors—neighborhood characteristics—leads to biased and inconsistent estimates of all model parameters. As Greg Duncan and his colleagues explain (see, for example, Duncan, Connell,

and Klebanov 1997), the problem of neighborhood selection is thus really one of omitted variables. Specific unobserved (omitted) parent and family factors are the parents' cognitive ability and family motivation and aspirations, which may influence the degree to which a family values its children's behavior and development (as well as its choice of place of residence). . . .

Neighborhood-Effects Analyses

The extensive neighborhood-effects literature published since 1990 has generally sought to answer one or more of the following three questions:

1. Are children who grow up in poor neighborhoods worse off than other children?
2. Are disparities in children's welfare by neighborhood poverty level due to differences in their families' characteristics, or do neighborhood conditions themselves play a role?
3. What mechanisms link concentrated-poverty neighborhoods to poorer outcomes for children?

Studies addressing these questions have generally been of two types. The largest group is non-experimental or observational studies, generally based on sample survey data. More recently, several experimental studies have assessed the consequences for poor families of moving into nonpoor neighborhoods. Both types of studies have usually sought to investigate the first two questions—whether children's outcomes differ by neighborhood characteristics and whether this variation persists if family attributes are held constant. A smaller number of studies have attempted to answer the third question by exploring the mechanisms that may link the characteristics of concentrated-poverty neighborhoods to poorer outcomes for children.

Recent reviews by psychologists, economists, and sociologists have thoroughly cataloged and critiqued this literature (Leventhal and Brooks-Gunn 2000; Ginter, Haveman, and Wolfe 2000; Duncan and Raudenbush 1999, 2001; Sampson, Morenoff, and Gannon-Rowley 2002). In this section, we draw on these critiques and our own reading of the literature to summarize the results of non-experimental research on neighborhood effects. We then consider more recent experimental studies.

Observational Studies

Observational studies are typically based on individual and household data from sample surveys linked with census data on the local areas (usually census tracts) in which children and families live. These studies have employed a wide range of study designs, survey datasets, theoretical approaches, neighborhood, family, and outcome measures, and statistical methods. Here we summarize the results of this very diverse group of studies.

First, basic descriptive analyses have shown that many dimensions of children's well-being (including teen sexual behavior, substance abuse, mental health, cognitive and achievement scores, high school completion, youth violence, delinquency, and child abuse) vary significantly by neighborhood income levels and, less often, by other neighborhood characteristics such as residential stability, high school completion rate, female headship, social disorder, and social cohesion. Children and teens living in poorer neighborhoods generally have worse outcomes.

Second, a substantial part of the variation in children's outcomes by neighborhood income level is accounted for by differences in family income and other family characteristics. In other words, when family characteristics such as income, family structure, and parents' educational attainment are held constant, the relationship between children's outcomes and neighborhood income levels is

substantially reduced. Moreover, Ginter and her colleagues (2000) show that the more complete the set of family characteristics that is held constant, the greater the decline in the size and significance of coefficients on neighborhood variables. They conclude that the results of many neighborhood-effects studies are likely due, at least in part, to omitted variables at the family level. Nonetheless, these researchers and others find that some neighborhood characteristics retain significant effects even after extensive controls for family and individual characteristics are introduced.

Third, the size of neighborhood effects on children's outcomes is generally modest and considerably smaller than the effects of family and individual characteristics. For example, in studies reviewed by Leventhal and Brooks-Gunn (2000), neighborhood characteristics accounted for 5 to 10 percent of the variance in children's outcomes. Duncan and Raudenbush (2001, 132) argue, however, that "the degree of neighborhood-based 'action' may still be large enough to be consistent with cost-effective, neighborhood based interventions."

Fourth, results from these studies, not surprisingly, suggest that the effects of neighborhood conditions (net of family SES) vary by type of child outcome investigated (behavior problems, school readiness, teen sex, delinquency) and by the child's age, ethnicity, and gender. For example, Brooks-Gunn, Duncan, and their colleagues examined an extensive set of child development indicators across a broader age range (Brooks-Gunn, Duncan, and Aber 1997). Duncan and Raudenbush (1999) summarize the results as follows: (1) neighborhood effects appear in the preschool years but are most consistent for school-age children; (2) neighborhood effects appear to be stronger for cognitive and achievement outcomes than for behavior and mental health measures; and (3) white children appear to be more affected by neighborhood conditions than African American chil-

dren. Sampson and his colleagues (2002) argue that the evidence of neighborhood effects on crime rates is stronger than the evidence for other types of outcomes.

Fifth, several studies suggest that the presence of affluent neighbors has a greater impact on children's outcomes than neighborhood poverty (Brooks-Gunn, Duncan, and Aber 1997; Duncan and Raudenbush 1999; Sampson, Morenoff, and Earls 1999). However, Ginter and her colleagues (2000) dispute this conclusion. Their reanalysis of the PSID data includes variables indicating the percentage of households with high and low income as well as the income of the child's family relative to that of other families in the neighborhood. Their results suggest that "the income of the family relative to that of its neighbors—rather than the extent to which the neighborhood is populated by high (low) income families—may be the more relevant consideration" (628). This is an important topic for future neighborhood-effects research.

Sixth, reliable methods for assessing neighborhood social and physical environments are not well developed and tested. Those studies that have examined intervening processes have investigated a broad range of potential mechanisms. For example, Sampson and his colleagues (Sampson et al. 1997; Sampson and Raudenbush 1999; Morenoff, Sampson, and Raudenbush 2001) show that informal social control, collective efficacy, and social ties are significantly related to outcomes such as delinquency, crime, and homicide. Scott South and Eric Baumer's (2000; Baumer and South 2001) results suggest that peer attitudes and behaviors account for a substantial proportion of neighborhood effects on adolescents, particularly teen parenthood and sexual activity. Dawn Upchurch and her colleagues (1999) and Carol Aneshensel and Clea Sucoff (1996) show that perceived "ambient hazards" (for example, neighborhood disorder, disorganization, and threats) are an important mediating factor

between neighborhood disadvantage and teen sexual behavior and mental health.

Finally, a few studies have tackled endogenous residential choice using non-experimental data and statistical models. William Evans, Wallace Oates, and Robert Schwab (1992), Anne Case and Lawrence Katz (1991), and Eric Foster and Sara McLanahan (1996) used instrumental variables to eliminate the correlation between unobserved parent attributes and neighborhood variables. However, finding credible and viable instruments is a very difficult task. Instead, Daniel Aaronson (1997, 1998) and Robert Plotnick and Saul Hoffman (1996) have used sibling fixed effects in analyses of educational attainment, adult economic status, and teen pregnancy in the PSID. While Aaronson found significant neighborhood effects once unobserved family characteristics were controlled, Plotnick and Hoffman did not. Aaronson (1998) suggests that the difference in results lies in the types of sibling pairs included and the measurement of neighborhood variables. Gary Solon, Marianne Page, and Greg Duncan (2000) take another approach: they compare correlations for sibling pairs with correlations among neighbors within sampling clusters in the PSID. Their results suggest that the size of neighborhood effects is small and considerably smaller than family effects.

Experimental Studies

More recently, several experimental or quasi-experimental studies have attempted to tackle the issue of endogenous neighborhood selection. The initial effort was the Gautreaux Program, in which low-income African American families from Chicago housing projects were given Section 8 housing vouchers that could be used only in predominantly white or multi-ethnic neighborhoods (typically in suburban areas). The control group was Section 8 voucher recipients who used their vouchers in the city of Chicago. James Rosenbaum (1991, 1995) shows that children who move to the suburbs

rather than cities are less likely to drop out of school and more likely to attend college, have a job, and receive higher pay. However, the study has several methodological limitations, including self-selection into the study and substantial sample attrition.

The Moving To Opportunity (MTO) experiment was a more carefully designed outgrowth of the Gautreaux Program developed by the U.S. Department of Housing and Urban Development (HUD) and implemented by local public housing authorities and non-profits between 1994 and 1999 in Baltimore, Boston, Chicago, Los Angeles, and New York (Brennan 2002). Participants were volunteers from very low-income families with children in public housing or Section 8 project-based housing in inner-city, high-poverty neighborhoods. Each participant family was assigned randomly to one of three groups: the *experimental group*, which received vouchers that could be used only in low-poverty areas plus counseling and assistance locating housing; the *comparison group*, which received geographically unrestricted vouchers and standard housing authority briefings and assistance; and the *control group*, which continued to receive project-based assistance. The study sought to answer two questions: What impact does mobility counseling have on families' residential choices and housing and neighborhood conditions? And what are the effects of neighborhood conditions on the well-being of MTO families?

The follow-up design and analyses of MTO in each city have been conducted by separate groups of researchers using different data collection and analytic strategies. This approach has the disadvantage that it is harder to make comparisons across cities (and hence generalizations beyond each city). But it also has the advantage that the multiple research strategies used provide a richer picture of the experimental process and outcomes. The one commonality among all five sites is that HUD conducted a self-administered baseline

survey of all families who volunteered to participate. Researchers in most study sites conducted follow-up telephone surveys two to three years after families were assigned to treatment groups. The Boston and Los Angeles projects also conducted qualitative studies with a sample of participants. In contrast, analyses of Boston participants have relied on baseline passive and active tracking of respondents and administrative data on arrests and school performance. Furthermore, the project in each city focused on a somewhat different set of children's outcomes.

As in almost all social experiments, the MTO project encountered significant problems in implementation of the experimental treatment (Matulef 1999). Large proportions of families who were offered vouchers did not move during the period when the vouchers were valid, and analyses comparing movers to nonmovers in the experimental and comparison groups show that movers are significantly different from nonmovers. As a result, most (but not all) MTO analyses adopt analytic strategies that account for this selection. For example, Jens Ludwig and his colleagues (Ludwig, Duncan, and Pinkston 2000; Ludwig, Ladd, and Duncan 2001) in Baltimore and Lawrence Katz and his colleagues (Katz, Kling, and Liebman 2001) in Boston produce both intent-to-treat (ITT) and treatment-on-treated (TOT) estimates. ITT analyses compare outcomes for families assigned to the two treatment groups whether or not they actually moved with outcomes for the control group. Thus, ITT results are "lower bounds" on the effects of the treatment because the two treatment groups include substantial proportions of families who never moved. These researchers also estimate an "effects of TOT" parameter, which is a measure of the effect of moving on those who actually moved during the program. The TOT analysis uses instrumental variables methods to estimate the difference between families in the treatment groups

who moved with those in the control group who would have moved if offered the opportunity. In the New York study, Leventhal and Brooks-Gunn (2003) use a conceptually similar approach by comparing treatment group movers with both treatment group non-movers and those who were assigned to the control group. . . .

In general, most studies show some improvements in children's outcomes in the treatment groups compared with the controls. In particular, the Boston and Baltimore studies show significant differences in behavior problems, including juvenile arrests and respondent-reported behavior problems. This result is particularly striking since both studies report that children in the experimental group were more likely than those in the other groups to be arrested prior to their move. Experimental group children in New York, especially boys, experienced fewer depressive and anxiety-related behaviors.

Baltimore children in the treatment groups also had better test scores. Compared with the control group, children in the experimental group had better test scores overall, while those in the comparison group had better reading scores. There is some evidence in the Baltimore results that experimental-group children were more likely to be suspended from and drop out of school. Ludwig, Ladd, and Duncan (2001) suggest that middle-income schools are less likely to tolerate behavior that is acceptable in schools in poor neighborhoods. Health outcomes were better for children in the experimental group in Boston. The Boston and Chicago studies also report significant declines in fears about safety and increases in feelings of safety, a point of view echoed in the Los Angeles and New York studies.

An important concern of the New York and Los Angeles studies was the impact of moving into middle-class neighborhoods on poor children's social adjustment, social capital, and friendship patterns. If children move to better neighborhoods but feel left out or

are socially isolated, they may not be better off in the long run. In general, the results to date are reassuring. Children in all three groups were as likely to have a friend in the neighborhood. In some cases, children were less likely to participate in extracurricular activities in the experimental group. Maria Hanratty, Sara McLanahan, and Becky Pettit (1998) speculate that experimental group families may face more stringent financial situations because of higher rents and large security deposits compared with other groups.

In summary, the early results of the MTO experiments provide important new evidence that neighborhood social and physical conditions affect family life and at least some aspects of children's well-being. The results of these experimental studies are limited by implementation problems and unexpected events as well as by difficulties in generalizing to the rest of the population. Nonetheless, the results of experiments combined with those from observational studies will play an important role over the next several years in helping us understand the role of residential patterns in children's well-being.

Discussion

Despite the serious methodological problems that are only beginning to be addressed adequately, a review of previous experimental and observational studies suggests that growing up in a poor neighborhood negatively affects children's outcomes over and above the effects of family socioeconomic status. However, the effects may be complex and difficult to observe. For example, the MTO results suggest that a major effect of moving to a better neighborhood is feeling safer and less anxious and depressed. Although we might expect a greater sense of safety and lower anxiety and depression to have very important long-run effects on children's emotional development and outlook on life, the effects may be less immediately apparent on school performance, skills acquisition, and

behaviors, outcomes that are more typically measured in surveys and administrative data.

Research to date also suggests that family effects on children's outcomes are significantly larger than neighborhood effects. However, it is important to keep in mind that the measurement of neighborhood characteristics is at a much more rudimentary stage of development than measurement of family processes in large-scale surveys. Because of their pervasive role in most children's lives, it makes sense that families would have a greater influence on children's well-being than neighborhoods or other social environments. However, public policy generally has considerably less ability to influence parents' behavior and attributes directly than to affect neighborhood quality. Hence, even modest neighborhood effects may be of considerable interest to policymakers.

Moreover, it is important to consider residential segregation and neighborhood and family effects on children's well-being in a larger context. The finding that neighborhood effects are more modest in size than family effects can be misleading to the extent that neighborhood conditions, and residential segregation more generally, have an important influence on families' socioeconomic status and family dynamics. Residential segregation has been implicated by many scholars as a key mechanism for the intergenerational transmission of inequality (Massey and Denton 1993; Wilson 1987, 1996; Jarrogowsky 1997). The argument is that restriction to concentrated-poverty neighborhoods compounds the difficulty that poor, minority families face in escaping poverty because in poor neighborhoods housing values remain low, the chances of criminal victimization remain higher, high-paying jobs are less available, exposure to disease and substance abuse is greater, and individuals are more socially isolated. Thus, residential segregation and residence in concentrated-poverty neighborhoods may be an important determinant of

the family socioeconomic status and a major indirect influence on children's outcomes.

If residence in a poor and dangerous neighborhood affects parents' attitudes, mental health, and parenting practices (Furstenberg et al. 1999; Kling, Liebman, and Katz 2001; Brooks-Gunn, Duncan, and Aber 1997), it is even more difficult to disentangle "family" effects from "neighborhood" effects. Nonetheless, several recent experimental and observational studies promise to provide clearer answers than past research on the direct and indirect pathways through which residential segregation affects children growing up in poor neighborhoods.

REFERENCES

- Aaronson, Daniel. 1997. "Sibling Estimates of Neighborhood Effects." In *Neighborhood Poverty: Policy Implications in Studying Neighborhoods*, vol. 2, edited by Jeanne Brooks-Gunn, Greg J. Duncan, and J. Lawrence Aber. New York: Russell Sage Foundation.
- . 1998. "Using Sibling Data to Estimate the Impact of Neighborhoods on Children's Educational Outcomes." *Journal of Human Resources* 33(4): 915–46.
- Aber, J. Lawrence, Martha A. Gephart, Jeanne Brooks-Gunn, and James P. Connell. 1997. "Development in Context: Implications for Studying Neighborhood Effects." In *Neighborhood Poverty: Context and Consequences for Children*, vol. 1, edited by Jeanne Brooks-Gunn, Greg J. Duncan, and J. Lawrence Aber. New York: Russell Sage Foundation.
- Aneshensel, Carol S., and Clea A. Sucoff. 1996. "The Neighborhood Context of Adolescent Mental Health." *Journal of Health and Social Behavior* 37: 293–310.
- Baumer, Eric P., and Scott J. South. 2001. "Community Effects on Youth Sexual Activity." *Journal of Marriage and the Family* 63: 540–54.
- Brennan, Brian. 2002. "Background on MTO." *Moving To Opportunity Research*. Created August 30, 2000; last modified August 22, 2002. Available at: www.princeton.edu/~kling/mto/background.htm.
- Bronfenbrenner, Urie. 1986. "Ecology of the Family as Context for Human Development." *Developmental Psychology* 22(6): 723–42.

- Brooks-Gunn, Jeanne, Greg J. Duncan, and J. Lawrence Aber, eds. 1997. *Neighborhood Poverty*. New York: Russell Sage Foundation.
- Brooks-Gunn, Jeanne, Greg J. Duncan, Pamela K. Klebanov, and Naomi Sealand. 1993. "Do Neighborhoods Influence Child and Adolescent Development?" *American Journal of Sociology* 99(2): 353–95.
- Case, Anne C., and Lawrence F. Katz. 1991. "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youth." Working paper. Cambridge, Mass.: National Bureau of Economic Research.
- Coleman, James S. 1988. "Social Capital in the Creation of Human Capital." *American Sociological Review* 94 (supp.): S95–120.
- Coulton, Claudia J. 1996. "Effects of Neighborhoods on Families and Children: Implications for Services." In *Children and Their Families in Big Cities: Strategies for Service Reform*, edited by A. J. Kahn and S. B. Kamerman. New York: Columbia University.
- Crane, Jonathan. 1991. "The Epidemic Theory of Ghettos and Neighborhood Effects on Dropping Out and Teenage Childbearing." *American Sociological Review* 96(5): 1226–59.
- Duncan, Greg J., James P. Connell, and Pamela K. Klebanov. 1997. "Conceptual and Methodological Issues in Estimating Causal Effects of Neighborhoods and Family Conditions on Individual Development." In *Neighborhood Poverty: Context and Consequences for Children*, vol. 1, edited by Jeanne Brooks-Gunn, Greg J. Duncan, and J. Lawrence Aber. New York: Russell Sage Foundation.
- Duncan, Greg J., and Saul D. Hoffman. 1990. "Welfare Benefits, Economic Opportunities, and Out-of-Wedlock Births Among Black Teenage Girls." *Demography* 27(4): 519–35.
- Duncan, Greg J., and Stephen W. Raudenbush. 1999. "Assessing the Effects of Context in Studies of Child and Youth Development." *Educational Psychologist* 34(1): 29–41.
- _____. 2001. "Neighborhoods and Adolescent Development: How Can We Determine the Links?" In *Does It Take a Village? Community Effects on Children, Adolescents, and Families*, edited by Alan Booth and Ann C. Crouter. Mahwah, N.J.: Lawrence Erlbaum Associates.
- Edin, Kathryn. 1991. "Surviving the Welfare System: How AFDC Recipients Make Ends Meet in Chicago." *Social Problems* 38(4): 462–74.
- Evans, William N., Wallace E. Oates, and Robert M. Schwab. 1992. "Measuring Peer Group Effects: A Study of Teenage Behavior." *Journal of Political Economy* 100(3): 966–91.
- Fischer, Claude S. 1984. *The Urban Experience*. New York: Harcourt, Brace, Jovanovich.
- Fordham, Signithia, and John U. Ogbu. 1986. "Black Students' School Success: Coping with the Burden of 'Acting White.'" *The Urban Review* 18: 176–206.
- Foster, Eric M., and Sara McLanahan. 1996. "An Illustration of the Use of Instrumental Variables: Do Neighborhood Conditions Affect a Young Person's Chance of Finishing High School?" Unpublished paper. Princeton University.
- Furstenberg, Frank F., Jr., Thomas D. Cook, Jacquelynne Eccles, Glen H. Elder Jr., and Arnold J. Sameroff. 1999. *Managing to Make It: Urban Families and Adolescent Success*. Chicago: University of Chicago Press.
- Ginther, Donna, Robert Haveman, and Barbara Wolfe. 2000. "Neighborhood Attributes as Determinants of Children's Outcomes: How Robust Are the Relationships?" *Journal of Human Resources* 35(4): 603–42.
- Hanratty, Maria, Sara McLanahan, and Becky Pettit. 1998. "The Impact of the Los Angeles Moving To Opportunity Program on Residential Mobility, Neighborhood Characteristics, and Early Child and Parent Outcomes." Working paper 98–18. Princeton, N.J.: Princeton University, Center for Research on Child Wellbeing.
- Haveman, Robert, and Barbara Wolfe. 1994. *Succeeding Generations: On the Effects of Investments in Children*. New York: Russell Sage Foundation.
- Jargowsky, Paul A. 1997. *Poverty and Place: Ghettos, Barrios, and the American City*. New York: Russell Sage Foundation.
- Jencks, Christopher, and Susan E. Mayer. 1990. "The Social Consequences of Growing up in a Poor Neighborhood." In *Inner-City Poverty in the United States*, edited by Laurence E. Lynn Jr. and Michael G. H. McGeary. Washington, D.C.: National Academy Press.
- Katz, Lawrence F., Jeffrey R. Kling, and Jeffrey B. Liebman. 2001. "Moving To Opportunity in Boston: Early Results of a Randomized Mobility Experiment." *Quarterly Journal of Economics* (May): 607–54.
- Klebanov, Pamela K., Jeanne Brooks-Gunn, P. Lindsay Chase-Lansdale, and R. A. Gordon. 1997. "Are Neighborhood Effects on Young Children Mediated by Features of the Home Environment?" In *Neighborhood Poverty*, vol. 1, edited by Jeanne Brooks-Gunn, Greg J. Duncan,

- and J. Lawrence Aber. New York: Russell Sage Foundation.
- Klebanov, Pamela K., Jeanne Brooks-Gunn, and Greg J. Duncan. 1994. "Does Neighborhood and Family Poverty Affect Mother's Parenting, Mental Health and Social Support?" *Journal of Marriage and the Family* 56: 441–55.
- Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. 2001. "Bullets Don't Got No Name: Consequences of Fear in the Ghetto." Working paper 225. Chicago: Northwestern University and the University of Chicago, Joint Center for Poverty Research.
- Korbin, Jill E., and Claudia J. Coulton. 1997. "Understanding the Neighborhood Context for Children and Families: Combining Epidemiological and Ethnographic Approaches." In *Neighborhood Poverty*, vol. 2, edited by Jeanne Brooks-Gunn, Greg J. Duncan, and J. Lawrence Aber. New York: Russell Sage Foundation.
- Leventhal, Tama, and Jeanne Brooks-Gunn. 2000. "The Neighborhoods They Live In: The Effects of Neighborhood Residence on Child and Adolescent Outcomes." *Psychological Bulletin* 126(2): 309–37.
- _____. 2003. "The Early Impacts of Moving To Opportunity on Children and Youth in New York City." In *Choosing a Better Life: Evaluating the Moving To Opportunity Social Experiment*, edited by John Goering and Judith Feins. Washington, D.C.: Urban Institute Press.
- Ludwig, Jens, Greg J. Duncan, and Joshua C. Pinkston. 2000. "Evidence from a Randomized Housing-Mobility Experiment." Unpublished paper. Georgetown University, Washington, D.C.
- Ludwig, Jens, Helen F. Ladd, and Greg J. Duncan. 2001. "Urban Poverty and Educational Outcomes." *Brookings-Wharton Papers on Urban Affairs* 2001: 147–201.
- Massey, Douglas S., and Nancy A. Denton. 1993. *American Apartheid: Segregation and the Making of the Underclass*. Cambridge, Mass.: Harvard University Press.
- Massey, Douglas S., and Mitchell L. Eggers. 1990. "The Ecology of Inequality: Minorities and the Concentration of Poverty, 1970 to 1980." *American Journal of Sociology* 95(5): 1153–88.
- Massey, Douglas S., Andrew B. Gross, and M. L. Eggers. 1991. "Segregation, the Concentration of Poverty, and the Life Chances of Individuals." *Social Science Research* 20(4): 397–420.
- Massey, Douglas S., Andrew B. Gross, and K. Shibuya. 1994. "Migration, Segregation, and the Concentration of Poverty." *American Sociological Review* 59: 425–45.
- Matulef, Mark. 1999. "Moving To Opportunity (MTO) Demonstration for Fair Housing Program, Los Angeles Demonstration Site Interim Outcomes of Housing Search and Counseling Strategies: Early Lessons for Experimental Design and Implementation." Available at: www.wws.princeton.edu/~kling/mto/quick.htm.
- McLoyd, Vonnie C. 1990. "The Impact of Economic Hardship on Black Families and Children: Psychological Distress, Parenting, and Socioemotional Development." *Child Development* 61: 311–46.
- Morenoff, Jeffrey D., Robert J. Sampson, and Stephen W. Raudenbush. 2001. "Neighborhood Inequality, Collective Efficacy, and the Spatial Dynamics of Urban Violence." *Criminology* 39(3): 517–59.
- Oliver, Melvin L. 1988. "The Urban Black Community as Network: Toward a Social Network Perspective." *Sociological Quarterly* 29(4): 623–45.
- Rosenbaum, James E. 1991. "Black Pioneers: Do Their Moves to the Suburbs Increase Economic Opportunity for Mothers and Children?" *Housing Policy Debate* 2(4): 1179–1213.
- _____. 1995. "Changing the Geography of Opportunity by Expanding Residential Choice: Lessons from the Gautreaux Program." *Housing Policy Debate* 6(1): 231–70.
- Sampson, Robert J., and John H. Laub. 1994. "Urban Poverty and the Family Context of Delinquency: A New Look at Structure and Process in a Classic Study." *Child Development* 65: 523–40.
- Sampson, Robert J., Jeffrey D. Morenoff, and Felton Earls. 1999. "Beyond Social Capital: Spatial Dynamics of Collective Efficacy for Children." *American Sociological Review* 64(5): 633–60.
- Sampson, Robert J., Jeffrey D. Morenoff, and Thomas Gannon-Rowley. 2002. "Assessing 'Neighborhood Effects': Social Processes and New Directions in Research." *Annual Review of Sociology* 28: 443–78.
- Sampson, Robert J., and Stephen W. Raudenbush. 1999. "Systematic Social Observation of Public Spaces: A New Look at Disorder in Urban Neighborhoods." *American Journal of Sociology* 105(3): 603–51.
- Sampson, Robert J., Stephen W. Raudenbush, and Felton Earls. 1997. "Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy." *Science* 277 (August 15): 918–24.
- Shaw, Clifford R., and McKay, Henry D. 1969. *Juvenile Delinquency and Urban Areas*. Chicago: The University of Chicago Press.

- Solon, Gary, Marianne E. Page, and Greg J. Duncan. 2000. "Correlations Between Neighboring Children in Their Subsequent Educational Attainment." *Review of Economics and Statistics* 82(3): 383–92.
- South, Scott J. 2001. "Issues in the Analysis of Neighborhoods, Families and Children." In *Does It Take A Village? Community Effects on Children, Adolescents, and Families*, edited by Alan Booth and Ann C. Crouter. Mahwah, N.J.: Lawrence Erlbaum Associates, Inc.
- South, Scott J., and Eric P. Baumer. 2000. "Deciphering Community and Race Effects on Adolescent Premarital Childbearing." *Social Forces* 78: 1379–1407.
- Stack, Carol. 1974. *All Our Kin: Survival Strategies*. New York: Harper Torchback.
- Tigges, Leann M., Irene Browne, and Gary P. Green. 1998. "Social Isolation of the Urban Poor: Race, Class and Neighborhood Effects on Social Resources." *Sociological Quarterly* 39(1): 53–77.
- Upchurch, Dawn M., Carol S. Aneshensel, Clea A. Sucoff, and Lene Levy-Storms. 1999. "Neighborhood and Family Contexts of Adolescent Sexual Activity." *Journal of Marriage and the Family* 61: 920–33.
- Waldinger, Roger. 1996. "Ethnicity and Opportunity in the Plural City." In *Ethnic Los Angeles*, edited by Roger Waldinger and Mehdi Bozorgmehr. New York: Russell Sage Foundation.
- Wellman, Barry. 1996. "Are Personal Communities Local? A Dumptarian Reconsideration." *Social Networks* 17(2): 423–36.
- _____. 1999. "Preface." In *Networks in the Global Village*, edited by Barry Wellman. Boulder, Colo.: Westview Press.
- Wilson, William Julius. 1987. *The Truly Disadvantaged*. Chicago: University of Chicago Press.
- _____. 1996. *When Work Disappears: The World of the New Urban Poor*. New York: Alfred A. Knopf.

40 ■ Robert J. Sampson and Jeffrey D. Morenoff

Durable Inequality

Spatial Dynamics, Social Processes, and the Persistence of Poverty in Chicago Neighborhoods

Poverty can trap entire nations and social groups, not just individuals. The persistence of poverty among social groupings is perhaps the more intriguing and surprisingly understudied puzzle, especially in the case when it is not necessarily the same individuals that make up the groups over time (see also Tilly 1998). In this chapter we consider urban neighborhoods as one such social grouping and investigate the durable consequences of concentrated poverty. . . .

The Persistence of Poverty Regimes

Most urban analyses would agree that poverty concentration has increased in recent decades. For example, Wilson (1987) argues that the geographical concentration of low-income residents, especially of African Americans and female-headed families, stems from macroeconomic changes related to the deindustrialization of central cities and the outmigration of middle-class residents. Massey and Denton (1993) argue that the greater the

race/class segregation in a metropolitan area, the fewer the neighborhoods that can absorb economic shocks and the more severe the resulting concentration of poverty. In either case, economic stratification by race and residence fuels the neighborhood concentration of cumulative forms of disadvantage, intensifying the social isolation of low-income, minority, and single-parent residents from resources that could support collective social control (see also Jargowsky 1997).

Most research on the “Wilson-Massey” debate has attempted to assess whether the out-migration of the black middle class or racial segregation best explain the social transformation of urban communities. Of course it turns out that both mechanisms appear to have been operative, with the weight of recent evidence pointing to differential out-migration of the black nonpoor (Quillian 1999). In this chapter we do not attempt a causal analysis that adjudicates between these interpretations. Instead we step back and ask, just what has been transformed and by how much? The metaphor of social transformation implies radical change, a repositioning of neighborhoods in the ecological system. What do the data tell us?

We begin with some basic illustrations of how neighborhoods in Chicago have changed over the past three decades with respect to poverty and racial/ethnic composition. We conceptualize neighborhood as an ecological subsection of a larger community—a collection of both people and institutions occupying a spatially defined area that is conditioned by a set of ecological, socio-demographic, and often political forces. Operationally, we use the neighborhood boundaries that were constructed for the Project on Human Development in Chicago Neighborhoods (PHDCN), an ongoing study explicitly designed to examine social context. The PHDCN research team combined Chicago’s 847 populated census tracts to create 343 “Neighborhood Clusters” (NCs). The overriding consideration in

forming NCs was that they should be ecologically meaningful units composed of geographically contiguous census tracts and should be internally homogeneous with regard to distributions on a variety of census indicators. The study settled on an ecological unit smaller than the established 77 community areas in Chicago (average size = 40,000) but large enough to approximate local neighborhoods—on average, around 8,000 people. Major geographic boundaries (e.g., railroad tracks, parks, freeways), knowledge of Chicago’s local neighborhoods, and cluster analyses of census data were used to guide the construction of relatively homogeneous NCs with respect to distributions of racial-ethnic mix, SES, housing density, and family organization.

Our initial analysis is based on Census data from 1970 to 1990. Figure 1 tells a story of both stability and change in neighborhood poverty.¹ On the one hand, there is a very high correlation ($r = .87$) between neighborhood poverty rates in 1970 and 1990, as evidenced by the strong linear relationship on the scatterplot. Neighborhoods that were poor in 1970 generally continued to be poor in 1990. In fact, in supplemental analysis, we decomposed the total variance in neighborhood poverty rates over the decades 1970, 1980, and 1990 and found that most of this variance (67%) is due to differences *between* neighborhoods rather than differences over time within neighborhoods, implying that between-neighborhood differences in poverty were quite stable over time. On the other hand, there was also significant change during this time period, as the poverty rate for the average neighborhood increased from 11 percent in 1970 to 20 percent in 1990. This change was even more pronounced at the upper tail of the neighborhood poverty distribution. The 75th percentile of the distribution corresponded to a poverty rate of only 14 percent in 1970, but by 1990 it had more than doubled, increasing to 30 percent. In short, there was a dramatic growth in neighborhood poverty between

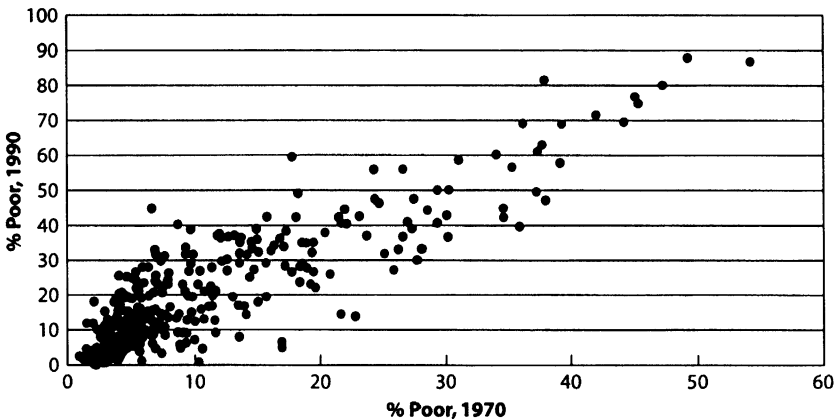


Figure 1. Persistence of Poverty, 1970–1990: Chicago Neighborhood Clusters

1970 and 1990, but despite this change, there was stability in the relative rank order of neighborhoods vis-à-vis poverty. Neighborhood poverty was both a persistent and an increasingly prevalent condition.

To bring the paths of neighborhood change into sharper relief, we analyze decade-to-decade change in neighborhood poverty in table 1, using the following six-category typology of neighborhood poverty: 0–4.9 percent, 5–9.9 percent, 10–19.9 percent, 20–29.9 percent, 30–39.9 percent, and 40 percent or above. Like mobility tables, the cross-tabulations displayed in table 1 depict inflows and outflows among poverty categories over time. The upper panel of table 1 displays patterns of neighborhood change from 1970 to 1980, and the lower panel covers the time period from 1980 to 1990. The trend toward more concentrated neighborhood poverty is reflected in the margins of both tables. Whereas in 1970 only 7 of the 342 NCs (2.0%) had poverty rates of 40 percent or above (a conventional threshold for defining “ghetto poverty” areas), this category expanded over time to include 29 NCs (8.5%) in 1980 and 42 (12.3%) in 1990. Thus, the number of high-poverty areas increased sixfold from 1970 to 1990. In percentage terms, the higher poverty categories grew more during the 1970s, but the growth in concentrated poverty continued during

the 1980s. Perhaps the most distinctive feature of neighborhood change over both decades was its asymmetrical nature: neighborhoods tended either to stay in the same poverty category (i.e., cells on the main diagonal) or move to a higher poverty category (i.e., cells to the right of the main diagonal) over time, but “upgrading” (i.e., cells to the left of the main diagonal) was quite rare.

Given this asymmetry, it is still important to distinguish between which neighborhoods remained relatively stable in terms of poverty and which got worse over time. We now turn our attention toward understanding change in poverty, rather than the level of poverty at any given period, as our main outcome of interest. Since the trends depicted in table 1 toward higher poverty were pervasive across both decades, we focus on the combined time period, 1970–1990. One important question about change in poverty is how it is related to “initial” conditions. In other words, having established that neighborhood poverty is a persistent condition, we now ask whether the neighborhoods that experienced the largest increases in poverty between 1970 and 1990 were those that were already very poor or those that were in transition toward becoming poor.

Figure 2 addresses this question by plotting the neighborhood poverty rate in 1970 against the change score (% poor in 1990—% poor

Table 1. Neighborhood “Mobility” in Poverty, 1970–1990 (Row Percentages in Parentheses)

| Neighborhood Poverty 1970 | Neighborhood Poverty 1980 | | | | | | Total |
|------------------------------|---------------------------|---------------|---------------|---------------|---------------|---------------|-----------------|
| | 0–4.9% | 5–9.9% | 10–19.9% | 20–29.9% | 30–39.9% | 40%+ | |
| 0–4.9% | 66 (58.4%) | 33 (29.2%) | 12 (10.6%) | 2 (1.8%) | 0 (0.0%) | 0 (0.0%) | 113 (100.0%) |
| 5–9.9% | 5 (4.6%) | 34 (31.2%) | 53 (48.6%) | 15 (13.8%) | 2 (1.8%) | 0 (0.0%) | 109 (100.0%) |
| 10–19.9% | 0 (0.0%) | 5 (7.1%) | 14 (20.0%) | 34 (48.6%) | 15 (21.4%) | 2 (2.9%) | 70 (100.0%) |
| 20–29.9% | 0 (0.0%) | 0 (0.0%) | 3 (11.5%) | 1 (3.8%) | 15 (57.7%) | 7 (26.9%) | 26 (100.0%) |
| 30–39.9% | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 4 (23.5%) | 13 (76.5%) | 17 (100.0%) |
| 40%+ | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 7 (100.0%) | 7 (100.0%) |
| Total | 71 (20.8%) | 72 (21.1%) | 82 (24.0%) | 52 (15.2%) | 36 (10.5%) | 29 (8.5%) | 342 (100.0%) |

| Neighborhood Poverty 1980 | Neighborhood Poverty 1990 | | | | | | Total |
|------------------------------|---------------------------|---------------|---------------|---------------|---------------|---------------|-----------------|
| | 0–4.9% | 5–9.9% | 10–19.9% | 20–29.9% | 30–39.9% | 40%+ | |
| 0–4.9% | 56 (78.9%) | 10 (14.1%) | 5 (7.0%) | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 71 (100.0%) |
| 5–9.9% | 8 (11.1%) | 29 (40.3%) | 32 (44.4%) | 3 (4.2%) | 0 (0.0%) | 0 (0.0%) | 72 (100.0%) |
| 10–19.9% | 2 (2.4%) | 11 (13.4%) | 45 (54.9%) | 20 (24.4%) | 2 (2.4%) | 2 (2.4%) | 82 (100.0%) |
| 20–29.9% | 0 (0.0%) | 0 (0.0%) | 7 (13.5%) | 24 (46.2%) | 19 (36.5%) | 2 (3.8%) | 52 (100.0%) |
| 30–39.9% | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 5 (13.9%) | 18 (50.0%) | 13 (36.1%) | 36 (100.0%) |
| 40%+ | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 4 (13.8%) | 25 (86.2%) | 29 (100.0%) |
| Total | 66 (19.3%) | 50 (14.6%) | 89 (26.0%) | 52 (15.2%) | 43 (12.6%) | 42 (12.3%) | 342 (100.0%) |

in 1970). The graph shows that there is a positive and moderately strong correlation ($r = .52$) between the level of poverty in 1970 and change between 1970 and 1990. That is to say, poverty increased the most among those neighborhoods that were already poor in 1970. In this regard, neighborhood poverty appears to represent a spiraling trap. For example, neighborhoods that were already 40 percent poor in 1970 experienced an average poverty increase on the order of 31.6

percentage points between 1970 and 1990. Figure 2 also shows that the relationship between initial conditions and change was not uniform across all neighborhoods. Instead, it is conditioned by a third factor, racial composition. The association was strongest among neighborhoods that were at least 75 percent black in 1970 ($r = .61$) but weaker among neighborhoods that were at least 75 percent white ($r = .39$), and non-significant among other “mixed” neighbor-

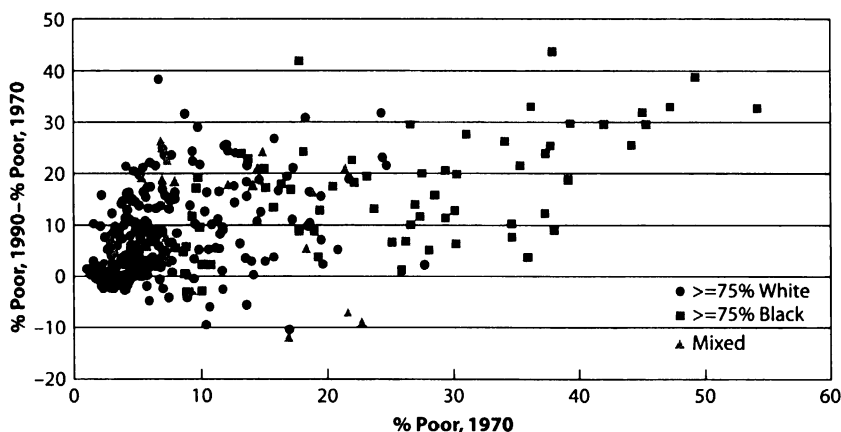


Figure 2. Initial Condition and Change in Poverty, 1970–1990: Chicago Neighborhood Clusters

hoods ($r = -.08$).² Thus, the trap of neighborhood poverty was particularly nettlesome for poor black neighborhoods.

To further elaborate the backdrop of racial segregation against which the changes in poverty occurred, we plot the percentage of a neighborhood's population that was black in 1970 against changes in percent black from 1970 to 1990 (% black 1990 - % black 1970) in figure 3. This graph shows that racial change is even more asymmetric than poverty change—the amazing lack of observations in the lower-right quadrant of the graph indicates that none of the NCs that had large percentages of black population in 1970 lost shares over time. In fact, there appears to be a clear threshold effect of around

40 percent black, above which all neighborhoods either maintained or increased their share of black population. As was the case with poverty, figure 3 tells a story of change within a stable ecological system: there were great shifts in neighborhood racial composition from 1970 to 1990, but neighborhoods that were initially black stayed that way over time, while at the same time many areas of the city remained off limits to blacks. In light of this pattern, it comes as no surprise that as of 1990 Chicago had the distinction of being one of the most racially segregated cities in America (Massey and Denton 1993).³

Another key to understanding the paths of neighborhood change with regard to poverty is geography. Here, the motivating question

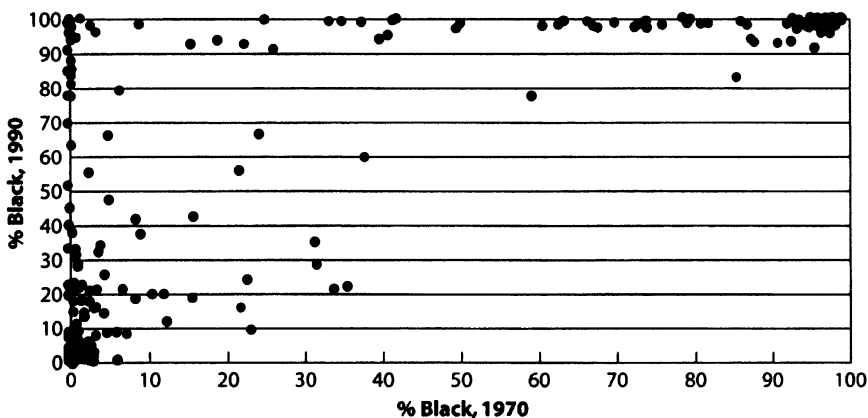


Figure 3. Asymmetry of Racial Change, 1970–1990: Chicago Neighborhood Clusters

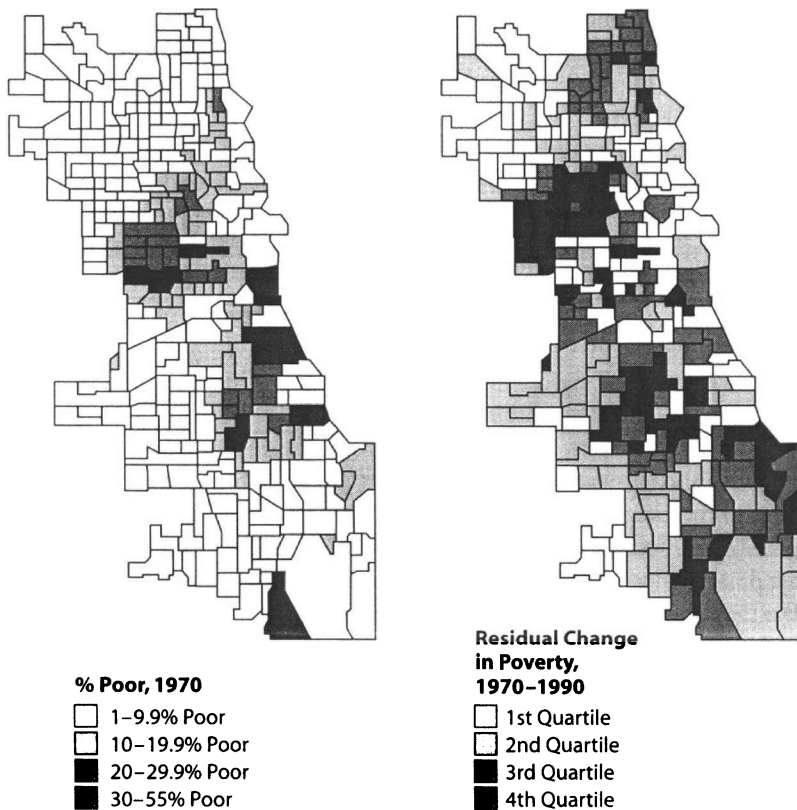


Figure 4. The Geography of Neighborhood Poverty (1970), Poverty Change (1970–1990), and Racial Composition (1970): Chicago Neighborhood Clusters

is whether changes in neighborhood poverty from 1970 to 1990 occurred in a spatially random fashion or whether they were concentrated within certain geographic areas of the city. If location in space didn't matter, but only a neighborhood cluster's own history of poverty did, then a good predictor of a neighborhood's 1990 poverty rate would be its 1970 poverty rate and variations in poverty from the predicted rate would not have any spatial pattern. To test this hypothesis, we first estimate the equation:

$$\text{Poverty}_{90} = \alpha + \beta * \text{Poverty}_{70} + \epsilon.$$

The expression $\alpha + \beta * \text{Poverty}_{70}$ is the 1990 *predicted poverty rate* and ϵ is the deviation or *residual* from the predicted rate. The residual reflects the amount of change in a given neighborhood's poverty that is unac-

counted for or unexplained by its initial level of poverty.⁴ We plot the residual for each neighborhood cluster in the map in the right panel of figure 4, and in the left panel the map displays the spatial distribution of the level of poverty in 1970.

Neighborhoods where the population was at least 75 percent black in 1970 were also very poor (shaded in dark gray). Initial concentrations of poverty—displayed in the 1970 map—were geographically clustered at the core of Chicago's black belt, in community areas such as Grand Boulevard, Oakland, Woodlawn, and Englewood to the south; and Garfield Park and North Lawndale along the western corridor. The second map displays residual change in poverty from 1970 to 1990. The change map shows that between 1970 and 1990 relative increases in poverty were found, not at the core of Chicago's black

ghetto, but in neighborhoods along the periphery, which pushed further outward with each succeeding decade (see Morenoff and Sampson 1997). Side-by-side, these two maps fit together like a jigsaw puzzle, revealing that most of the “unexpected” change in poverty was occurring in close geographic proximity to high initial rates of poverty.

Explaining Poverty Trajectories

Our next set of analyses tackles the multidimensional predictors of change in neighborhood poverty. In so doing, we pay special attention to two factors highlighted in our descriptive analyses: *initial conditions* and the *spatial dynamics* of poverty. Our dependent variables are the two types of change scores described above: the “raw” change score used in figure 2 (% poor 1990 – % poor 1970) and residual change score used in figure 4. In modeling raw change, we use the 1970 poverty rate as a predictor to examine the extent to which neighborhood poverty is a path-dependent process. Based on prior research we include a core set of other predictors in this model: 1970 Census measures of percent black, percent Hispanic, percent owner-occupied homes, as well as raw change between 1970 and 1990 in percent black, percent Hispanic, and percent owner-occupied. The residual change score already filters out the effect of path dependence, leaving us to model the “unexpected” change in poverty between 1970 and 1990. We include the same control variables, only here all 1970–1990 change measures are expressed as residual change scores. Thus, our statistical models allow us to assess the robustness of the overall process by predicting either raw change or residual change as a function of initial level and change in racial/ethnic composition and owner occupancy, controlling for initial poverty status.

Building on the exploratory analysis presented in figure 4, we also address the notion of *spatial dependence* in our statistical models. In this case, spatial dependence

refers to the idea that change in a given neighborhood’s poverty rate is conditional on poverty changes in surrounding neighborhoods. Such dependence could come about if changes in poverty are the result of ecological dynamics that extend beyond the boundaries of a given neighborhood, such as the distribution of institutional resources that influence housing decisions—such as the quality of local schools—and other factors that trigger patterns of residential mobility, such as violent crime (Morenoff and Sampson 1997). The important theoretical point is that neighborhoods should not be treated as islands unto themselves, as they are in conventional regression analyses, because it is likely that factors affecting poverty change in one place also affect changes in nearby places, much like a ripple or diffusion process.

We estimate spatial dependence by constructing “spatially lagged” versions of change in poverty. We define y_i as the change in poverty from 1970 to 1990 (either raw or residual change) in NC_i , and w_{ij} as element i,j of a spatial weights matrix that expresses the geographical proximity of NC_i to NC_j (Anselin 1988, 11). For a given observation i , a spatial lag, defined as $\sum_j w_{ij} y_j$, is the weighted average of poverty change in neighboring locations. The weights matrix is expressed as first-order contiguity, which defines neighbors as those NCs that share a common border. Thus, $w_{ij} = 1$ if i and j are contiguous, 0 if not. We then test formally for the independent role of spatial dependence in a multivariate model by introducing the spatial lag as an explanatory variable. The spatial lag regression model is defined as

$$y = \rho W y + X \beta + \epsilon, \quad (1)$$

where y is an N by 1 vector of observations on the dependent variable; $W y$ is an N by 1 vector composed of elements $\sum_j w_{ij} y_j$, the spatial lags for the dependent variable; ρ is the spatial autoregressive coefficient; X is an N by

K matrix of exogenous explanatory variables with an associated K -by-1 vector of regression coefficients β ; and ϵ is an N -by-1 vector of normally distributed random error terms, with means 0 and constant (homoskedastic) variances.

A common but misleading interpretation of ρ is that for a given neighborhood, i , it simply represents the effect of a one-unit change in the average rate of change of i 's first-order neighbors on the poverty change rate of i . This interpretation would seem to suggest a diffusion process, whereby a high rate of "ghetto formation" in one neighborhood diffuses outward and affects the growth in poverty in surrounding neighborhoods. However, the interpretation of ρ as a pure diffusion (or feedback) mechanism—the effect of a one-unit change in Wy on y —is incorrect and does not capture the complexity of the spatial process specified in equation 1. By extending the logic of equation 1, we can demonstrate that the spatial lag model also incorporates the values of the measured X variables and the ϵ term (i.e., unmeasured characteristics) in spatially proximate neighborhoods. According to equation 1, the value of y at location i depends on the values of X and ϵ at location i and on values of y in i 's first-order neighbors. In turn, the first-order neighbors' values of y are functions of X and ϵ in i 's first-order neighbors and y in i 's second-order neighbors, and so on. This process continues in a steplike fashion, incorporating the neighborhood characteristics of successively higher-order neighbors of i (see also Tolnay et al. 1996), so that a change in X or ϵ at location i influences not only the value of y at location i but also (indirectly) at all other locations in Chicago. This spatial process can be expressed mathematically by rewriting equation 1 as follows:

$$(2) \quad y = X\beta + \rho WX\beta + \rho^2 W^2 X\beta + \dots \\ + \rho^m W^m X\beta + \epsilon + \rho W\epsilon + \rho^2 W^2 \epsilon + \dots \\ + \rho^m W^m \epsilon,$$

where $m \rightarrow \infty$. Equation 2 is also known as the "spatial multiplier" process, because it shows that the spatial regression model treats spatial dependence as a sort of ripple effect.⁵ The ρ coefficient from the spatial lag model thus captures spatial "exposure" to the observed (X) variables, spatial exposure to unobserved predictors, and endogenous feedback in poverty change. Similar to the inclusion of endogenous explanatory variables in systems of simultaneous equations, we estimated the spatial lag models with maximum likelihood (Anselin 1995).

The results for the spatial lag models of raw change and residual change, displayed in table 2, show that consistent findings emerge across the two specifications of change. First, poverty change appears to be both a path-dependent and spatially dependent process. The path dependence is reflected in the significant coefficient for percent poor in 1970 on raw change in poverty. The spatial dependence coefficient is also highly significant in both models, suggesting poverty change in one place is conditioned by similar changes in nearby places. Both models also imply that neighborhoods that experienced greater increases in percent black between 1970 and 1990 experienced significant growth in poverty change over that time period, but prior initial concentrations of black population were associated only with raw poverty change, not residual change. This finding dovetails with the residual change map in figure 4, which showed that most poverty change occurred around rather than in Chicago's traditional black belt. The models also suggest that owner occupancy acts as a buffer against poverty—higher prior levels and change in owner occupancy were both associated with significant declines in poverty over time. The initial distribution of the Hispanic population was not a significant predictor of poverty change in either model, but increases in Hispanic composition over time were associated with growing poverty rates. This implies that the substan-

Table 2. Maximum Likelihood Spatial Regression Models of Change in Poverty, 1970–1990

| <i>Statistic</i> | <i>Raw Change in %Poor 1970–1990</i> | | <i>Residual Change in %Poor 1970–1990</i> | |
|--------------------------------------|--|------------------|---|------------------|
| | <i>Coeff.</i> | <i>Std. Err.</i> | <i>Coeff.</i> | <i>Std. Err.</i> |
| %Poor 1970 | 0.25 | (0.06)** | | |
| Spatial Dependence | 0.27 | (0.06)** | 0.31 | (0.06)** |
| %Black 1970 | 0.04 | (0.01)** | 0.01 | (0.01) |
| Change (raw/residual) %Black 1970–90 | 0.14 | (0.02)** | 0.14 | (0.02)** |
| %Hispanic 1970 | -0.01 | (0.03) | 0.03 | (0.03) |
| Change (raw/residual) %Hisp 1970–90 | 0.11 | (0.02)** | 0.12 | (0.02)** |
| %Owner Occupied 1970 | -0.12 | (0.02)** | -0.06 | (0.01)** |
| Change %Own Occ 1970–90 | -0.34 | (0.06)** | -0.26 | (0.05)** |
| Intercept | 6.03 | (1.56)** | 2.11 | (0.89)* |
| <i>R</i> ² | 0.62 | | 0.47 | |

* $p < .05$; ** $p < .01$

tial inflow of Hispanic migrants to Chicago from 1970 to 1990 (most of it occurring from Mexican migration) does explain part of the increase in neighborhood poverty over time.

Consequences of Inequality

Now that we have a reasonable sense of temporal and spatial dynamics, the third part of our analysis tackles the legacy of durable inequality for key social processes in urban neighborhoods. This analysis speaks most directly to the idea of neighborhood poverty traps. We begin with a focus on how poverty undermines mechanisms of social organization in modern cities that may be facilitated by, but do not require, strong (or dense) ties and associations.

Erosion of Collective Efficacy

As Warren (1975, 50) noted some time back, the common belief that neighborhoods have declined in importance as social units “is predicated on the assumption that the neighborhood is exclusively a primary group and therefore should possess the ‘face-to-face,’ intimate, affective relations which characterize all primary groups.” Rejecting this outmoded assumption about the function of local communities in mass society, Sampson

et al. (1997) highlighted the combination of a working trust and shared willingness of residents to intervene in situations that bear on social order and the sustainability of communities. This linkage of trust and cohesion with shared expectations for control was defined as “collective efficacy.” We argued that just as self-efficacy is situated rather than global (one has self-efficacy relative to a particular task), a neighborhood’s efficacy exists relative to specific tasks such as maintaining public order. . . .

We suggest that collective efficacy, like social capital, is endogenous to structural and cultural contexts (Bourdieu 1986). Sampson et al. (1997) argue that extreme resource deprivation and racial exclusion act as a centrifugal force that hinders collective efficacy. Even when personal ties are strong in areas of concentrated disadvantage, daily experiences with uncertainty, danger, and economic dependency are likely to reduce expectations for taking effective collective action (Woolcock 1998, 207). Wilson’s (1987) “socially isolated” areas, for example, are thought to be characterized by dense personal ties that are nonetheless disconnected from the capacity to capture resources from the larger society. Lack of collective capacity in turn renders a neighborhood vulnerable to further decay and a relative lack of desirability in the

pecking order of places to live. Therefore, we hypothesize that if concentrated poverty serves as a trap, it does so partly through a “vicious circle” whereby collective efficacy is undermined, itself setting in motion a cascading set of disadvantages (outmigration, violence) that contribute to deepening and reinforcing poverty. We are not able to model these feedback processes, but we can take a first step by looking at the predictive effects of initial conditions and change in poverty on later collective efficacy. If these prove consequential, there is reason to explore more deeply the connection of declines in collective efficacy with reciprocal declines in the social position of neighborhoods and ultimately the stratification of places.⁶

“Anomie” and Cynicism

Our focus on another outcome draws inspiration from the classic Durkheimian notion of anomie—a state where the rules of the dominant society, and hence the legal system, are no longer perceived to be binding in a community or for a population subgroup (Kapsis 1978, 1139). Anomie in this sense is conceived as part of a social system and not merely a property of the individual, breeding a sense of normlessness, perceived lack of control, and ultimately moral cynicism about the rules of the society and their application, regardless of individual values. A key ingredient of cynicism is lack of trust in the larger society. While conceptually different, there is thus an analogy between the emphasis in collective efficacy on active control and trust in neighbors with the idea that underlying cynicism is a lack of trust in societal institutions.

There are strong theoretical reasons to expect that moral and legal cynicism are consequences of deprivation. Bullough (1967) and Kapsis (1978) suggest that members of economically and racially isolated communities, especially those least able to exercise political influence to obtain community services, are more likely than others to report alienation

and powerlessness. Sampson and Bartusch (1998) argue that concentrated poverty fosters a sense of cynicism and detachment from notions about self- and societal control. There is also speculation in the literature on poverty traps that attenuation in perceived locus of control might be one of the causal links that serves to reinforce poverty (Bowles 2000). It follows that pronounced cynicism and anomie set in motion a self-fulfilling prophecy of sorts; to the extent that communities maintain a reputation or actual conduct that sets them apart from striving for societal success in traditional ways, their economic status will be further undermined. . . .

In short, we propose that subcultures of moral and legal cynicism and the attenuation of social efficacy are not necessarily (or only) about individual-level variations, especially those by race/ethnicity. Individual-level variations still matter, on the other hand, a point Fischer (1995, 548–49) carefully acknowledges in a call for contextually based research on subcultural orientations. We therefore integrate these analytical perspectives by simultaneously examining individual-level variations in moral/legal cynicism and collective efficacy in conjunction with an analysis of whether structural characteristics explain their neighborhood-level variance—above and beyond the socio-demographic characteristics of the people residing in those neighborhoods. In this way we partition the variance in efficacy and cynicism into between-individual and between-neighborhood components, with our main theoretical interest residing in the latter.

Survey Data and Methodology

We use the Community Survey of the Project on Human Development in Chicago Neighborhoods (PHDCN), a multidimensional assessment by residents of the structural and cultural properties of their neighborhoods in 1995. . . .

Modifying Srole's (1956) anomie scale, *cynicism* is measured by five items assessing beliefs about the legitimacy of law and social norms (see also Sampson and Bartusch 1998). Respondents reported their level of agreement with five statements: "Laws were made to be broken"; "It's okay to do anything you want as long as you don't hurt anyone"; "To make money, there are no right and wrong ways anymore, only easy ways and hard ways"; "Fighting between friends or within families is nobody else's business"; and "Nowadays a person has to live pretty much for today and let tomorrow take care of itself" (1 = strongly disagree, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, and 5 = strongly agree). The common idea is the sense in which laws or societal rules are not trusted or considered binding in the present lives of respondents. Taken as a whole, that is, the items tap variation in respondents' ratification of acting in ways that are fatalistic and "outside" of common-law understanding.

To assess *collective efficacy*, we replicate Sampson et al. (1997) and combine two related scales. The first is a five-item scale tapping shared expectations for social control. Residents were asked about the likelihood that their neighbors could be counted on to take action if: children were skipping school and hanging out on a street corner, children were spraypainting graffiti on a local building, children were showing disrespect to an adult, a fight broke out in front of their house, and the fire station closest to home was threatened with budget cuts. Social cohesion/trust was measured by asking respondents how strongly they agreed that "People around here are willing to help their neighbors"; "This is a close-knit neighborhood"; "People in this neighborhood can be trusted"; "People in this neighborhood generally don't get along with each other" (reverse coded); and "People in this neighborhood do not share the same values" (reverse coded). Social cohesion and informal social

control were strongly related across neighborhood clusters ($r = .80$), and, following Sampson et al. (1997), were combined into a summary measure of the higher-order construct, "collective efficacy." The aggregate-level or "ecometric" reliability (see Raudenbush and Sampson 1999) of collective efficacy is .85.

The nested structure of the PHDCN data is addressed by adapting hierarchical linear models (HLM) that account for the nonindependence of observations within neighborhood contexts (Bryk and Raudenbush 1992). We estimate two equations simultaneously: within-neighborhood and between-neighborhood.

Our analysis strategy accounts for a rich array of individual and neighborhood covariates. The within-neighborhood model regresses our measures of cynicism and collective efficacy on ten person-level characteristics: *race/ethnicity* (composed of indicators for Latino-American, non-Latino African American, and "other"; non-Latino Caucasian is the reference category); a composite measure of *socioeconomic status* (first principal component of education, income, and occupational prestige); *sex* (1 = female, 0 = male), *marital status* (composed of separate indicators for married and separated or divorced); *home ownership*, *mobility* (number of moves in the past five years); and *age*. Using cynicism as the example, the within-neighborhood model can be written as:

$$(\text{Cynicism})_{ij} = \beta_{0j} + \sum_{q=1-10} \beta_q X_{qij} + e_{ij},$$

where β_{0j} is the intercept; X_{qij} is the value of covariate q associated with respondent i in neighborhood j ; and β_q is the partial effect of that covariate on cynicism. The error term, e_{ij} , is the unique contribution of each individual, which is assumed to be independently, normally distributed with constant

variance σ^2 . Importantly, because the person-level covariates are centered about the sample means, β_{0j} is the mean cynicism in a neighborhood after the effects of the ten person-level covariates have been adjusted.

The core between-neighborhood model can be written as:

$$\begin{aligned} \beta_{0j} = & \theta_{00} + \theta_{01} \text{ (1970 Poverty)} \\ & + \theta_{02} \text{ (\Delta in Poverty 70–90)} \\ & + \theta_{03} \text{ (1970 Black)} \\ & + \theta_{04} \text{ (\Delta in Black 70–90)} \\ & + \theta_{05} \text{ (1970 Latino)} \\ & + \theta_{06} \text{ (\Delta in Latino 70–90)} \\ & + \theta_{07} \text{ (1970 Home-ownership)} \\ & + \theta_{08} \text{ (\Delta in Home-ownership 70–90)} \\ & + \theta_{09} \text{ (1990 Density)} \\ & + \theta_{10} \text{ (1990 Log Population Size)} + U_{0j} \end{aligned}$$

where θ_{00} is the average cynicism score, and θ_{01} to θ_{10} are the regression coefficients of the estimated effects of the structural predictors (1970 and change scores) on the adjusted neighborhood level of cynicism. We also include population density and size in 1990 based on previous research linking them to 1995 social processes in the Chicago data (Sampson et al. 1999). U_{0j} is the neighborhood-level error term, or the unique contribution of each neighborhood, assumed to be normally distributed with variance τ . Based on preliminary analysis, we constrain all within-neighborhood slopes to be constant across neighborhoods. Our interest is in the main effects of poverty on variance across neighborhoods in legal cynicism and collective efficacy, adjusting for individual-level differences in socio-demographic composition.

Results

Table 3 presents the coefficients and standard errors of the HLM decomposition of variance for both cynicism and collective efficacy. Model 1 presents just the person-level predictors to reveal differential patterns in

the data as we integrate levels of analysis. Note that the coefficient for African Americans is both positive and significant ($p < .01$), meaning that blacks report higher levels of cynicism and lower levels of collective efficacy than do whites. The other pattern is that higher-SES respondents, females, older respondents, and those either married or separated/divorced report lower levels of estrangement from societal norms. For collective efficacy, on the other hand, sex and marital status are not important, while home-ownership and residential mobility emerge as strong predictors.

Model 2 introduces neighborhood context into the picture. Race/ethnic composition, home ownership, density, and population size all fail to predict neighborhood variance in cynicism. By contrast, concentrated poverty in 1970 and increases in poverty from 1970 to 1990 emerge as independent and significant predictors of both cynicism and collective efficacy in 1990.⁷ The greater the initial level of concentrated poverty, and the greater the unexpected increases in concentrated poverty, the more cynical and collectively inefficacious the neighborhood turns out to be, adjusting for the individual characteristics of the survey respondents (see also Cohen and Dawson 1993). Moreover, once neighborhood-level differences are accounted for, the coefficient for blacks is reduced to insignificance. Other person-level predictors do not change. For example, the coefficients for SES, marital status, and sex remain invariant, whereas the coefficient for African Americans is cut by almost 50 percent.

A similar pattern obtains for collective efficacy, although here the reduction in the coefficient for blacks is even greater in scope—it declines by approximately 70 percent when neighborhood characteristics are added to the model. The coefficient for “other” race drops as well for both outcomes, although it remains marginally significant

Table 3. Individual- and Neighborhood-Level Predictors of Cynicism and Collective Efficacy: PHDCN, 1995

| | <i>Cynicism</i> | | | | | | <i>Collective Efficacy</i> | | | | | |
|---------------------------------|-----------------|-----------|--------|-----------|--------|-----------|----------------------------|-----------|--------|-----------|--------|-----------|
| | (1) | | (2) | | (3) | | (1) | | (2) | | (3) | |
| | Coeff. | Std. Err. | Coeff. | Std. Err. | Coeff. | Std. Err. | Coeff. | Std. Err. | Coeff. | Std. Err. | Coeff. | Std. Err. |
| Intercept | 2.36 | (0.01)** | 2.32 | (0.21)** | 2.30 | (0.21)** | 0.02 | (0.04) | -0.03 | (0.71) | -0.24 | (0.71) |
| Individual-Level | | | | | | | | | | | | |
| Female | -0.07 | (0.02)** | -0.07 | (0.02)** | -0.07 | (0.02)** | -0.03 | (0.04) | -0.03 | (0.04) | -0.03 | (0.04) |
| Married | -0.11 | (0.02)** | -0.11 | (0.02)** | -0.11 | (0.02)** | 0.05 | (0.04) | 0.03 | (0.04) | 0.03 | (0.04) |
| Separated/Divorced | -0.07 | (0.02)** | -0.07 | (0.02)** | -0.07 | (0.02)** | -0.06 | (0.06) | -0.06 | (0.05) | -0.06 | (0.05) |
| Own Home | -0.01 | (0.02) | -0.01 | (0.02) | -0.01 | (0.02) | 0.41 | (0.05)** | 0.33 | (0.05)** | 0.33 | (0.05)** |
| Black | 0.09 | (0.02)** | 0.05 | (0.03) | 0.05 | (0.03) | -0.31 | (0.08)** | -0.09 | (0.09) | -0.08 | (0.09) |
| Hispanic | 0.04 | (0.02) | 0.03 | (0.03) | 0.02 | (0.03) | -0.03 | (0.07) | 0.10 | (0.07) | 0.11 | (0.07) |
| Mobility | 0.01 | (0.01) | 0.01 | (0.01) | 0.01 | (0.01) | -0.06 | (0.02)** | -0.06 | (0.02)** | -0.06 | (0.02)** |
| Age | 0.00 | (0.00)** | 0.00 | (0.00)** | 0.00 | (0.00)** | 0.01 | (0.00)** | 0.01 | (0.00)** | 0.01 | (0.00)** |
| SES | -0.10 | (0.01)** | -0.10 | (0.01)** | -0.10 | (0.01)** | 0.11 | (0.02)** | 0.09 | (0.02)** | 0.09 | (0.02)** |
| Other Race | 0.09 | (0.03)** | 0.07 | (0.03)* | 0.07 | (0.03)* | -0.24 | (0.08)** | -0.11 | (0.08) | -0.11 | (0.08) |
| Neighborhood Level | | | | | | | | | | | | |
| Population Density 1990 | | | 0.00 | (0.00) | 0.00 | (0.00) | | | 0.00 | (0.00)** | 0.00 | (0.00)** |
| %Owner Occupied 1970 | | | 0.09 | (0.06) | 0.09 | (0.06) | | | 0.39 | (0.18)* | 0.28 | (0.18) |
| %Poor 1970 | | | 0.42 | (0.19)* | 0.40 | (0.19)* | | | -1.44 | (0.54)** | -1.07 | (0.54)* |
| %Black 1970 | | | 0.02 | (0.05) | 0.02 | (0.05) | | | -0.37 | (0.14)** | -0.16 | (0.16) |
| %Hispanic 1970 | | | 0.01 | (0.08) | 0.00 | (0.09) | | | -0.51 | (0.24)* | -0.17 | (0.26) |
| Resid Change %Poor 1970-90 | | | 0.34 | (0.15)* | 0.32 | (0.14)* | | | -2.47 | (0.46)** | -2.22 | (0.44)** |
| Resid Change %Black 1970-90 | | | -0.04 | (0.07) | -0.06 | (0.08) | | | -0.54 | (0.22)* | -0.31 | (0.24) |
| Resid Change %Own Occ 1970-90 | | | -0.06 | (0.16) | -0.05 | (0.16) | | | 1.12 | (0.49)* | 1.24 | (0.48)** |
| Resid Change %Hisp 1970-90 | | | 0.06 | (0.07) | 0.05 | (0.07) | | | -0.60 | (0.18)** | -0.44 | (0.19)* |
| LN Total Pop 1970 | | | -0.37 | (2.33) | -0.29 | (2.40) | | | 5.20 | (7.79) | 10.38 | (7.77) |
| LN Homicide Rate 1970 | | | | | -0.39 | (0.82) | | | | | -4.08 | (2.51) |
| LN Homicide Rate 1990 | | | | | 0.76 | (0.86) | | | | | -6.22 | (2.75)* |
| %Var explained at individ level | 4.48 | | 4.56 | | 4.55 | | 2.19 | | 2.58 | | 2.59 | |
| %Var explained at neighb level | 57.36 | | 58.81 | | 58.49 | | 42.08 | | 74.95 | | 75.97 | |

Note: All neighborhood-level coefficients and standard errors have been multiplied by 100.

* $p < .05$; ** $p < .01$

for cynicism. It seems, then, that African American status is confounded with neighborhood context—blacks appear more cynical and perceive lower efficacy because they are disproportionately likely to live in residential environments of concentrated and increasing poverty. As we have already shown, the magnitude of difference in residential niches by race/ethnicity is striking—African Americans, relative to both whites and Latinos, are much more likely to reside in ecologically distinct environments of poverty. The data suggest that it is precisely this contextual reality of ecologically structured disadvantage—and not race at the individual level—that is the driving component of the legal cynicism and collective efficacy results.

Another interesting fact in table 3 is that collective efficacy is shaped more by initial structural context and ensuing dynamics than is cynicism. More of the neighborhood variance is explained in collective efficacy compared to cynicism (75% and 59%, respectively), and a larger number of structural-dynamic factors predict collective efficacy. For example, areas characterized by high levels of minorities (Latino and black) in 1970 and residual changes in minorities from 1970 to 1990 experience lower levels of collective efficacy in 1995, regardless of individual-level race/ethnicity. It may be that segregated minority areas are stigmatized in ways that lead to reduced trust and expectations for social control. Another factor that stands out is one often overlooked in the explanation of poverty traps—home ownership (see Hoff and Sen 2000; Galster 1987). In our data, both the initial level and residual change in home ownership are clear contextual predictors of increased collective efficacy.

In any case, the main message from the results in table 3 that we would like to emphasize is the consistent predictive power of initial poverty and changes in poverty for

explaining later social-process outcomes. This pattern holds up even when we introduce measures of neighborhood violent crime (Model 3). Because we know that concentrated poverty is linked to violent crime (Sampson et al. 1997), it is important to explicitly entertain this rival hypothesis. For example, it may be that experience with personal victimization or the perception of rampant crime in the neighborhood breeds hopelessness and cynicism about social norms of responsibility and action. We construct measures of violent crime by taking the log of the homicide rate in 1970 and 1990. Although concentrated poverty is rather highly correlated with the homicide rate ($r = .65$ in 1970 and $.64$ in 1990, $p < .01$), the results are invariant to the latter's consideration. This test reveals that lagged concentrated poverty and recent increase in poverty continue to be robust predictors of legal cynicism and collective efficacy, adjusting not only compositional differences but also rates of violent crime in the neighborhood.

Conclusions

Based on our three sets of analyses, we offer the following tentative conclusions:

First, poverty is stubbornly persistent in terms of its neighborhood concentration, which is somewhat surprising when considered in relation to the common emphasis in urban poverty writings on the notion of urban social transformation (Wilson 1987). While poverty did increase in the 1970s and 1980s, its pathways were set by initial conditions.

Second, and relatedly, neighborhoods remain remarkably stable in their relative economic standing despite the inflow and outflow of individual residents. In fact, about 70 percent of the variance in poverty lies between neighborhoods rather than within neighborhoods over time, which means that

the overall pattern of neighborhood inequality did not change much over time. There is something enduring about the poverty vulnerability of neighborhoods that is not simply a matter of the current income of residents.

Third, what change does occur reveals strong patterns of asymmetry by race and class, suggesting that once a neighborhood is beyond a certain threshold or “tipping point” of either percent black or percent poor—but especially the former—further change is invariably in the direction of greater racial homogeneity and more poverty. Amazingly, not one neighborhood more than 40 percent black in 1970 became predominantly white at a later time (figure 3). By contrast, a large number of white neighborhoods turned black even as the polar extremes (all-black and all-white neighborhoods) remained the dominant pattern. Neighborhoods also tended either to stay in the same poverty category or move to a higher poverty category over time (table 1)—upgrading (e.g., gentrification) was quite rare at least from 1970 to 1990.⁸ Linear models fail to capture these dimensions of asymmetric change.

Fourth, in multivariate analyses we demonstrated that initial conditions of poverty independently predict the rate of increase in later poverty controlling for race, ethnicity, and home ownership. This conclusion holds whether or not we examine raw change or residual changes that take account of the city’s baseline levels of poverty. Furthermore, black neighborhoods are at special risk for rapid poverty increases, whereas home ownership emerges as a major protective factor at the neighborhood level.

Fifth, neighborhoods that changed the most were those that were spatially proximate to initial concentrations of poverty, so poverty change was geographically systematic—so much so that spatial proximity to change in poverty matters as much as the internal characteristics of a neighborhood, in-

cluding its initial level of poverty. In particular, proximity to increasing poverty is directly related to the rate of change in poverty—whether raw or residual changes relative to the city as a whole.

Finally, the consequences of durable and increasing poverty appear to be long-lasting, at least with respect to predicting key social processes. Controlling for the socio-demographic location of individual respondents, both persistent poverty and increases in poverty from 1970 to 1990 predict lower collective efficacy and the moral cynicism of neighborhood residents in 1995—a span of some twenty-five years. This finding is consistent with the scenario that certain urban neighborhoods get locked into structural dynamics that generate systematic social processes that in turn may contribute to their further stigmatization and deepening of poverty.

NOTES

1. We measure neighborhood poverty as the percentage of families in poverty and refer to this measure hereafter as either neighborhood poverty or the neighborhood poverty rate.

2. This “mixed” category does include one neighborhood that was 79.7 percent Hispanic in 1970. However, no other neighborhoods in this category had a 75 percent majority of any one racial/ethnic group. If this one Hispanic neighborhood, which increased from 18.9 percent poor in 1970 to 35.1 percent poor in 1990, is removed and the correlation between poverty in 1970 and change between 1970 and 1990 is recalculated for the mixed category, the correlation becomes even more negative ($r = -.12$), but still not significant, in part because of the small sample size ($n = 31$). Nonetheless, this analysis suggests that the dynamics of neighborhood change were somewhat different in mixed neighborhoods; we intend to investigate this issue further in future research.

3. We would note, however, that segregation seems to increase with increasing percent black in a metropolitan area. Many cities in places such as Maine, New Hampshire, and Montana have extremely low segregation levels, but then again these places are also almost lily white. We would be hard pressed to consider them racially progressive.

4. Residual change scores have the advantage over raw change scores of being statistically independent of initial levels of poverty. In addition, since all NCs in Chicago were used to estimate the regression equation upon which the residuals are computed, these change scores incorporate the dynamics of the entire ecological system (see also Bursik and Webb 1982).

5. Because ρ is multiplied by the β coefficient for each X variable in equation (2), and $0 \leq \rho \leq 1$, it is possible to think of ρ as the rate at which the “effects” of each X variable are “discounted” in contiguous neighbors. Thus, if $\rho = .50$, then the effects of the average level of X in the first-order neighbors (W_x) will be half as strong as they are in the focal neighborhood. In the second-order neighbors, the effect will be reduced by one-quarter the size of β ($.50^2 = .25$), and so on for each successive order of contiguity.

6. There are alternative interpretations of how neighborhoods become trapped in poverty, the most common one being the differential selection of individuals into and out of neighborhoods. Poor neighborhoods are often seen as undesirable, especially with respect to housing (Galster 1987), and so it follows that richer individuals are better able to capitalize on the desire for living in better endowed neighborhoods. Such a natural “sorting” process would imply that poverty changes are unevenly spread throughout the city. On the other hand, racial segregation is clearly a barrier to the pure selection argument (Massey and Denton 1993)—neighborhoods may be said to select individuals as well, constraining their sorting decisions. Moreover, our theoretical argument implies that neighborhoods that are the poorest have the least collective capacity to counter threats to neighborhood quality, and so their public spaces, local government services, level of crime, and rate of housing abandonment will make them the least desirable neighborhoods and increase their stigma, leading to further decline. As noted earlier, the data needed to fully disentangle these processes are not available (for a discussion of different approaches and an assessment of the evidence on migration, see Quillian 1999).

7. Because of the correlation between initial status and percentage change in poverty, for these analyses we examine residual poverty change from 1970 to 1990, which is, by definition, independent of 1970 poverty. We prefer residual poverty on theoretical grounds as well, because the ongoing ecological dynamics of Chicago neighborhoods are taken as the backdrop against which changes in poverty are assessed (see also Bursik and Webb 1982).

8. We acknowledge that focusing solely on poverty rates, which tap only one end of the distribution of neighborhood socioeconomic status, may lead us to underestimate the amount of gentrification that did take place. Moreover, gentrification may have occurred at a geographic level lower than that of the neighborhood cluster, such as the census tract or block group. We are currently exploring this issue, along with an analysis of concentrated affluence that explicitly takes into account the upper end of the income distribution. Interestingly, our preliminary analysis reveals that initial conditions of poverty in 1970 do predict later pathways of economic change at the upper end of the distribution. Specifically, using an inequality measure that taps the proportional balance of incomes (affluent families – poor families/total families), poverty in 1970 shows a strong negative relationship to later concentrated affluence. Once a neighborhood reached approximately a 25 percent level of poverty in 1970, it never climbed above 0 (equality) on the proportional inequality index in later decades.

REFERENCES

- Anselin, Luc. 1988. *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic.
- . 1995. *SpaceStat Version 1.80: User's Guide*. Morgantown: West Virginia University Press.
- Bourdieu, Pierre. 1986. “The Forms of Capital.” In *Handbook of Theory and Research for the Sociology of Education*, ed. J. Richardson, 241–58. New York: Greenwood Press.
- Bowles, Samuel. 2000. “Persistent Inequality in a Competitive World: Causes, Consequences, Remedies.” Proposal to the Russell Sage Foundation and Santa Fe Institute.
- Bryk, Anthony, and Stephen W. Raudenbush. 1992. *Hierarchical Linear Models*. Newbury Park, Calif.: Sage Publications.
- Bullough, Bonnie. 1967. “Alienation in the Ghetto.” *American Journal of Sociology* 72: 469–78.
- Bursik, Robert J. Jr., and Jim Webb. 1982. “Community Change and Patterns of Delinquency.” *American Journal of Sociology* 88: 24–42.
- Cohen, Cathy, and Michael Dawson. 1993. “Neighborhood Poverty and African American Politics.” *American Political Science Review* 87: 286–303.
- Fischer, Claude S. 1995. “The Subcultural Theory of Urbanism: A Twentieth-Year Assessment.” *American Journal of Sociology* 101: 543–77.
- Galster, George. 1987. *Homeowners and Neighborhood Reinvestment*. Durham, N.C.: Duke University Press.

- Garner, Catherine, and Stephen W. Raudenbush. 1991. "Neighborhood Effects on Educational Attainment: A Multilevel Analysis." *Sociology of Education* 64: 251–62.
- Hoff, Karla, and Arijit Sen. "Homeownership, Community Interactions, and Segregation." *American Economic Review* 95 (4), 2005, 1167–89.
- Jargowsky, Paul. 1997. *Poverty and Place: Ghettos, Barrios, and the American City*. New York: Russell Sage Foundation.
- Kapsis, Robert E. 1978. "Black Ghetto Diversity and Anomie: A Sociopolitical View." *American Journal of Sociology* 83: 1132–53.
- Massey, Douglas S., and Nancy Denton. 1993. *American Apartheid: Segregation and the Making of the Underclass*. Cambridge, Mass.: Harvard University Press.
- Morenoff, Jeffrey, and Robert J. Sampson. 1997. "Violent Crime and the Spatial Dynamics of Neighborhood Transition: Chicago, 1970–1990." *Social Forces* 76: 31–64.
- Morenoff, Jeffrey, Robert J. Sampson, and Stephen Raudenbush. 2001. "Neighborhood Inequality, Collective Efficacy, and the Spatial Dynamics of Urban Violence." *Criminology* 40: 213–30.
- Quillian, Lincoln. 1999. "Migration Patterns and the Growth of High-Poverty Neighborhoods, 1970–1990." *American Journal of Sociology* 105: 1–37.
- Raudenbush, Stephen, and Robert J. Sampson. 1999. "Ecometrics: Toward a Science of Assessing Ecological Settings, with Application to the Systematic Social Observation of Neighborhoods." *Sociological Methodology* 29:1–41.
- Sampson, Robert J., and Dawn Jeglum Bartusch. 1998. "Legal Cynicism and (Subcultural?) Tolerance of Deviance: The Neighborhood Context of Racial Differences." *Law and Society Review* 32: 777–804.
- Sampson, Robert J., Stephen W. Raudenbush, and Felton Earls. 1997. "Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy." *Science* 277: 918–24.
- Srole, Leo. 1956. "Social Integration and Certain Corollaries: An Exploratory Study." *American Sociological Review* 21: 709–16.
- Tilly, Charles. 1998. *Durable Inequality*. Berkeley: University of California Press.
- Tolnay, Stewart, Glenn Deane, and E. M. Beck. 1996. "Vicarious Violence: Spatial Effects on Southern Lynchings, 1890–1919." *American Journal of Sociology* 102: 788–815.
- Warren, Donald. 1975. *Black Neighborhoods: An Assessment of Community Power*. Ann Arbor: University of Michigan Press.
- Wilson, William Julius. 1987. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press.
- Woolcock, Michael. 1998. "Social Capital and Economic Development: Toward a Theoretical Synthesis and Policy Framework." *Theory and Society* 27: 151–208.

■ POVERTY AND SOCIAL POLICY ■

41 ■ Bruce Western

Incarceration, Unemployment, and Inequality

In the later 1990s, two remarkable conditions influenced the employment opportunities of young, unskilled men in the United States. First, the unemployment rate fell to its lowest level in 30 years—around 4.5 percent by the summer of 1998. Second, the incarceration rate rose to its highest level in U.S. history. In 1998 1.78 million men were detained in American prisons and jails.

In removing so many from the labor force, U.S. incarceration policy has had significant but largely overlooked effects on unemployment. In the short run, it has lowered conventional measures of unemployment by concealing joblessness among a large group of able-bodied, working-age men. In the long run, it may raise unemployment rates by curtailing the job prospects of ex-convicts and will certainly deepen economic inequality, because its effects are increasingly detrimental to young, black, unskilled men, whose incarceration rates are the highest of all and whose market power is weak. In this article I explore both short- and long-run effects.

The Penal System as a Labor Market Institution

In industrial relations and welfare state development, the United States stands apart from other major western nations. Its employment system is far more market-driven: rates of

unionization and unemployment insurance coverage are lower than in most other countries in the Organisation for Economic Cooperation and Development (OECD), and social welfare spending accounts for only about 15 percent of gross domestic product (GDP), compared to about a quarter of GDP in the large European countries.

In recent years, the highly regulated European labor market has been burdened with persistent unemployment and apparently stagnant rates of economic growth, whereas the unregulated U.S. labor market has seen unemployment steadily falling in a booming economy. These trends have been used to buttress the argument that unregulated labor markets yield stronger economic performance than highly regulated markets.

But the argument is too narrowly framed. Labor markets are embedded in a wide array of social arrangements that extend beyond the welfare state or industrial relations. The weakness of social protection mechanisms does not alone explain the superior U.S. employment record. In the United States, criminal justice policy constitutes a significant state intervention, providing a sizeable, non-market reallocation of labor that has significant effects on employment trends.

The magnitude of this intervention is reflected in incarceration and budget figures. Rates of incarceration began to increase in

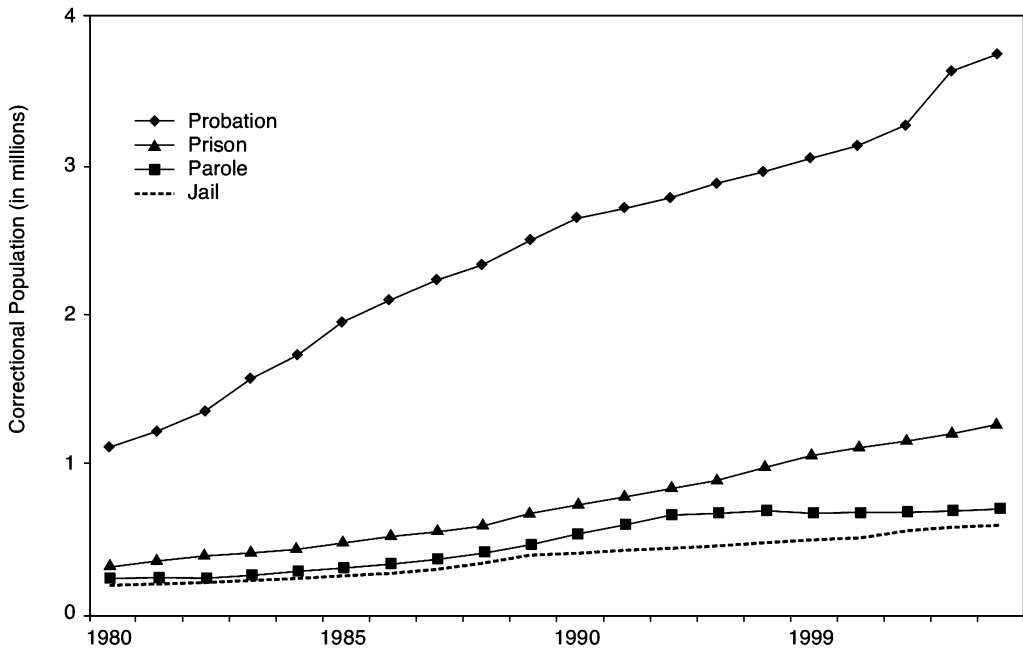


Figure 1. Adult Correctional Populations, 1980–99.

Source: Bureau of Justice Statistics, <http://www.ojp.usdoj.gov/bjs/glance/corr2.txt>.

the early 1970s, but the most rapid growth took place in the 1980s and 1990s (see Figure 1). Between 1980 and 1999, the number of people in prisons and jails in the United States grew from half a million to almost two million. . . .

Incarceration is spread unevenly across the adult population. Men make up more than 90 percent of all inmates. In the mid-1990s, about two-thirds of those inmates were under 35, and about half had not completed high school. The dramatic expansion of the prison population most seriously affected young African Americans. In 1930, blacks accounted for 22 percent of all those in prison. In 1992, over half the prison population was black. By 1995, one out of three black male youths was under some form of state supervision and nearly 7 percent of all black males were incarcerated.

The U.S. incarceration rate—the number incarcerated on a single day per 100,000 of the adult population—is very much greater

than incarceration rates in other industrialized democracies. In 1992–93, for instance, the overall U.S. rate was 5–10 times greater than the rate for other OECD countries; among American blacks, the incarceration rate was 20 times greater (Figure 2). These high rates correspond to large absolute numbers. The entire Western European prison population is measured in hundreds of thousands, the prison and jail population in the United States in millions. This disparity does not reflect higher crime rates in the United States. The evidence, indeed, suggests that U.S. crime rates are only slightly above the average among industrialized countries. Thus U.S. incarceration rates appear to result from more aggressive prosecutorial practices, tougher sentencing standards, and intensified criminalization of drug-related activity.

The Short-Run Effect of Incarceration

The performance of national labor markets is commonly summarized by the unemployment

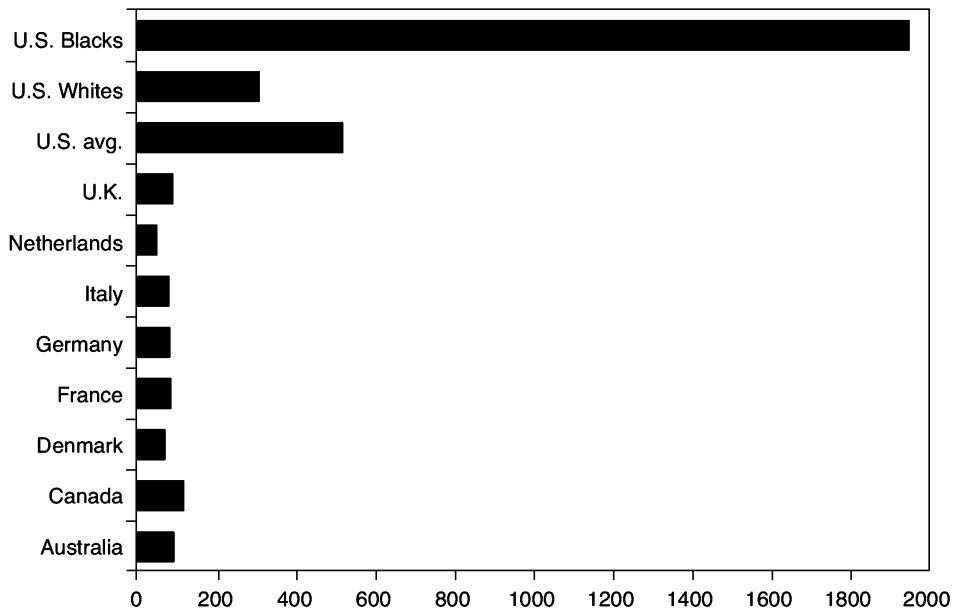


Figure 2. Incarceration Rate in Selected OECD Countries, 1992–93.

Source: B. Western and K. Beckett, "How Unregulated Is the U.S. Labor Market? The Penal System as a Labor Market Institution," *American Journal of Sociology* 104, no. 4 (1999): 130–60, Table 2.

rate—the percentage of unemployed individuals among the civilian labor force. This definition excludes from its calculations those in the military and those in prison and jail, as well as the “discouraged,” those no longer looking for work and therefore out of the labor force. A broader concept tries to tap the idleness or economic dependence of a group by extending the definition of “unemployed” to include those incarcerated.

In the short run, incarceration keeps those with a high risk of unemployment out of the labor market. National surveys of prisons and jails from the early 1990s indicate that, on average, more than a third of male inmates were unemployed at the time they were sent to prison. If, for example, we count among the unemployed those inmates who were not working when incarcerated, the adjusted unemployment rate for 1995 was 6.2 percent, versus the conventional rate of 5.6 percent.

If we pursue the calculation further, including all inmates among the unemployed,

the adjusted U.S. unemployment rate for 1995 rises to 7.5 percent, an increase of 1.9 percentage points over the conventional rate.¹ According to this adjusted measure, U.S. rates of labor inactivity never fell below 7 percent throughout the 1980s. In the economically buoyant period of the mid-1990s, the rate of inactivity was about 8 percent, higher than any conventional unemployment rate since the recession of the early 1980s.

In Europe, the short-run effect of incarceration is tiny, because incarceration rates are so low. In all European countries, unemployed males outnumber imprisoned males by very large ratios—between 20:1 and 50:1 (in the United States the ratio is 3:1). Including prison inmates in the jobless count, therefore, changes the unemployment rate for most European countries by only a few tenths of a percentage point, a striking contrast with the large U.S. difference.

By the conventional measure, the United States enjoyed consistently lower unemploy-

ment than Europe after the mid-1980s. But adjusted figures that count the incarcerated population as unemployed suggest that the U.S. labor market performed worse, not better, than Europe for most of the period between 1976 and 1994.

The Long-Run Effect of Incarceration

The long-run effects of incarceration highlight the employment experiences of convicts after they are released. Ex-convicts must reintegrate themselves into mainstream society, surmounting the psychological, social, and financial consequences of imprisonment. In this process, the ability to find stable, legal employment is crucial, yet job prospects for ex-convicts are poorer than the prospects of applicants with no criminal record. Incarceration erodes the value of vocational skills, and the increasingly violent and overcrowded state of prisons and jails produces attitudes and practices that may enhance survival in the prison but are not compatible with success in the conventional job market. Job prospects may be even worse in the current context, because resources for educational and vocational training in prisons have declined. Such experiences are better observed through survey data than through aggregate labor market statistics; I drew upon the National Longitudinal Survey of Youth (NLSY), which surveys a national sample of Americans aged 14 to 21 in 1979, to follow workers as they moved from prison to the labor market. I examined employment as a function of youth incarceration, jail time, and work experience, controlling also for personal and regional characteristics such as the extent of juvenile contact with the criminal justice system and the local unemployment rate. I included employment status both before and after incarceration, because the same characteristics that place men at risk of unemployment or low wages also raise their chances of criminal conviction.

The regression analysis indicates that youth who spent time in prison worked less four years later than youth who had no prison time. Incarceration, on average, reduced employment by about 5 percentage points, or three weeks a year. The effect is particularly large for black youth, whose employment was reduced by about 9 percentage points (5 weeks) by juvenile jail time. The effect of jail time on adult employment was greater than the effect of dropping out of high school or living in a high-unemployment area. Moreover, these effects did not decay over time. Even after 15 years, men incarcerated as juveniles worked between 5 and 10 percentage points less than their counterparts who never went to prison. The effects of adult incarceration, although large in the short run (5–10 weeks a year), were less persistent, and largely disappeared within four to five years of release.

With nearly two million men now in prison, these effects are not small. The penal system, when viewed as a labor market institution, appears to significantly undermine the productivity and employment chances of the male workforce.

The Penal System and Racial Inequality in Employment

A large and growing proportion of young black men has had experience with the penal system; can we estimate the likely effect on their future prospects and on black-white economic inequality?

The short-run effects of incarceration are very much worse for black than for white men. As the prison and jail population grew throughout the 1980s, the labor market effects of incarceration for black men becomes much larger if all those incarcerated are counted among the unemployed. In 1990–94, the adjusted unemployment rate is only one percentage point higher for white men, but it is seven percentage points higher

for black men. During the 1990s almost one in five African American men, on average, was without a job.

Standard labor force data report a persistent gap in joblessness between black and white men that dates at least from the late 1960s.² Employment-population ratios calculated from these data show that employment inequality grew most sharply for young high school dropouts.³ By including the incarcerated in calculating these ratios, we gain a more accurate picture, because marginalized groups at the fringes of the labor market have especially high incarceration rates.

If we include men in prison or jail, employment among black high school dropouts aged 20–35 declines from 46 to 29.3 percent in 1996. Furthermore, the adjusted employment ratios for black high school dropouts show a steady decline over time, whereas conventional ratios show black employment as stable or even rising slightly.

The long-run effects are also serious. The incarceration of youths disrupts transitions from school to a career and hinders the acquisition of work experience. Ex-prisoners find it difficult to access jobs in the primary sector that offer opportunities for training, pay schedules that rise strongly with experience and age, and other characteristics of stable employment.⁴ Instead, many ex-inmates find themselves stuck in low-wage job trajectories, confined to casual or illegitimate employment in the secondary labor market.

These kinds of disruptions materially affect earnings. In general, black ex-convicts earn about 10 percent less per hour than comparable men who were never incarcerated, after adjusting for such factors as work experience, schooling, youth delinquency,

and drug use. They also have generally flat earnings profiles. My estimates suggest that, by 1998, the disproportionate incarceration of black men and the low-wage job trajectory of ex-inmates, taken together, had raised black-white earnings inequality by about 15 percent. . . .

By the end of the twentieth century, the prison experience was routinely shaping the working lives of young, less-educated, minority men. Because their incarceration rates are so high, labor market and earnings statistics as conventionally measured may significantly underestimate the extent of racial inequality in employment. The evidence suggests that the U.S. prison system may be exercising a systematic influence on large-scale patterns of economic inequality.

NOTES

1. The unemployed are usually defined as those without paid employment who are actively seeking work in the month before the survey.

2. F. Wilson, M. Tienda, and L. Wu, "Race and Unemployment: Labor Market Experiences of Black and White Men, 1968–1988," *Work and Occupations* 22, no. 3 (1995): 245–70.

3. J. Bound and R. Freeman, "What Went Wrong? The Erosion of Relative Earnings and Employment among Young Black Men in the 1980s," *Quarterly Journal of Economics* 107 (1992): 201–32.

4. M. Duncier, *Sidewalk* (New York: Farrar, Strauss, and Giroux, 1999), R. Sampson and J. Laub, *Crime in the Making: Pathways and Turning Points Through Life* (Cambridge, MA: Harvard University Press, 1993), J. Waldfogel, "The Effect of Criminal Conviction on Income and the Trust 'Reposed in the Workmen'," *Journal of Human Resources* 29 (1994): 62–81, and "Does Conviction Have a Persistent Effect on Income and Employment?" *International Review of Law and Economics* 14 (1994): 103–19.

42 ■ James E. Rosenbaum and Stefanie DeLuca

Does Changing Neighborhoods Change Lives?

The Chicago Gautreaux Housing Program

Neighborhood effects on the lives of families and young people have long been an important topic of research, and communities are also an important topic for current public policy. Theoretically, neighborhoods are important contexts for socialization and development as well as places where we see structures of inequality and opportunity in action. Neighborhoods are also significant because they are closely tied to schooling opportunities, given the zoning of public schools. The possibility of choosing different schools, including schools in different neighborhoods, is intended to be a central piece of the No Child Left Behind legislation, and federal courts are currently considering whether to mandate racial or socioeconomic integration in housing and school settings (*Thompson v. HUD*; *Meredith* and *Parents* cases). Neighborhoods have also become the focus of many recent policy discussions. Residential mobility and housing policy garnered national attention after the hurricane disaster in New Orleans, and HOPE VI demolitions are leading to the relocation of inner city families all over the country.

Despite years of research on these topics, it is hard to know for sure if neighborhoods can be used as policy levers to improve youth and family well-being. This is due in large part to two related issues. First, despite relatively high levels of residential mobility in the United States, we see little variation in the types of communities low-income mi-

nority families inhabit. Often, poor families are trapped in dangerous neighborhoods and their children are trapped in poor schools (South and Deane 1993; South and Crowder 1997; Massey and Denton 1993). Therefore, we don't get the chance to observe how a *different* environment might affect their life chances. Second, families choose neighborhoods, and the characteristics of families that lead them to choose certain neighborhoods are also likely to affect family and child well-being. This leads to the selection problem (endogeneity), which plagues our attempts to recover causal estimates of environmental effects. However, there have been some opportunities to study what happens when parents and children experience moderate to radical changes in their neighborhood or schooling environments. Residential mobility programs, where poor families relocate to opportunity-rich communities via housing vouchers, provide one way we can begin to separate the effects of family background and neighborhood conditions. In this chapter, we review one particularly important mobility plan—Chicago's Gautreaux program—and examine a decade of research following the fortunes of the families who moved as a part of this intervention.

The Gautreaux Program

As a result of a 1976 Supreme Court decision, the Gautreaux program allowed low-income

black public housing residents in Chicago to receive Section 8 housing certificates (or vouchers) and move to private-sector apartments either in mostly-white suburbs or within the city. Between 1976 and 1998, over 7,000 families participated, and over half moved to suburban communities. Because of its design, the Gautreaux program presents an unusual opportunity: it allows us to examine whether individual outcomes change when low-income black families move to safer neighborhoods with better labor markets and higher quality schools.

Gautreaux participants circumvented the typical barriers to living in suburbs, not by their jobs, finances, or values, but by acceptance into the program and quasi-random assignment to the suburbs. The program provided housing subsidy vouchers and housing support services, but not employment or transportation assistance. Unlike the usual case of working-class blacks living in working-class suburbs, Gautreaux permitted low-income blacks to live in middle- and upper-income white suburbs. Participants moved to more than 115 suburbs throughout the six counties surrounding Chicago. Suburbs with a population that was more than 30 percent black were excluded by the consent decree. A few very high-rent suburbs were excluded by funding limitations of Section 8 certificates.

Early Findings

Early research on Gautreaux had shown large and significant relationships between placement neighborhoods and subsequent gains in employment and education. A study of 330 Gautreaux mothers in the early 1990s found that suburban movers had higher employment levels than city movers, but not higher earnings, and the employment difference was especially large for adults who were unemployed prior to the move (Rosenbaum 1997). Another study found that as young adults, Gautreaux children who moved to the suburbs were more likely than city movers to

graduate from high school, attend college, attend four-year colleges (vs. two-year colleges), and if they were not in college, to be employed and to have jobs with better pay and with benefits (Rosenbaum 1995). These differences were very large, often larger than the effects of education and training programs targeted at these specific outcomes.

Analyses indicated that children moving to suburbs were just as likely to interact with neighbors as city movers, but the suburb movers interacted with white children while city movers interacted mostly with black children. The program seems to have been effective at integrating low-income black children into middle-class white suburbs. Although suburban schools were often far ahead of city schools in terms of curriculum level, mothers reported that suburban teachers often extended extra efforts to help their children catch up with the class. Initial concerns that these children would not be accepted were unsupported by the evidence.

Recent Research

To improve upon the design and data quality of the earlier work, more recent research used administrative data to locate recent addresses for a 50 percent random sample of Gautreaux movers who had relocated before 1990, as well as track economic outcomes for mothers.¹ Additionally, multiple census measures were used to characterize neighborhoods and a more comprehensive accounting for preprogram characteristics was employed in the regression models. The use of administrative records permitted us to locate 1,504 of 1,507 families, and we found that 66 percent of suburban movers remained in the suburbs an average of 15 years after placement. After premove individual and neighborhood attributes were controlled, the racial composition of placement neighborhood predicted racial composition of current neighborhood (DeLuca and Rosenbaum 2003), and mothers continued to live in areas with much

lower poverty rates and higher household incomes (Keels et al. 2005).

Individual level economic outcomes, such as welfare receipt, employment, and earnings, were also influenced by the income and racial characteristics of placement neighborhoods. Women who moved to racially mixed or predominantly white neighborhoods with higher levels of socioeconomic resources did slightly better than their counterparts in areas with low resources and high levels of black residents (Mendenhall, DeLuca, and Duncan 2006). Research on the children of the original Gautreaux families has demonstrated that the neighborhoods where they resided in the late 1990s were substantially more integrated than their overwhelmingly minority origin neighborhoods (Keels 2007a). However, the effects of relocating to lower poverty, more integrated areas had a mixed effect on the delinquent behaviors and arrest rates of boys versus girls. Suburban boys were much less likely to become involved in the criminal justice system, while girls who moved to the suburbs were more likely to be convicted for criminal offenses (Keels 2007b).

How Did Gautreaux “Work”?

The findings described above focus on the advances made in recent quantitative work. We had employed techniques to approximate the assessment of Gautreaux as a “treatment”—a social intervention with effects we might measure with statistical corrections and design comparisons. However, the stories Gautreaux participants tell about their experiences can contribute greatly to our understanding. The long-term family outcomes we observed appear to be significantly linked to the mobility program and the characteristics of the placement neighborhoods. However, administrative data cannot tell us *how* these outcomes occurred, or the mechanisms through which neighborhoods have their impact. This is a problem common to neigh-

borhood research, and one that makes improving mobility programs especially difficult. However, in several qualitative studies (Rosenbaum, Reynolds, and DeLuca 2002; Rosenbaum, DeLuca, and Tuck 2005), we analyzed interviews with mothers who described how these neighborhoods helped improve their lives and the lives of their children. Was it a matter of just increasing access to better resources, or was it necessary to interact with neighbors to obtain the full benefit of these new resources?

We analyzed interviews with 150 Gautreaux mothers and found that after the move, they described a new sense of efficacy and control over their lives and that the major changes in their environments helped them to see that they had the ability to make improvements in their lives. Certain features of the new suburban neighborhoods changed their perception of what was possible. Specifically, the women reported that they felt better about having an address in the suburbs, and not having to put down a public housing address on job applications. Other women noted that by moving to areas with more white residents, they and their children got to know more white people, and racial stereotypes were debunked. One child whose only exposure to white people were those she saw on TV reported that after moving, she discovered that not all whites looked like TV actors.

Social interactions with whites allowed some of these women to feel that they had more social and cultural know-how and feel much less intimidated by future contexts in which they might have to interact with whites. Additionally, working through some of the initial difficulties of the transitions to the suburbs allowed these women to realize that they could handle manageable challenges along the way to better jobs and more schooling. In comparison, the drugs or gang violence in their old city neighborhoods seemed to be forces too big for them to control and therefore permanent impediments to the advancements they were trying to

make in their lives. These findings suggest to us that one's repertoire of capabilities can vary depending on the type of neighborhood one lives and works in.

Many of the mothers we interviewed also noted that they had to change their way of behaving to comply with the social norms of the new neighborhoods. Several women noted initial difficulties in adjusting to suburban norms, which were unfamiliar and intolerant of some of their prior behaviors. These mothers, who have lived all their lives in housing projects where these norms did not exist, saw benefits to complying with these expectations, and they decided to adopt them. Ironically, some of these normative constraints, such as low tolerance for drugs and parties, were liberating because the trade-off was community safety. This meant that mothers did not have to spend all their time watching their children, and these norms allowed mothers to give their children more freedom.

Similarly, mothers reported social responsiveness from their neighbors. They received the benefits of reciprocal relations related to child care, and neighbors' watchfulness promoted the safety of their children, their property, and themselves. They were also given favors in terms of transportation and some acts of charity. It is remarkable that these new residents, who generally differed in race and class from their neighbors, were awarded this collective generosity, and the interviews suggest that it may have been conditional on their showing a willingness to abide by community norms.

Most important, the new suburban social contexts provided a form of capital that enhanced people's capabilities. Some mothers reported that they could count on neighbors if their child misbehaved or seemed at risk of getting into trouble, if their child was sick and couldn't attend school, or if there was some threat to their children, apartments, or themselves. This was not just interpersonal support, it was systemic, and enabled these mothers to take actions and make commit-

ments that otherwise would be difficult or risky. For instance, some mothers reported a willingness to take jobs because they could count on a neighbor to watch their child in case they were late getting home from work. It is through some of these mechanisms—some social, some psychological—that we believe the Gautreaux families were able to permanently escape the contexts and consequences of segregated poverty and unsafe inner-city neighborhoods.

More recent interviews with Gautreaux mothers suggest that some aspects of the city-suburban divide were also important for shaping how the placement community affected their children's behavior (Keels 2007b; Mendenhall 2004). City movers placed in both moderate- and low-poverty neighborhoods found that although their immediate neighborhood was safe, the larger community to which their children had easy access continued to be dangerous. In comparison, children placed in the suburbs had less direct neighborhood exposure to drugs and illegal activities and attended higher-performing public schools with greater financial and teacher resources. Interviews revealed that affluent suburban neighborhoods also had substantially fewer opportunities for involvement in delinquent criminal activities and gangs.

Was Gautreaux a Social Experiment?

Methodologically, we often rely on observational data and regression analyses to provide estimates of the "effect" of neighborhood contexts and interventions. These approaches have their weaknesses; it is complicated, if not impossible to infer causal effects when we know that there are unobservable characteristics of families that lead not only to their selection of neighborhood, but also to the outcomes of interest. As a result, there has been an increased push to employ experimental designs to control for and assign social and economic "treatments," be they neighborhoods, school programs, or income subsidies.

Along these lines, the Gautreaux program resembled a quasi-experiment. Although the program was not designed as an experiment and families were not formally randomly assigned to conditions, aspects of the program administration break the link between family preferences and neighborhood placement. In principle, participants had choices about where they moved. In practice, qualifying rental units were secured by rental agents working for the Gautreaux program and offered to families according to their position on a waiting list, regardless of their locational preference. Although participants could refuse an offer, few did so, since they were unlikely to ever get another.² As a result, participants' preferences for placement neighborhoods had relatively little to do with where they ended up moving, providing a degree of exogenous variability in neighborhood placement that undergirds Gautreaux research. Few significant differences were found between suburban and city movers' individual characteristics, but premove neighborhood attributes show small, but statistically significant differences on two of nine comparisons. This may indicate selection bias, although random assignment studies by the HUD-sponsored Moving to Opportunity (MTO) also find some substantial differences (Goering and Feins 2003, Table 7.1). It is not clear whether the observed premove differences explain much of the outcome difference. For instance, while suburban movers came from *slightly* lower-poverty tracts than city movers (poverty rate of 40.6 percent vs. 43.8 percent), they moved to census tracts with *dramatically* lower poverty rates (5.0 percent vs. 27.3 percent; DeLuca and Rosenbaum 2003). While small (3 percentage points) differences in initial neighborhoods may account for a portion of the outcome differences, it is hard to dismiss the possible influence of the vast differences in placement neighborhoods. Current papers have discussed these issues at length and examine multiple neighborhood level indicators, detailed pre-

program neighborhood differences, and intergenerational effects (DeLuca and Rosenbaum 2003; Keels et al. 2005; Mendenhall, DeLuca, and Duncan 2006; DeLuca et al. forthcoming; Keels 2007a and 2007b).

In contrast, MTO was an experiment, with the random assignment of low-income families to three conditions—an experimental group (who moved to low-poverty census tracts), an open-choice Housing Voucher group, and a “no move” control group. MTO was developed to formally test the Gautreaux findings, with more rigorous design, and pre/post move data collection. Unfortunately, while MTO was a stronger study, it was a weaker “neighborhood change treatment” in some respects. The Gautreaux program moved families an average of 25 miles away from their original neighborhood, to radically different labor markets, where nearly all children attended schools with above-average achievement and were too far away to interact with prior friends. In comparison, most MTO families moved less than 10 miles away in mostly city neighborhoods, most children attended schools with very low achievement, and many children continued interacting with old friends. In addition, MTO occurred in the hot labor market of the late 1990s, and large numbers of families in the control group moved out of high-rise housing projects through the federal HOPE VI program, so the control group was experiencing unusual benefits and atypical circumstances which may not generalize to more ordinary times.

While early Gautreaux analyses showed that suburban children attended much better schools and enjoyed improvements in educational outcomes relative to the city movers, the MTO program did not have such an effect on educational outcomes. Compared to the control group, the MTO treatment group showed no difference in test scores, school dropout, or self-reported measures of school engagement four to seven years after random assignment (Sanbonmatsu et al. 2006). This was due in part to the fact that

many MTO experimental families sent their children to schools in the same school district (often the same schools), and even when they changed schools, the new schools were not much better than the original schools.

While Gautreaux was associated with gains in mothers' employment, the MTO treatment group showed no impact compared with the control group—both groups showed large gains of comparable magnitude. However, MTO outcomes were measured in the late 1990s, during a strong labor market and strong welfare reform, so, although MTO found no difference between groups, it found a 100 percent employment gain for the control group. One possible interpretation is that virtually everyone who could work was doing so, and residential moves had no additional effect for that reason.

Both Gautreaux and MTO also found large effects on mothers' and children's feelings of safety. Despite the shorter moves and less dramatic change in social environment, MTO also showed significant reductions in depression and obesity among mothers, and lower levels of anxiety, arrests, and substance use among daughters (but an increase in arrests for sons) (Kling, Liebman, and Katz 2007; Kling, Ludwig, and Katz 2005). Gautreaux studied neither of these outcomes.

When comparing the two programs, it is crucial to understand the nature of the comparisons that are being made. Although social scientists have been concerned with learning what are likely benefits of certain kinds of neighborhood moves, what policy makers need to know is how a family fares when policy offers them the opportunity to move to a lower-poverty or less segregated neighborhood, relative to what would have happened to that family had it not been given that opportunity. Gautreaux research studies can only compare subgroups of families that moved in conjunction with the program and experienced variation in neighborhood contexts—there is no comparison group of similar families who did not move as part of the

program. MTO's evaluation design is much stronger since it tracked the fortunes of a randomly assigned control group of families who expressed interest in the program but, owing to the luck of the draw, were not assigned to the move (although some managed to move on their own). At the same time, however, the program-driven assignment procedures of Gautreaux can inform us about what happens when families move long distances to radically different neighborhoods, moves which changed their social context in many ways—new neighborhoods were more racially integrated and less poor, and the communities had higher levels of school quality, labor market strength, and safety. While both studies indicate how powerful the effects of residential moves can be for some families, the differences in findings indicates the importance of program design features, historical context influences and concurrent policy effects. For example, alternative forms of mobility, such as those under the involuntary conditions of HOPE VI, may have different results.

Many policy reforms have tried to improve individuals' education or employability while they remain in the same poor schools or labor markets, but these reforms have often failed. Such policies may be fighting an uphill battle as long as families remain in the same social contexts and opportunity structures. In contrast, Gautreaux findings suggest that housing policy is one possible lever to assist poor families, by moving them into much better neighborhoods, with much better schools and labor markets. The initial gains in neighborhood quality that many of the Gautreaux families achieved persisted for at least one to two decades. The Gautreaux findings suggest that it is possible for low-income black families to make permanent escapes from neighborhoods with concentrated racial segregation, crime, and poverty, and that these moves are associated with significant gains in education, employment, and racially integrated friendships, particularly for children. However, as the MTO findings

suggest, there is much that we still need to learn about what kinds of moves are required to make major changes in outcomes, and like MTO, strong research designs will be needed to remove alternative interpretations.

NOTES

1. The collection and preparation of the data used for the more recent studies executed from 2000–2005 were done as a collaborative effort involving: Stefanie DeLuca, Greg J. Duncan, James E. Rosenbaum, Ruby Mendenhall, and Micere Keels. Various authors collaborated on each study, as cited in the references.

2. Although only about 20 percent of the eligible applicants ended up moving through the program, self-selection does not appear to have played a big role in program take-up (Peterson and Williams 1995). Rather than opting out of the program, most nonmoving families were not offered a housing unit and thus not given the chance to participate. Housing counselors were forbidden by the consent decree from making offers selectively among eligible families, and there is no evidence that they did so.

REFERENCES

- DeLuca, Stefanie, Greg Duncan, Ruby Mendenhall, and Micere Keels. Forthcoming. "Gautreaux Mothers and Their Children: An Update." *Housing Policy Debate*.
- DeLuca, Stefanie, and James Rosenbaum. 2003. "If Low-Income Blacks Are Given a Chance to Live in White Neighborhoods, Will They Stay? Examining Mobility Patterns with Quasi-Experimental Data." *Housing Policy Debate* 14: 305–345.
- Goering, John, and Judith Feins. 2003. *Choosing a Better Life?* Washington, DC: Urban Institute Press.
- Keels, Micere. 2007a. "Residential Attainment of Now-Adult Gautreaux Children: Do They Gain, Hold, or Lose Ground in Neighborhood Ethnic and Economic Segregation?" Working paper.
- _____. 2007b. "Second-Generation Effects of Chicago's Gautreaux Residential Mobility Program on Children's Participation in Crime." Working paper.
- Keels, Micere, Greg J. Duncan, Stefanie DeLuca, Ruby Mendenhall, and James E. Rosenbaum. 2005. "Fifteen Years Later: Can Residential Mobility Programs Provide a Permanent Escape from Neighborhood Crime and Poverty?" *Demography* 42 (1): 51–73.
- Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75: 83–119.
- Kling, Jeffrey R., Jens Ludwig, and Lawrence F. Katz. 2005. "Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Voucher Experiment." *Quarterly Journal of Economics* 120: 87–130.
- Massey, Douglas, and Nancy Denton. 1993. *American Apartheid*. Cambridge, MA: Harvard University Press.
- Mendenhall, Ruby. 2004. "Black Women in Gautreaux's Housing Desegregation Program: The Role of Neighborhoods and Networks in Economic Independence." Unpublished dissertation. Evanston, IL: Northwestern University.
- Mendenhall, Ruby, Stefanie DeLuca, and Greg Duncan. 2006. "Neighborhood Resources and Economic Mobility: Results from the Gautreaux Program." *Social Science Research* 35: 892–923.
- Crystal D. Meredith v. Jefferson County Board of Education*, United States Supreme Court, Docket No. 05-915.
- Parents Involved in Community Schools v. Seattle School District*, United States Supreme Court, Docket No. 05-908.
- Rosenbaum, James E. 1995. "Housing Mobility Strategies for Changing the Geography of Opportunity." *Housing Policy Debate* 6 (1): 231–70.
- Rosenbaum, James E., Stefanie DeLuca, and Tammy Tuck. 2005. "Crossing Borders and Adapting: Low-Income Black Families in Suburbia." Pp. 150–75 in *The Geography of Opportunity: Race and Housing Choice in Metropolitan America*, edited by Xavier de Souza Briggs. Washington, DC: Brookings Institution.
- Rosenbaum, James E., Lisa Reynolds, and Stefanie DeLuca. 2002. "How Do Places Matter? The Geography of Opportunity, Self-Efficacy, and a Look Inside the Black Box of Residential Mobility." *Housing Studies* 17: 71–82.
- Sanbonmatsu, Lisa, Jeffrey R. Kling, Greg J. Duncan, Jeanne Brooks-Gunn. 2006. "Neighborhoods and Academic Achievement." *Journal of Human Resources* 41: 649–691.
- South, Scott, and Glenn Deane. 1993. "Race and Residential Mobility: Individual and Structural Determinants." *Social Forces* 72: 147–167.
- South, Scott, and Kyle Crowder. 1997. "Escaping Distressed Neighborhoods: Individual, Community and Metropolitan Influences." *American Journal of Sociology* 102: 1040–1084.
- Thompson v. Department of Housing and Urban Development*. Civil Action No. MJG-95-309 (2005–2006).

43 ■ Elizabeth Clark-Kauffman, Greg J. Duncan, and Pamela Morris

How Welfare Policies Affect Child and Adolescent Achievement

Recent research based on random-assignment experiments has found that the effects of welfare and employment policies appear to vary by children's developmental stage. Programs that increase parents' employment and income have been found to have either neutral or positive effects for preschool and early-school-age children in poverty, depending on the policy approach utilized (Pamela Morris et al., 2001). At the same time, negative effects have been observed for adolescent children (Lisa Gennetian et al., 2002), and the limited research on very young children has shown neutral effects (Morris and Charles Michalopoulos, 2000).

We pool data on over 30,000 achievement reports for children in families enrolled in 14 different random-assignment welfare and work programs to examine how the impacts of policies aimed at increasing parents' employment and their income vary across childhood. Our efforts extend past work by including more programs, using longer-run follow-up data from some of the programs that had been included in prior studies, and pooling micro-level data rather than applying meta-analytic techniques to study-wide impact estimates.

Both economic and psychological theories suggest that changes in parents' economic and employment circumstances may affect children's development, although each emphasizes

a different pathway of influence (Gary S. Becker, 1981; James S. Coleman, 1988; Vonnie McLoyd, 1998). But children in different developmental periods may respond differently to the same changes in employment on the part of their parents. For example, employment in the first year of a child's life may be associated with more negative child outcomes (Jeanne Brooks-Gunn et al., 2002), especially if it involves many hours of work (Elizabeth Harvey, 1999) or is associated with long hours in nonmaternal care (Michael Lamb, 1998; McLoyd, 1998). After-school arrangements are associated with positive outcomes for both pre-adolescent and adolescent children, keeping them in structured care and away from peers (Gregory S. Pettit et al., 1999; Jill K. Posner and Deborah Lowe Vandell, 1999).

At the same time, adolescents may have difficulties if left alone after school and into the evening hours as mothers take on off-hour and shift work; maternal employment has been found to be associated with reduced parental supervision and increased adolescent delinquency among low-income families (Robert J. Sampson and John H. Laub, 1994). Finally, adolescent children may also be asked to take on greater household responsibilities when their single mothers move into employment, which may result in negative academic performance.

Since some of the welfare programs in our pooled data provided financial incentives and boosted not only employment, but also family income, we are able to ascertain whether income-augmenting programs benefit children more than programs that merely increase employment. Note, however, that none of our studies randomly allocated participants into either a financial incentive or a nonfinancial incentive program, so our comparison of impacts between these two types of programs is nonexperimental. Experimental research to date suggests that programs that increase both parents' employment and their income produce positive effects for preschool and early-school-age children even as they negatively affect adolescents (Morris et al., 2001; Gennetian et al., 2002). In contrast, programs without these same effects on parents' income have few effects on preschool and early-school-age children.

Based on this past work, we expect that the effects of welfare and employment programs will change over the childhood age span, with more positive effects for preschool and early-school-age children and more negative effects for adolescents; that this pattern of impacts will be most pronounced for programs with earnings supplements; and that programs that increase parents' transitions from welfare to employment may have their strongest effects on children when these impacts coincide with transition points in children's development (i.e., the transition from preschool to school-age or that from pre-adolescence to adolescence).

Understanding the Source of the Variation in Effects

Age-specific patterns of effects of welfare programs may arise because of differences in the way child characteristics interact with microsystem changes brought about by policy initiatives targeting maternal employment and income. Prior theoretical work would suggest that differences in developmental

tasks, sensitivity to change, and interactions with the environment across childhood may explain differences in the way younger, middle childhood, and adolescent children are affected by maternal employment.

First, responses by younger and older children to changes in parents' employment and income may reflect differences in the extent to which changes in employment and income affect the achievement of developmental tasks, that is, the goodness of the stage-environment fit (Eccles et al., 1993). Children across developmental periods may vary in their responses to policies that affect employment and income because of differences in the fit between these changes in children's proximal environments and the attainment of stage-salient developmental tasks (Sroufe, 1979; Waters and Sroufe, 1983). For example, researchers finding negative effects of maternal employment in the second and third quarter of the first year of life have argued that the parent-child separation that occurs with parental employment interferes with the child's formation of the representation of the parent at the critical moment when these representations are being established (Baydar and Brooks-Gunn, 1991). For older children, parents' transitions to employment that occur at the time of identity development may foster the increasing autonomy needed at this point in development.

Differences across developmental responses to policy changes may also reflect differential sensitivity of children of different ages to changes in development. A number of researchers have argued that early childhood is a sensitive period in development, amenable to intervention (Shonkoff and Phillips, 2000; Sroufe, 1979). Waddington (1957) has described development as proceeding along the branches of a tree; whereas changes in developmental trajectories can occur at any point at which a new branch is formed, the ability of the individual to substantially alter his or her developmental course becomes increasingly difficult over time.

Transitions in development may also serve as important periods in which change may be most likely to occur. Because developmental transitions involve a qualitative reorganization of an individual's organizing systems, these transitions are points at which development is most in flux and open to change in direction (Graber and Brooks-Gunn, 1996). For example, some researchers have found that the transition to adolescence is a time of an increased number of stressful life events (Brooks-Gunn, 1991; Ge, Lorenz, Conger, Elder, and Simons, 1994). Experiencing changes in maternal employment along with the pubertal and life-events stresses that may accompany the transition to adolescence may have negative implications for the development of young adolescents.

Also important is the extent to which the interaction between the child and the environment may change with development. Bronfenbrenner has described the development of the individual as occurring within a nested and interactive set of systems (Bronfenbrenner and Morris, 1998). However, the nature of the interactions among these systems is affected by the developmental stage of the individual. For example, young children are primarily affected by their family context and only indirectly, through interactions with their family, by community or neighborhood contexts. In contrast, older children may be more directly affected by their own interactions within their neighborhoods. Because welfare and employment policy interventions target changes in the family (rather than school or community contexts), they may play a relatively stronger role for younger children's developmental outcomes (McCall, 1981; Yoshikawa, 1994).

Data and Measures

Our pooled sample consists of children aged 0–15 at the point of random assignment whose families were enrolled in seven experimental studies that together tested the effects of 14 different welfare and work programs. Al-

though most of the studies were under way by 1996, they were designed to test the effects of many program features that have been implemented by the states since the federal welfare law of 1996 was passed. The following studies were included in this analysis: Connecticut Jobs-First (abbreviated CT Jobs First in subsequent discussion), Florida's Family Transition Program (FTP), Los Angeles Jobs-First GAIN (LA GAIN), Minnesota Family Investment Program (MFIP; testing the effects of two programs, Full MFIP and MFIP Incentives Only), National Evaluation of Welfare to Work Strategies (NEWWS; testing the effects of six programs in three sites across two follow-up points), New Hope, and the Canadian Self-Sufficiency Project (SSP; testing the effects of one program at two follow-up points).

Taken together, the experimental studies tested three basic approaches that are currently used in many state welfare policies to increase the self-sufficiency of welfare recipients: earnings supplements, mandatory employment services, and time limits on welfare receipt, at times combining these approaches into a single package. Five of the 14 programs (two variants of the MFIP program, New Hope, CT Jobs First and SSP) offered generous earnings supplements designed to make work more financially rewarding by providing families with monthly cash supplements or by increasing the amount welfare recipients could keep when they went to work. Ten of the programs (LA GAIN, the six programs evaluated as part of the NEWWS evaluation, the Full MFIP program, CT Jobs First, and FTP) provided such mandatory employment services as education, training, or immediate job search in which parents were required to participate to be eligible to receive cash welfare benefits. Two of the programs under study put time limits on families' eligibility for welfare benefits (CT Jobs First and FTP), restricting eligibility to a certain number of months in a specified period. As confirmed in our empirical analysis below, all program types increased the employment of welfare-

recipient parents, but only programs with generous earnings supplements were intended to, and were successful at, increasing both parents' employment and their family incomes (Howard Bloom and Michalopoulos, 2001).

Achievement Outcomes

Children's cognitive performance or school achievement were assessed 2–5 years after parents' entry into the programs. We utilize all of the achievement measures available in the studies of the 14 programs. Achievement measures are based on parent report (available in estimates of seven programs assessed in six studies), teacher report (available in estimates of nine programs in three studies), and test scores (available in estimates of eight programs in two studies), with some studies including multiple reports per child, and two of the studies assessing children at multiple points in time. All achievement scores were standardized using study-specific control-group means and standard deviations to provide comparability. We include in our regressions dummy variables indicating the source of the achievement report. All told, our data include 30,623 achievement reports of 24,050 individual children residing in 16,456 families.

Earnings and Family Income Outcomes

For all sample members, administrative records provided data on monthly cash assistance and Food Stamp benefits as well as quarterly earnings in jobs covered by the Unemployment Insurance system. We use data covering the first two years following random assignment to construct measures of (i) average annual earnings and (ii) average annual family income, based on the sum of earnings, AFDC/TANF payments, and Food Stamp payments for both parents. Note that our family-income measure omits self-employment and informal earnings, other public transfers, private transfers, and earnings from family members other than the sample member and, if present, his or her spouse or partner.

Baseline Controls

Baseline surveys and administrative data sources provided a set of comparable pre-random-assignment parental and family control measures: whether the family received cash assistance in the two years prior to baseline; average earnings in the two years prior to baseline, measured in thousands of 2001 dollars; average earnings in the two years prior to baseline squared; whether employed in the year prior to baseline; whether the parent has a high-school degree or GED; whether the parent was a teenager at the child's birth; the marital status of the parent; the number of children in the family; the age of the youngest child in the family; and the race/ethnicity of the parent.

Analysis Strategy

We test our hypotheses by estimating ordinary least-squares (OLS) models in which treatment status (e.g., whether in the experimental or control group) varies with children's age. We accomplish this by constructing a set of interaction terms of child age group (0–2, 3–5, 6–8, 9–11, or 12–15 at baseline) with the experimental dummy. We also include dummy variables for these age groups and type of test, plus baseline controls and a measure of the time between baseline and the given achievement assessment. Huber-White methods are employed to adjust standard errors for nonindependence of multiple reports per child and multiple children per family.

Results

The first column of Table 1 presents coefficients and standard errors from regression models estimated for the achievement scores of the entire sample of children. We find statistically significant and positive experimental impacts for the two youngest groups of children: those aged 0–2 and 3–5 at baseline. These positive effects of welfare-reform strategies are small: a 0.05–0.07 standard-deviation increase in achievement scores brought about by parents' assignment to the program group.

Table 1. Regression Coefficients and Standard Errors for Models of Achievement

| Variable | Dependent Variable = Child Achievement | | |
|-------------------------------|--|------------------------------|--------------------|
| | All Programs | Earnings-Supplement Programs | All Other Programs |
| Exp × age 0–2 | 0.070* (0.031) | 0.082* (0.034) | -0.016 (0.074) |
| Exp × age 3–5 | 0.051** (0.019) | 0.080** (0.026) | 0.035 (0.026) |
| Exp × age 6–8 | -0.018 (0.030) | -0.025 (0.033) | -0.015 (0.070) |
| Exp × age 9–11 | -0.045 (0.036) | -0.043 (0.040) | -0.046 (0.082) |
| Exp × age 12–15 | -0.073 (0.051) | -0.039 (0.060) | -0.167 (0.102) |
| Age 0–2 | 0.139** (0.040) | 0.171** (0.045) | 0.182* (0.087) |
| Age 3–5 | 0.134** (0.033) | 0.147** (0.038) | 0.104 (0.071) |
| Age 6–8 | 0.177** (0.034) | 0.215** (0.038) | 0.059 (0.074) |
| Age 9–11 ^a | — | — | — |
| Age 12–15 | -0.101* (0.047) | -0.132* (0.054) | -0.019 (0.091) |
| Baseline demographic controls | yes | yes | yes |
| Type of achievement report | yes | yes | yes |
| Study dummies | yes | yes | yes |
| R ² | 0.0318 | 0.0346 | 0.0409 |
| N | 30,623 | 18,641 | 11,982 |

Note: “Exp” indicates the experimental dummy. Baseline demographic controls include follow-up length, prior earnings, prior earnings squared, prior AFDC receipt, prior years of employment, high-school degree, teen parent, marital status, number of children, and age of youngest child.

^aOmitted from the regression.

* Statistically significant at the 5-percent level.

** Statistically significant at the 1-percent level.

For the other age groups the estimated coefficients are negative, but none attains statistical significance at conventional levels.

The second and third columns present the results separately for programs with generous earnings supplements and other program models (programs with mandatory employment services and programs with time limits, both without generous supplements). Only programs with earnings supplements produce positive effects on the younger two groups of children. Apart from a marginally significant ($p = 0.101$) negative impact on youth 12–15 years old at baseline, the non-earnings-supplement programs have no impacts, either

positive or negative, on sample children. None of the t tests performed on coefficient differences between the two samples was statistically significant at conventional levels. Control variables produce predictable effects, with lower achievement associated with older ages, pre-baseline AFDC receipt, lower maternal schooling, and larger family sizes.

Table 2 presents the effects of these same welfare reform strategies on parents’ earnings and family income. Consistent with past analyses of the individual programs, the effects of welfare-reform programs on earnings are positive and significant for both earnings-supplement and non-earnings-supplement

Table 2. Regression Coefficients and Standard Errors for Models of Earnings and Family Income

| Variable | Dependent Variable = Earnings in \$1,000's | | Dependent Variable = Family Income in \$1,000's | |
|----------------------------------|---|-----------------------|--|-----------------------|
| | Earnings- Supplement Programs | All Other Programs | Earnings- Supplement Programs | All Other Programs |
| Exp × age 0–2 | 0.818** (0.175) | 0.762* (0.328) | 1.535** (0.182) | 0.235 (0.355) |
| Exp × age 3–5 | 0.900** (0.156) | 0.749** (0.126) | 1.673** (0.158) | 0.104 (0.125) |
| Exp × age 6–8 | 1.003** (0.208) | 1.093** (0.378) | 1.967** (0.211) | 0.163 (0.385) |
| Exp × age 9–11 | 0.758** (0.237) | 0.921* (0.421) | 1.602** (0.236) | 0.162 (0.426) |
| Exp × age 12–15 | 0.718* (0.349) | 1.124* (0.558) | 1.924** (0.332) | 0.235 (0.523) |
| Age 0–2 | 0.004 (0.244) | -0.211 (0.396) | 0.262 (0.240) | -0.248 (0.403) |
| Age 3–5 | -0.073 (0.209) | 0.120 (0.322) | 0.118 (0.202) | -0.144 (0.319) |
| Age 6–8 | -0.274 (0.200) | -0.118 (0.327) | -0.211 (0.190) | -0.107 (0.312) |
| Age 9–11* | — | — | — | — |
| Age 12–15 | 0.040 (0.295) | -0.415 (0.444) | 0.184 (0.274) | -0.520 (0.416) |
| Baseline demographic controls | yes | yes | yes | yes |
| Type of achievement report | no | no | no | no |
| Study dummies | yes | yes | yes | yes |
| R ² | 0.3146 | 0.2966 | 0.2184 | 0.3242 |
| N | 18,582 | 11,982 | 18,641 | 11,982 |

Note: “Exp” indicates the experimental dummy. Baseline demographic controls include follow-up length, prior earnings, prior earnings squared, prior AFDC receipt, prior years of employment, high school degree, teen parent, marital status, number of children, and age of youngest child.

* Omitted from the regression.

* Statistically significant at the 5-percent level.

** Statistically significant at the 1-percent level.

programs for all the age groups of children. Earnings impacts per year over the first two years following random assignments ranged from \$700 to \$1,100 for both sets of programs and in no case were the age-specific earnings impacts significantly different between the two sets of programs.

In contrast, family-income impacts were significantly higher for all age groups of children in the earnings-supplement programs compared with the other programs. Impacts

ranged from \$1,500 to \$2,000 per year for the family incomes of families enrolled in the earnings-supplement programs but never exceeded \$250 per year for families enrolled in the other programs.

Discussion

Our analysis suggests that welfare reform and antipoverty programs affect children and adolescents differently. When we pool across all

program models, we observe positive impacts for the two youngest age groups, children 0–5 at the beginning of these studies. Ironically, younger children were viewed as most at risk for negative impacts of maternal employment in the debate that preceded welfare reform. Moreover, we find that the positive impacts on young children's achievement appear to be confined to family-income-boosting programs offering generous earnings supplements as compared with programs with mandatory employment services and time limits but no generous supplements. That younger children's achievement appears to benefit from higher family incomes has also been documented in the nonexperimental literature (Duncan and Brooks-Gunn, 1997).

Why might effects on achievement differ for children of differing developmental periods? Our investigation of whether effects on earnings and income impacts differed by child age revealed no such differences. Both kinds of programs boosted parents' earnings, and earnings-supplement programs alone increased family income as well as earnings. Thus, the differing effects observed across the age groups are likely due to different responses on the part of children to similar changes in their parents' economic circumstances. Other work (Gennetian et al., 2002) has suggested that some negative effects are observed for adolescents in both programs with earnings supplements and those without such supplements, although the precise outcome affected might differ by program model.

REFERENCES

- Baydar, N., and Brooks-Gunn, J. "Effects of Maternal Employment and Child-care Arrangements in Infancy on Preschoolers' Cognitive and Behavioral Outcomes: Evidence from the Children of the NLSY." *Developmental Psychology*, 1991, 27, pp. 932–45.
- Becker, Gary S. *A treatise on the family*. Cambridge, MA: Harvard University Press, 1981.
- Bloom, Howard and Michalopoulos, Charles. *How welfare and work policies affect parents: A synthesis of research*. New York: Manpower Demonstration Research Corporation, 2001.
- Bronfenbrenner, U., and Morris, P. "The Ecology of Developmental Processes," in W. Damon, series ed., and R.M. Lerner, vol. ed., *Theoretical models of human development: Vol. I, Handbook of child psychology*, 5th ed. New York: Wiley, 1998, pp. 993–1028.
- Brooks-Gunn, J. "How Stressful Is the Transition to Adolescence for Girls?" in M.E. Colten and S. Gore, eds., *Adolescent stress: Causes and consequences*. Hawthorne, New York: Aldine de Gruyter, 1991, pp. 131–49.
- Brooks-Gunn, Jeanne; Han, Wen-Jui and Waldfogel, Jane. "Maternal Employment and Child Cognitive Outcomes in the First Three Years of Life." *Child Development*, July/August 2002, 73(4), pp. 1052–72.
- Coleman, James S. "Social Capital in the Creation of Human Capital." *American Journal of Sociology*, Supplement 1988, 94, pp. S95–S120.
- Duncan, Greg J. and Brooks-Gunn, Jeanne, eds., *Consequences of growing up poor*. New York: Russell Sage, 1997.
- Eccles, J.S., Midgley, C., Wigfield, A., Buchanan, C.M., Reuman, D., Flanagan, C. and Maclver, D. "Development During Adolescence: The Impact of Stage-Environment in Fit in Young Adolescents' Experiences in Schools and in Families." *American Psychologist*, 2000, 48, pp. 90–101.
- Ge, X., Lorenz, F.O., Conger, R.D., Elder, G.H., Jr., and Simons, R.L. "Trajectories of Stressful Life Events and Depressive Symptoms During Adolescence." *Developmental Psychology*, 1994, 30, pp. 467–83.
- Gennetian, Lisa; Duncan, Greg; Knox, Virginia; Vargas, Wanda; Clark-Kauffman, Elizabeth and London, Andrew. *How welfare and work policies for parents affect adolescents: A synthesis of research*. New York: Manpower Demonstration Research Corporation, 2002.
- Graber, J.A., and Brooks-Gunn, J. "Transitions and Turning Points: Navigating the Passage from Childhood Through Adolescence." *Developmental Psychology*, 1996, 32, pp. 768–76.
- Harvey, Elizabeth. "Short-Term and Long-Term Effects of Early Parental Employment on Children of the National Longitudinal Survey of Youth." *Developmental Psychology*, March 1999, 35(2), pp. 445–59.
- Lamb, Michael. "Nonparental Child Care, Context, Quality, Correlates and Consequences," in I. E. Siegel and K. A. Renninger, eds., *Handbook of*

- child psychology*, 4th Ed. New York: Wiley, 1998, pp. 73–134.
- McCall, R.B. “Nature-Nurture and the Two Realms of Development: A Proposed Integration with Respect to Mental Development.” *Child Development*, 1981, 52, pp. 1–12.
- McLoyd, Vonnie. “Children in Poverty, Development, Public Policy and Practice,” in I. E. Siegel and K. A. Renninger, eds., *Handbook of child psychology*, 4th Ed. New York: Wiley, 1998, pp. 135–210.
- Morris, Pamela; Huston, Aletha; Duncan, Greg; Crosby, Danielle and Bos, Hans. *How welfare and work policies affect children: A synthesis of research*. New York: Manpower Demonstration Research Corporation, 2001.
- Morris, Pamela and Michalopoulos, Charles. *The Self-Sufficiency Project at 36 months: Effects on children of a program that increased employment and income*. Ottawa, Canada: Social Research and Demonstration Corporation, 2000.
- Pettit, Gregory S.; Bates, John E.; Dodge, Kenneth A. and Meece, Darrell W. “The Impact of After-School Peer Contact on Early Adolescent Externalizing Problems Is Moderated by Parental Monitoring, Perceived Neighborhood Safety and Prior Adjustment.” *Child Development*, June 1999, 70(3), pp. 768–78.
- Posner, Jill K. and Vandell, Deborah Lowe. “After-School Activities and the Development of Low-Income Urban Children: A Longitudinal Study.” *Developmental Psychology*, May 1999, 35(3), pp. 868–79.
- Sampson, Robert J. and Laub, John H. “Urban Poverty and the Family Context of Delinquency: A New Look at Structure and Process in a Classic Study.” *Child Development*, April 1994, 65(2), pp. 523–40.
- Shonkoff, J.P. and Phillips, D.A. *From neurons to neighborhoods: The science of early childhood development*. Washington, D.C.: National Academy Press, 2000.
- Sroufe, L.A. “The Coherence of Individual Development: Early Care, Attachment and Subsequent Developmental Issues.” *American Psychologist*, 1979, 34, pp. 834–41.
- Waddington, C.H. *The strategy of genes*. London: Allen and Unwin, 1957.
- Waters, E., and Sroufe, L.A. “Social Competence as a Developmental Construct.” *Developmental Review*, 1983, 3, pp. 79–97.
- Yoshikawa, H. “Prevention as Cumulative Protection: Effects of Early Family Support and Education on Chronic Delinquency and Its Risks.” *Psychological Bulletin*, 1994, 115, pp. 28–54.

44 ■ Sharon Hays

Flat Broke with Children

Women in the Age of Welfare Reform

Money and Morality

A nation’s laws reflect a nation’s values. The 1996 federal law reforming welfare offered not just a statement of values to the thousands of local welfare offices across the nation, it also backed this up with something much more tangible. Welfare reform came with money. Lots of it. Every client and case-

worker in the welfare office experienced this. New social workers and employment counselors were hired. New signs were posted. New workshops were set up. In Arbordale and Sunbelt City, the two welfare offices I studied, every caseworker found a new computer on her desk. (Arbordale and Sunbelt City are pseudonyms for the two towns

where I studied the effects of welfare reform. I gave them these fictitious names to protect all the clients and caseworkers who shared with me their experiences of reform.) In small-town Arbordale, the whole office got a facelift: new carpets, new paint, a new conference room, new office chairs, and plush new office dividers. The reception area, completely remodeled with plants and posters and a children's play area, came to resemble the waiting room of an elite pediatrician's office more than the entrance to a state bureaucracy. Sunbelt City acquired new carpets, a new paint job, and new furniture as well. And all the public areas in that welfare office were newly decorated with images of nature's magnificence—glistening raindrops, majestic mountains, crashing waves, setting sun—captioned with inspirational phrases like “perseverance,” “seizing opportunities,” “determination,” “success.”

As I walked the halls of the Sunbelt City welfare office back in 1998, situated in one of the poorest and most dangerous neighborhoods of a western boom town, those scenes of nature's magnificence struck me as clearly out of place. But the inspirational messages they carried nonetheless seemed an apt symbolic representation of the new legislative strategy to train poor families in “mainstream” American values. Welfare reform, Congress had decreed, would “end the dependence of needy parents on government benefits by promoting job preparation, work, and marriage.”¹ Welfare mothers, those Sunbelt signs implied, simply needed a *push*—to get them out to work, to keep them from having children they couldn't afford to raise, to get them married and safely embedded in family life. Seizing opportunities.

States were awash in federal funds. And the economy was booming in those early years of reform. Everyone was feeling it. There was change in the air. A sense of possibilities—with just a tinge of foreboding.

The Personal Responsibility and Work Opportunity Reconciliation Act of 1996, the law that ended 61 years of poor families' entitlement to federal welfare benefits—the law that asserted and enforced a newly reformulated vision of the appropriate values of work and family life—provided all that additional funding as a way of demonstrating the depth of the nation's commitment to change in the welfare system. It provided state welfare programs with federal grants in amounts matching the peak years of national welfare caseloads (1992 to 1995)—even though those caseloads had everywhere since declined. This meant an average budget increase of 10 percent, before counting the tremendous amount of additional federal funding coming in for new childcare and welfare-to-work programs. Even though there was lots more money, most states did not pass it on to poor mothers in the form of larger welfare checks. In fact, only two states raised their benefit amounts, while two others lowered theirs at the inception of reform.

Most of the welfare caseworkers I met were optimistic about the new law, at least in the first year of its enactment. “Welfare reform is the best thing that ever happened,” was a phrase I heard frequently. A number of caseworkers, echoing popular sentiment, told me that “welfare had become a trap” and the clients had become “dependent.” Some focused on the tax money that would be saved. Others pointed out that lots of caseworkers are mothers too, and economic necessity forces them to come to work every day and leave their children in day care, so it seemed only fair that welfare mothers should be required to do the same. . . .

The purpose of this chapter is to explore the cultural norms, beliefs, and values embedded in welfare reform. While millions of dollars have been spent to track the outcomes of this legislation, and while scholars, politicians, and pundits have fiercely debated the effects of every policy contained

within it, I want to focus attention on the broad *cultural* significance of this reform effort. What does the Personal Responsibility Act tell us about the values of our society? How have its moral prescriptions been translated into concrete practices? What message does it send to the poor and to the nation? In particular, I was interested from the start in determining just what welfare reform is saying about work and family life in American society today. . . .

Invisibility and Inclusion

Invisibility was a problem experienced by the majority of the welfare mothers I encountered. Most emphasized, in one way or another, that they were not “born to welfare” (as one put it), and almost all felt fairly certain that—time limits or no time limits—they would find some way to survive. These mothers did not, in other words, tell me their stories primarily to convince listeners that they were worthy of continued welfare receipt. Many did, on the other hand, share their tales because they wanted people to understand how they came to participate in one of the most universally despised social programs in U.S. history. They had heard more than once the stereotypes labeling them as lazy, dependent, ignorant, promiscuous, and manipulative cheats. They told their stories, therefore, with the hope that they would be recognized not simply as a composite of clichés, but as whole persons. . . .

Sheila’s Downward Spiral

The spiral for Sheila began just after she finished high school. Sheila is white and was 29 years old at the time I met her in the old and notoriously dangerous housing project where she lived, not far from the Sunbelt City welfare office. From a working-class background, she was raised in a small town “with small town values.” She was engaged to be married to her high school sweetheart. The

summer after they graduated he was killed in an auto accident.

I wish that then and there I had just said, “Okay, forward,” instead of sitting and mourning and moping and weeping and thinking about what could have and should have and would have been. I should have just gone on ahead to college like I had intended. And I didn’t do it. But you know what they say, hindsight is 20–20.

She moved to Sunbelt City with her parents and took a part-time job. Less than a year later her father left her mother.

It was in February. He left a note on the kitchen table. It said, “I’m leaving you for good.” And he left the keys to the car that was not paid for. He left one month owing in rent. My mom at the time was not working, and I was still mourning my boyfriend.

Shortly after that me and my mom found jobs at a dry cleaners. We opened it in the morning, and we worked for 15 hours a day. We’d come home, go to sleep, get up, and go right back there. Six days a week. And we were doing good. We had paid up the back rent, and the car was being paid for.

And then my mother got blood clots. She almost lost her leg. And the doctor said she shouldn’t work any more. This meant we lost \$1,500 a month in income and we were trying to make it on my little \$1,200. With the medical bills and the car payments, I got a month behind on rent, and they evicted us.

We were homeless. In all the hoopla and everything, I lost my job. It just kind of dominoed. We were actually homeless, living with friends and things like that. And I mean we went hungry—we ate the throwaways from McDonald’s.

A lot of people don’t realize how close they live to being homeless. I mean, you’re just one or two paychecks away from the street. And once you hit bottom, you’re gonna hit and

you're gonna hit hard. But you have to remember that once you hit bottom you're as low as you can go. There's only one way to go, and that's up.

Unfortunately for Sheila, the downward spiral had not yet reached its lowest point. It was while she and her mother were homeless that she met the man who was to become the father of her daughter. He was, at that moment, her savior, but only a temporary one:

I met Sam, and I thought he was a very nice gentleman. Sam was living with a friend, and they took the two of us in. I found another job. We were hopeful; even though my mom still couldn't work, we were beginning to get back on our feet.

Sam had told me he was divorced. I was still young; I was 21 at the time. And, like I said, I come from a small community. Well, after we'd been together for almost a year, I got a phone call one morning—and it was his wife! So I used the money I had to put him on a bus back to his wife in Florida. And that's the last I heard of him.

A month later I found out I was pregnant. Sam still doesn't know he's got a daughter. The child support people haven't found him yet.

Putting Sam on that bus meant that pregnant Sheila and her disabled mom were homeless again. And then Sheila was raped:

For a while I wasn't sure if Sam was the father. There was the small issue of the fact that three weeks almost to the day after he left I was raped. That's a part of living on the street; that's a danger for women who live on the street.

And that's how I came to find out I was pregnant—I went to the clinic to check a few weeks after the rape. So there were two possible identities to the father. I had a hard time knowing what to do. But as soon as she was

born I saw what she looked like and I knew who she was; I knew who she belonged to. She looks too much like her dad.

Sheila was just 22 years old when she gave birth to her daughter. If things had gone as she had originally planned, she would have been starting her last year of college at that time.

Sheila first went to the welfare office in the last months of her pregnancy, hoping to get medical coverage for the birth.

I was a high-risk pregnancy all the way through my pregnancy. I went down to the welfare office and applied for medical assistance. My mother and I were still homeless. I worked a part-time job until things [with the pregnancy] got too bad. My mother had gotten a job [against her doctor's advice] and we managed, with my first welfare check, to finally get our very own small studio apartment. Just two weeks later my daughter was born.

Sheila's daughter (who was busy with her homework during most of my visit), was seven years old when I met her.

From the time she gave birth to the time of our interview, Sheila had a string of jobs. She went back to work when her daughter was just three months old. That job she described as a "really good one," where she worked her way up to a management position in a fast-food restaurant. The rest of the jobs were temporary or low-paying jobs, mainly entry-level fast foods and unskilled clerical work. She left every one. She quit the good one after more than two years, and she's still sorry about it, even though she remembers well all the time she spent agonizing over that decision. She left because the hours and the bus rides were so long that she was spending over 12 hours a day away from home, and she never had a chance to see her daughter. "My daughter was nearly three years old and she was calling my

mother ‘mom’ and calling me ‘Sheila.’ It was just too hard. I just wanted to get to know her, to do her ABCs and her 1–2–3s. I wanted to be there.”

She went back on welfare and spent almost a year getting to know her daughter before again seeking work. She spoke nostalgically of that time with her three-year-old, but she also emphasized that leaving that job was her “third big mistake,” alongside mourning rather than continuing on to college, and getting involved with a married man.

Sheila left subsequent jobs because she hurt her back, loading boxes (“the doctors say I’m not allowed to lift over 15 pounds now”), because the pay was too poor, because the jobs were only temporary (through a “temp” agency), and most recently, because her mother was diagnosed as terminally ill and needed to be cared for. Her mother suffered respiratory failure first, then a massive heart attack and, by the time I met her, Sheila was afraid to leave her alone most of the time. “That’s my greatest fear, that I’ll go to work, and I’ll be at work, and something will happen to my mom. I don’t know what to do.”

Sheila’s combined time on welfare, including the time when her child was born, between jobs, and since her mother’s illness, added up to about three and a half years, including the last year and a half since welfare reform. She’d been using the resources offered by the welfare office to train herself on computers and in accounting skills. She told me about the contacts she’d made with state agencies that might hire her, and she was feeling somewhat optimistic, though still quite worried about her mom. Her primary goal was to find an employer flexible enough to allow her to care for both her daughter and her mother.

The sheer number of tragedies in Sheila’s young life—her fiancé’s death, her father

leaving, her homelessness, her affair with a married man, her rape, her high-risk pregnancy, and now her mother’s terminal illness—testify to the unique circumstances that led her to go on welfare. But every mother I talked to had a story of hardship to tell. And every mother I met had experienced some version of the domino effect: one problem leading to another and compounding it, until too many dominoes fall and the situation becomes impossible to manage. In this, Sheila’s story represents the most prominent pattern in the road to welfare.

A second, partially hidden, pattern that Sheila shares with many welfare mothers who have children out of wedlock involves the issue of birth control. Most observers of Sheila’s life would agree that she has had some very tough luck. Yet many would also want to know why she allowed herself to get pregnant. After all, it is clear from this vantage point that the last thing Sheila needed was a child to support. She answered, “I did use birth control, but it must not have worked. I don’t know what happened.”

The truth is, it could very well be that Sheila simply did not use birth control faithfully enough, or that the methods she and her partner used were not sufficiently fool-proof. The crucial point here, however, is that a substantial number of sexually active young men and women do not use birth control faithfully enough. This fact does not appear to vary significantly by one’s race or economic status. About 50 percent of female teens are sexually active. About 70 percent of those say they used birth control the last time they had sex. But this percentage is based on self-reporting—in which case Sheila, for instance, would be included as a “yes.” And equally important, having used birth control recently is not the same thing as using it consistently. There tends to be a good deal of variation in answers to the questions “used at most recent sex,” and “used at first sex,” for example, which confirms that “recently” is

not the same as “always.”² Putting all this together, it becomes clear that a large number of sexually active young people do not use foolproof birth control every time they have sex. Sheila, in this sense, is a member of the majority.

The central factor separating poor and working-class youth from the middle and upper classes on this score is that financially privileged young women who find themselves pregnant before they are ready are more likely to get an abortion.³ I asked Sheila if she had thought about having an abortion.

Oh no, no. Well, I can't say it didn't cross my mind. But I'm a person who believes if you're gonna play, you're gonna pay. And it's not her fault. I'm the kind of person who thinks that, as soon as they have that heartbeat, which is like ten days after conception, then that's a live human being. I just couldn't do it. I love my daughter.

No matter what we might think of the consequences of this choice in the context of Sheila's life, the vast majority of Americans agree that she has a right to make this decision. And few would argue that her problem, in this instance, is a problem of bad values.

There is a third important pattern that Sheila shares with the majority of welfare mothers—the pushes toward work and the pulls toward home. The stress associated with those pushes and pulls is something welfare mothers have in common with parents of all classes and backgrounds. Sheila's desire to stay at home with her daughter is no different from all the other working moms who long to have more time to spend with their young children. Her longing for a job flexible enough to allow her to care for her child is shared by most working parents today. Her sense of regret over taking off those career-building years in order to be

with her daughter mimics all the stay-at-home mothers who worry that they will never be able to recoup their lost time in the labor market.⁴ And the fact that Sheila feels committed to staying at home with her terminally ill mother puts her in the same position as the millions of (mainly) women who care for their aging parents—many of whom suffer serious economic hardship because of it.⁵ What makes Sheila's case distinct is solely that these realities landed her in the welfare office.

No matter what the edicts of welfare reform might mean to Sheila in the coming years, and no matter what we might think of the paths she chose at the multiple crossroads of her early adulthood, the difficulties she has faced speak to much larger social problems. To the extent that the Personal Responsibility Act is our collective cultural response, it is clear that this law has done little to address the underlying causes of the strains Sheila has experienced and the choices she was forced to make. . . .

The “Success” of Welfare Reform

Most welfare mothers share the core values of most Americans. They share a concern with contemporary problems in work and family life and a commitment to finding solutions—including the overhaul of the welfare system. The trouble is, welfare reform was founded on the assumption that welfare mothers do not share American values and are, in fact, personally responsible for *undermining* our nation's moral principles. The policies and procedures instituted by welfare reform have thus been aimed at “fixing” these women. . . .

Shared Values, Symbolic Boundaries, and the Politics of Exclusion

In responding to welfare reform, the welfare mothers I met often offered a perfect mirror of the complex mix of higher values, genuine

concerns, exclusionary judgments, and cultural distortions that informed the Personal Responsibility Act. One mother, Denise, captured nearly all these elements in her response, offering the full range of the more prominent patterns I encountered and mimicking the words of welfare mothers you have heard throughout the book. A black woman with two daughters, at the time I met her Denise was recently employed at Mailboxes-R-Us for \$6.50 an hour and was making ends meet with the help of welfare reform's (time-limited) income supplement, transportation vouchers, and childcare subsidy. This is what she had to say when I asked her for her overall assessment of reform:

When I was younger, years ago, anybody could get on welfare. And I think that's what's good about welfare reform. People have to show some sort of initiative. Before, the welfare office didn't pressure you to find a job, but now they do. And I think that's a good system. They've really helped me out a lot.

Plus, I think people are sick of having to pay their tax money. They say, "Look, I am out here working, and I don't make that much money, and I have kids of my own. I'm tired of having to take care of your babies." People are getting upset and it's rightly so. I think it's rightly so.

And lots of people abuse the system. You see it every day. A lot of people that you run into and a lot of people that live in your neighborhood—I mean a lot of people do hair and get paid in cash. And I hear about these people who had children just to get a welfare check, just because they didn't want to go out and work. I've seen women that's on welfare, they're looking good and their children look poorly. I see that happening.

Some of them are lazy and don't want to work. I think that some just want to stay home with their kids. But then they should have thought about that before they had the children.

At this point in her argument, Denise had hit upon nearly all the concerns of hard-working Americans who conscientiously pay their taxes, raise their children, and struggle to make it all work. She had also hit upon nearly all the well-worn stereotypes of poor mothers—implicitly labeling them as welfare cheats, lazy couch potatoes, promiscuous breeders, and lousy parents. But Denise wasn't finished.

I think some people on welfare are being greedy—taking away from people that are homeless, people that really need the help. I mean there are truly people out there living at the Salvation Army. I hear tell that there are people who can't get in those shelters because they're so full. And I think that's the sad part about it. Those women that don't really need welfare shouldn't be taking money away from the homeless.

But there are gonna be problems. Like, there are women that want to go out there and get a job, but who's gonna watch their kids? And there are people who will still need that little extra help to pay the bills. So that's a glitch in the system. And some of these women are already pregnant, and they're already poor, and they really do need the help. I think that we have to weigh things and maybe investigate a bit more. There are a lot of people that are disabled and need welfare; there are women who have been abused. Some of those people that are in a lot of trouble, you know, their kids are gonna be the ones you see on TV, shooting up the schools and everything.⁶

I know a lot of people say that this welfare reform is a good thing—and it is really gonna help a lot of people. But in the end things are probably gonna get worse. There's gonna be more crime 'cause people can't get on welfare and they're not gonna have any money and they're gonna go out and rob people, and kill people. And it happens, it happens. So that's a problem with the system.

This same sort of ambivalence is evident in Americans' response to welfare reform. Although most are positive about reform, the majority of Americans also say that they are "very" concerned about poverty. Most additionally believe that the national standards for poverty are set too low, stating that a family of four with an income of less than \$20,000 is, in fact, "poor," even if the federal government does not label them as such. More significantly, a majority of Americans are in favor of further aid to the poor—including the expansion of job opportunities, tax credits, medical coverage, subsidies for childcare and housing, and the provision of better schools. Still, Americans worry about the government's ability to appropriately and effectively provide that aid, and many don't want to have to pay higher taxes to subsidize the poor.

Denise is also much like most Americans in that the central moral categories she uses to frame her response are work and family values, independence and commitment to others, self-sufficiency and concern for the common good. Women should take the "initiative," they should work, they should not rely on the help of others, they should support their own children, they should think twice before they give birth to children they cannot afford to raise. At the same time, people should not be "greedy," they should care for those who are more vulnerable than themselves, and they should consider the impact of their actions on the nation as a whole. All this makes perfect sense, and all this resonates perfectly with our nation's values. The trouble is that managing these commitments is hard enough if you have a spouse, a house in the suburbs, two cars in the garage, good health insurance, reliable childcare, a willingness to make compromises, a great deal of determination, empathy, and energy, and a household income of \$60,000. The more items on this list that you lack, the tougher it becomes to live up to

this demanding system of values. Denise, like most Americans, implicitly understands these "glitches." Yet her reasoning becomes a bit cloudy at this point—in large measure, I would argue, because of the loophole provided by the final significant element in her response to welfare reform.

It is hard to miss that Denise's support for the Personal Responsibility Act is predicated on the construction of a moral distinction between herself and all those "other" bad welfare mothers who fail to live up to social standards. Denise is making use of what Michèle Lamont has called "symbolic boundaries" to develop an implicit hierarchy of social worth. Like most people who use this strategy, she is not simply engaging in a mean-spirited attack on others or a self-interested attempt to highlight her own virtues. These symbolic boundaries also allow her to positively affirm shared values and specify the proper way to live one's life.⁷

Yet, given that many observers consider Denise herself a member of the deviant group she describes, the fact that she and other welfare mothers persist in this technique is curious. It testifies not just to the power and ubiquity of boundary making as a social strategy, it also speaks to the power and ubiquity of the demonization of poor single mothers. When welfare mothers distinguish themselves from those other "bad" women, they are calling on widely disseminated negative images of welfare mothers. These images seem to match all those strangers, those loud neighbors, those people who appear to spend their lives hanging out on street corners. The lives of the women they actually know, on the other hand, seem much more complex, their actions more understandable, their futures more redeemable.

The demonization of welfare mothers and the dichotomy between "us" and "them" can thus provide a dividing line that allows Denise and other Americans to say, if some welfare mothers can't make it, it's not be-

cause the problems they encounter in trying to manage work and family and still keep their heads above water are that bad or that widespread; it's because they didn't try hard enough or weren't good enough. Symbolic boundaries thus become *exclusionary* boundaries—simultaneously offering a means to affirm shared values and a means to think of “outsiders” in terms of individual blame. The obvious problem, in Denise's case, is that her own logic might ultimately leave her as one of the “accused.” In broader terms, this exclusionary process means that all those Americans who are suffering from childcare woes, second shifts, inadequate health insurance, precarious jobs, unmanageable debt, and unstable communities are left to feel that their problems are *personal* problems for which no public solutions can be found. . . .

Winners and Losers

The extent to which the facts about the declining welfare rolls are read as a success ultimately depends on one's primary goals. If the goal of reform was solely to trim the rolls, then it has surely succeeded. If the goal was to place more single mothers in jobs regardless of wages, that goal has been met. If we sought to ensure that more welfare mothers would face a double shift of paid work and childcare, placing them on an “equal” footing with their middle-class counterparts, then some celebrations are in order. If the aim was to ensure that poor men are prosecuted for failure to pay child support, then welfare reform has been relatively effective. If the goal was to make low-income single mothers more likely to seek out the help of men, no matter what the costs, there is some (inconclusive) evidence that this strategy may be working.⁸ If the goal was to decrease poverty overall, there is no indication that anything but the cycle of the economy has had an impact. Beyond this, the answers are more complicated.

Thinking about losers, one can start with the families who have left welfare. One-half are sometimes without enough money to buy food. One-third have to cut the size of meals. Almost half find themselves unable to pay their rent or utility bills. Many more families are turning to locally funded services, food banks, churches, and other charities for aid. Many of those charities are already overburdened. In some locales, homeless shelters and housing assistance programs are closing their doors to new customers, food banks are running out of food, and other charities are being forced to tighten their eligibility requirements.⁹

Among the former welfare families who are now living with little or no measurable income, will those charities be enough? At ground level, Nancy, the supervisor in Arbordale's welfare office, told me more than once that she was deeply concerned about these families, particularly the children. Melissa, the supervisor in Sunbelt City, on the other hand, repeatedly responded to my questions regarding the fate of former welfare recipients with the simple statement, “They have other resources.” Melissa was referring not only to all those (overloaded) charities, but also to all the boyfriends and family members who could help in paying the bills, and to all those unreported or underreported side jobs (doing hair, cleaning houses, caring for other people's children, selling sex or drugs).¹⁰ Between these two welfare supervisors, both of whom have spent many years working with poor mothers, who is right? And what about Denise, who both agreed with Melissa that many welfare mothers didn't *really* need the help, and who predicted that welfare reform would result in frightening hardship, including a rise in crime?

Consider the “other resources” available to the women I have introduced in this chapter. In the case of Sheila, the Sunbelt mother who was caring for her seven-year-old daughter

and her terminally ill mother, the three of them might be able to survive somehow on her mom's disability check (about \$550 per month) with the help of food stamps and local charities. If worse came to worst, she might be able to find some work on the graveyard shift so that she wouldn't have to leave her mom and daughter alone during the day (but she would be faced with leaving them alone at night in that very dangerous housing project).

Most welfare mothers *do* have other resources. Yet many of those resources are only temporary, and many are, at best, inadequate. Most will likely add greater instability and uncertainty to the lives of these families. And nearly all these resources have their own price tags—practical, emotional, moral, and social. . . .

The real winners in the story of welfare reform are all the restaurant, hotel, retail, and food service chains, and all the corporations, manufacturers, and small business owners across America who employ low-wage workers. These owners (and their stockholders) benefit not just from the availability of millions of poor women desperate to find work and willing to accept the lowest wages and the worst working conditions, they benefit not just from the additional availability of all those now more-desperate poor men, they also benefit because all this desperation creates more profitable labor market conditions overall. Welfare reform helps to convince all low-wage workers that they can be easily displaced by former welfare recipients and therefore makes them less likely to complain, change jobs, join unions, or demand higher wages. The logic of reform also means that low-wage employers can rest assured, for the moment at least, that no one will be calling into question the fact that their policies are less than family friendly and their workers are unable to support their children on the wages they take home.¹¹

On a superficial level, the “end of welfare” appears to hold in place the symbolic mes-

sages that work is better than welfare and marriage is better than single parenthood. But by no stretch of the imagination could one argue that welfare reform brings with it anything resembling the triumph of “family values.” And the practical reality of most low-wage employment no more offers “independence” and self-sufficiency to former welfare recipients than it does to all the middle-class teenagers who spend their summers working in fast-food restaurants and retail chains.

NOTES

1. U.S. Congress (1996, PL104–193, Title I, Section 401).
2. Terry and Manlove (1999); see also Luker (1996) and Waller (1999).
3. Luker (1996), Waller (1999), and Ventura et al. (2000).
4. See, for instance, Hays (1996), Hochschild (1989, 1997).
5. See Arno et al. (1999) and Harvard School of Public Health (2000) on the number of women caring for disabled or aged family members.
6. Denise is referring to then-prominent news stories on the Columbine school shooting and the other school shootings that followed.
7. On symbolic boundaries, see Lamont (1992, 2000).
8. See Sorenson and Zibman (2000); Cherlin and Fomby (2002).
9. See Loprest (1999), Boushev and Gunderson (2001), Sherman et al. (1998), National Campaign for Jobs and Income Support (2001A, 2001B). A study of major U.S. cities found that from 2000 to 2001, requests for food had increased by 23 percent and requests for emergency housing were up by 13 percent (U.S. Conference of Mayors 2001).
10. See Edin and Lein (1997).
11. See, for instance, Newman (1988, 1999), Piven (1999), Blau (1999), Edin and Lein (1997), Ehrenreich (2001).

REFERENCES

- Arno, Peter, Carol Levine, and Margaret Memmott. 1999. “The Economic Value of Informal Caregiving.” *Health Affairs* 18: 182–188.
- Boushev, Heather and Bethney Gunderson. 2001. *When Work Just Isn't Enough: Measuring Hardships*

- Faced by Families after Moving from Welfare to Work.* Washington, DC: Economic Policy Institute.
- Cherlin, Andrew J. and Paula Fomby. 2002. *A Closer Look at Changes in Children's Living Arrangements.* Welfare, Children, and Families: A Three-City Study, Working Paper 02-01. Baltimore, MD: Johns Hopkins University.
- Edin, Kathryn and Laura Lein. 1997. *Making Ends Meet: How Single Mothers Survive Welfare and Low-Wage Work.* New York: Russell Sage Foundation.
- Ehrenreich, Barbara. 2001. *Nickel and Dimed: On (Not) Getting By in America.* New York: Metropolitan Books.
- Harvard School of Public Health. 2000. *Caregiving by Women: A Disproportionately Large Load.* Project on Global Working Families. Cambridge, MA: President and Fellows of Harvard College.
- Hays, Sharon. 1996. *The Cultural Contradictions of Motherhood.* New Haven, CT: Yale University Press.
- Hochschild, Arlie Russell with Anne Machung. 1989. *The Second Shift: Working Parents and the Revolution at Home.* New York: Viking.
- Hochschild, Arlie Russell. 1997. *The Time Bind: When Work Becomes Home and Home Becomes Work.* New York: Metropolitan Books.
- Lamont, Michèle. 1992. *Money, Morals, and Manners.* Chicago: University of Chicago Press.
- Lamont, Michèle. 2000. *The Dignity of Working Men: Morality and the Boundaries of Race, Class, and Immigration.* New York: Russell Sage Foundation.
- Loprest, Pamela. 1999. *Families Who Left Welfare: Who Are They and How Are They Doing?* Washington, DC: Urban Institute.
- Luker, Kristin. 1996. *Dubious Conceptions: The Politics of Teenage Pregnancy.* Cambridge, MA: Harvard University Press.
- National Campaign for Jobs and Income Support. 2001A. *A Recession Like No Other: New Analysis Finds Safety Net in Tatters as Economic Slump Deepens.* Washington, DC: National Campaign for Jobs and Income Support.
- National Campaign for Jobs and Income Support. 2001B. *Leaving Welfare, Left Behind: Employment Status, Income, and Well-Being of Former TANF Recipients.* Washington, DC: National Campaign for Jobs and Income Support.
- Newman, Katherine S. 1988. *Falling from Grace.* New York: Free Press.
- Newman, Katherine S. 1999. *No Shame in My Game: The Working Poor in the Inner City.* New York: Vintage Books.
- Piven, Frances Fox. 1999. "Welfare and Work," pp. 83-99 in *Whose Welfare?* edited by Gwendolyn Mink. Ithaca, NY: Cornell University Press.
- Sherman, Arloc, Cheryl Amey, Barbara Duffield, Nancy Ebb, and Deborah Weinstein. 1998. *Welfare to What? Early Findings on Family Hardship and Well-Being.* National Coalition for the Homeless. Washington, DC: Children's Defense Fund.
- Sorensen, Elaine and Chava Zibman. 2000. *Child Support Offers Some Protection Against Poverty.* Assessing the New Federalism, Series B, No. B-10. Washington, DC: Urban Institute.
- Terry, Elizabeth and Jennifer Manlove. 1999. *Trends in Sexual Activity and Contraceptive Use Among Teens.* Washington, DC: Child Trends.
- U.S. Conference of Mayors. 2001. *A Status Report on Hunger and Homelessness in America's Cities: A 27-City Survey.* Washington, DC: Conference of Mayors.
- U.S. Congress. 1996. *Personal Responsibility and Work Opportunity Reconciliation Act of 1996.* Public Law 104-193, H.R. 3734.
- Ventura, Stephanie and Christine A. Bachrach. 2000. *Nonmarital Childbearing in the United States, 1940-99.* U.S. Department of Health and Human Services, National Vital Statistics Reports 48 (16). Washington, DC: U.S. Government Printing Office.
- Waller, Maureen R. 1999. "Meanings and Motives in New Family Stories: The Separation of Reproduction and Marriage Among Low-Income Black and White Parents," pp. 182-218 in *The Cultural Territories of Race: Black and White Boundaries*, edited by Michèle Lamont. Chicago: University of Chicago Press.

PART V

Generating Inequality

Social Mobility

■ CLASS MOBILITY ■

45 ■ Ralph H. Turner

Sponsored and Contest Mobility and the School System

This paper suggests a framework for relating certain differences between American and English systems of education to the prevailing norms of upward mobility in each country. Others have noted the tendency of educational systems to support prevailing schemes of stratification, but this discussion concerns specifically the manner in which the *accepted mode of upward mobility* shapes the school system directly and indirectly through its effects on the values which implement social control.

Two ideal-typical normative patterns of upward mobility are described and their ramifications in the general patterns of stratification and social control are suggested. In addition to showing relationships among a number of differences between American and English schooling, the ideal-types have broader implications than those developed in this paper: they suggest a major dimension of stratification which might be profitably incorporated into a variety of studies in social class; and they readily can be applied in further comparisons between other countries.

The Nature of Organizing Norms

Many investigators have concerned themselves with rates of upward mobility in specific countries or internationally,¹ and with the manner in which school systems facilitate or impede such mobility.² But preoccupation with the *extent* of mobility has precluded equal attention to the predominant *modes* of mobility. The central assumption underlying this paper is that within a formally open class system that provides for mass education the organizing folk norm which defines the accepted mode of upward mobility is a crucial factor in shaping the school system, and may be even more crucial than the extent of upward mobility. In England and the United States there appear to be different organizing folk norms, here termed *sponsored mobility* and *contest mobility*, respectively. *Contest* mobility is a system in which elite³ status is the prize in an open contest and is taken by the aspirants' own efforts. While the "contest" is governed by some rules of fair play, the contestants have wide latitude in the strategies they may employ. Since the "prize" of suc-

Ralph H. Turner. "Sponsored and Contest Mobility and the School System," in *American Sociological Review* 25 (December 1960), pp. 855–856, 859–862, 865–866.

cessful upward mobility is not in the hands of an established elite to give out, the latter cannot determine who shall attain it and who shall not. Under *sponsored* mobility elite recruits are chosen by the established elite or their agents, and elite status is *given* on the basis of some criterion of supposed merit and cannot be *taken* by any amount of effort or strategy. Upward mobility is like entry into a private club where each candidate must be “sponsored” by one or more of the members. Ultimately the members grant or deny upward mobility on the basis of whether they judge the candidate to have those qualities they wish to see in fellow members. . . .

Social Control and the Two Norms

Every society must cope with the problem of maintaining loyalty to its social system and does so in part through norms and values, only some of which vary by class position. Norms and values especially prevalent within a given class must direct behavior into channels that support the total system, while those that transcend strata must support the general class differential. The way in which upward mobility takes place determines in part the kinds of norms and values that serve the indicated purposes of social control in each class and throughout the society.

The most conspicuous control problem is that of ensuring loyalty in the disadvantaged classes toward a system in which their members receive less than a proportional share of society's goods. In a system of contest mobility this is accomplished by a combination of futuristic orientation, the norm of ambition, and a general sense of fellowship with the elite. Each individual is encouraged to think of himself as competing for an elite position so that loyalty to the system and conventional attitudes are cultivated in the process of preparation for this possibility. It is essential that this futuristic orientation be kept alive by delaying a sense of final irreparable failure to

reach elite status until attitudes are well established. By thinking of himself in the successful future the elite aspirant forms considerable identification with elitists, and evidence that they are merely ordinary human beings like himself helps to reinforce this identification as well as to keep alive the conviction that he himself may someday succeed in like manner. To forestall rebellion among the disadvantaged majority, then, a contest system must avoid absolute points of selection for mobility and immobility and must delay clear recognition of the realities of the situation until the individual is too committed to the system to change radically. A futuristic orientation cannot, of course, be inculcated successfully in all members of lower strata, but sufficient internalization of a norm of ambition tends to leave the unambitious as individual deviants and to forestall the latter's formation of a genuine subcultural group able to offer collective threat to the established system. Where this kind of control system operates rather effectively it is notable that organized or gang deviancy is more likely to take the form of an attack upon the conventional or moral order rather than upon the class system itself. Thus the United States has its “beatniks”⁴ who repudiate ambition and most worldly values and its delinquent and criminal gangs who try to evade the limitations imposed by conventional means,⁵ but very few active revolutionaries.

These social controls are inappropriate in a system of sponsorship since the elite recruits are chosen from above. The principal threat to the system would lie in the existence of a strong group the members of whom sought to *take* elite positions themselves. Control under this system is maintained by training the “masses” to regard themselves as relatively incompetent to manage society, by restricting access to the skills and manners of the elite, and by cultivating belief in the superior competence of the elite. The earlier that selection of the elite recruits is made the sooner

others can be taught to accept their inferiority and to make “realistic” rather than fantasy plans. Early selection prevents raising the hopes of large numbers of people who might otherwise become the discontented leaders of a class challenging the sovereignty of the established elite. If it is assumed that the difference in competence between masses and elite is seldom so great as to support the usual differences in the advantages accruing to each,⁶ then the differences must be artificially augmented by discouraging acquisition of elite skills by the masses. Thus a sense of mystery about the elite is a common device for supporting in the masses the illusion of a much greater hiatus of competence than in fact exists.

While elitists are unlikely to reject a system that benefits them, they must still be restrained from taking such advantage of their favorable situation as to jeopardize the entire elite. Under the sponsorship system the elite recruits—who are selected early, freed from the strain of competitive struggle, and kept under close supervision—may be thoroughly indoctrinated in elite culture. A norm of paternalism toward inferiors may be inculcated, a heightened sensitivity to the good opinion of fellow elitists and elite recruits may be cultivated, and the appreciation of the more complex forms of aesthetic, literary, intellectual, and sporting activities may be taught. Norms of courtesy and altruism easily can be maintained under sponsorship since elite recruits are not required to compete for their standing and since the elite may deny high standing to those who strive for position by “unseemly” methods. The system of sponsorship provides an almost perfect setting for the development of an elite culture characterized by a sense of responsibility for “inferiors” and for preservation of the “finer things” of life.

Elite control in the contest system is more difficult since there is no controlled induction and apprenticeship. The principal regu-

lation seems to lie in the insecurity of elite position. In a sense there is no “final arrival” because each person may be displaced by newcomers throughout his life. The limited control of high standing from above prevents the clear delimitation of levels in the class system, so that success itself becomes relative: each success, rather than an accomplishment, serves to qualify the participant for competition at the next higher level.⁷ The restraints upon the behavior of a person of high standing, therefore, are principally those applicable to a contestant who must not risk the “ganging up” of other contestants, and who must pay some attention to the masses who are frequently in a position to impose penalties upon him. But any special norm of paternalism is hard to establish since there is no dependable procedure for examining the means by which one achieves elite credentials. While mass esteem is an effective brake upon over-exploitation of position, it rewards scrupulously ethical and altruistic behavior much less than evidence of fellow-feeling with the masses themselves.

Under both systems, unscrupulous or disreputable persons may become or remain members of the elite, but for different reasons. In contest mobility, popular tolerance of a little craftiness in the successful newcomer, together with the fact that he does not have to undergo the close scrutiny of the old elite, leaves considerable leeway for unscrupulous success. In sponsored mobility, the unpromising recruit reflects unfavorably on the judgments of his sponsors and threatens the myth of elite omniscience; consequently he may be tolerated and others may “cover up” for his deficiencies in order to protect the unified front of the elite to the outer world.

Certain of the general values and norms of any society reflect emulation of elite values by the masses. Under sponsored mobility, a good deal of the protective attitudes toward and interest in classical subjects percolates to

the masses. Under contest mobility, however, there is not the same degree of homogeneity of moral, aesthetic, and intellectual values to be emulated, so that the conspicuous attribute of the elite is its high level of material consumption—emulation itself follows this course. There is neither effective incentive nor punishment for the elitist who fails to interest himself in promoting the arts or literary excellence, or who continues to maintain the vulgar manners and mode of speech of his class origin. The elite has relatively less power and the masses relatively more power to punish or reward a man for his adoption or disregard of any special elite culture. The great importance of accent and of grammatical excellence in the attainment of high status in England as contrasted with the twangs and drawls and grammatical ineptitude among American elites is the most striking example of this difference. In a contest system, the class order does not function to support the *quality* of aesthetic, literary, and intellectual activities; only those well versed in such matters are qualified to distinguish authentic products from cheap imitations. Unless those who claim superiority in these areas are forced to submit their credentials to the elite for evaluation, poor quality is often honored equally with high quality and class prestige does not serve to maintain an effective norm of high quality.

This is not to imply that there are no groups in a “contest” society devoted to the protection and fostering of high standards in art, music, literature, and intellectual pursuits, but that such standards lack the support of the class system which is frequently found when sponsored mobility prevails. In California, the selection by official welcoming committees of a torch singer to entertain a visiting king and queen and “can-can” dancers to entertain Mr. Khrushchev illustrates how American elites can assume that high prestige and popular taste go together.

Formal Education

Returning to the conception of an organizing ideal norm, we assume that to the extent to which one such norm of upward mobility is prevalent in a society there are constant strains to shape the educational system into conformity with that norm. These strains operate in two fashions: directly, by blinding people to alternatives and coloring their judgments of successful and unsuccessful solutions to recurring educational problems; indirectly, through the functional interrelationships between school systems and the class structure, systems of social control, and other features of the social structure which are neglected in this paper.

The most obvious application of the distinction between sponsored and contest mobility norms affords a partial explanation for the different policies of student selection in the English and American secondary schools. Although American high school students follow different courses of study and a few attend specialized schools, a major educational preoccupation has been to avoid any sharp social separation between the superior and inferior students and to keep the channels of movement between courses of study as open as possible. Recent criticisms of the way in which superior students may be thereby held back in their development usually are nevertheless qualified by the insistence that these students must not be withdrawn from the mainstream of student life.⁸ Such segregation offends the sense of fairness implicit in the contest norm and also arouses the fear that the elite and future elite will lose their sense of fellowship with the masses. Perhaps the most important point, however, is that schooling is presented as an opportunity, and making use of it depends primarily on the student’s own initiative and enterprise.

The English system has undergone a succession of liberalizing changes during this century, but all of them have retained the attempt

to sort out early in the educational program the promising from the unpromising so that the former may be segregated and given a special form of training to fit them for higher standing in their adult years. Under the Education Act of 1944, a minority of students has been selected each year by means of a battery of examinations popularly known as “eleven plus,” supplemented in varying degrees by grade school records and personal interviews, for admission to grammar schools.⁹ The remaining students attend secondary modern or technical schools in which the opportunities to prepare for college or to train for the more prestigious occupations are minimal. The grammar schools supply what by comparative standards is a high quality of college preparatory education. Of course, such a scheme embodies the logic of sponsorship, with early selection of those destined for middle-class and higher-status occupations, and specialized training to prepare each group for its destined class position. This plan facilitates considerable mobility, and recent research reveals surprisingly little bias against children from manual laboring-class families in the selection for grammar school, when related to measured intelligence.¹⁰ It is altogether possible that adequate comparative study would show a closer correlation of school success with measured intelligence and a lesser correlation between school success and family background in England than in the United States. While selection of superior students for mobility opportunity is probably more efficient under such a system, the obstacles for persons not so selected of “making the grade” on the basis of their own initiative or enterprise are probably correspondingly greater. . . .

Effects of Mobility on Personality

Brief note may be made of the importance of the distinction between sponsored and contest mobility with relation to the sup-

posed effects of upward mobility on personality development. Not a great deal is yet known about the “mobile personality” nor about the specific features of importance to the personality in the mobility experience.¹¹ However, today three aspects of this experience are most frequently stressed: first, the stress or tension involved in striving for status higher than that of others under more difficult conditions than they; second, the complication of interpersonal relations introduced by the necessity to abandon lower-level friends in favor of uncertain acceptance into higher-level circles; third, the problem of working out an adequate personal scheme of values in the face of movement between classes marked by somewhat variant or even contradictory value systems.¹² The impact of each of these three mobility problems, it is suggested, differs depending upon whether the pattern is that of the contest or of sponsorship.

Under the sponsorship system, recruits are selected early, segregated from their class peers, grouped with other recruits and with youth from the class to which they are moving, and trained specifically for membership in this class. Since the selection is made early, the mobility experience should be relatively free from the strain that comes with a series of elimination tests and long-extended uncertainty of success. The segregation and the integrated group life of the “public” school or grammar school should help to clarify the mobile person’s social ties. (One investigator failed to discover clique formation along lines of social class in a sociometric study of a number of grammar schools.)¹³ The problem of a system of values may be largely met when the elite recruit is taken from his parents and peers to be placed in a boarding school, though it may be less well clarified for the grammar school boy who returns each evening to his working-class family. Undoubtedly this latter limitation has something to do with the observed fail-

ure of working-class boys to continue through the last years of grammar school and into the universities.¹⁴ In general, then, the factors stressed as affecting personality formation among the upwardly mobile probably are rather specific to the contest system, or to an incompletely functioning sponsorship system.

NOTES

This is an expanded version of a paper presented at the Fourth World Congress of Sociology, 1959, and abstracted in the *Transactions* of the Congress. Special indebtedness should be expressed to Jean Floud and Hilde Himmelweit for helping to acquaint the author with the English school system.

1. A comprehensive summary of such studies appears in Seymour M. Lipset and Reinhard Bendix, *Social Mobility in Industrial Society*, Berkeley and Los Angeles: University of California Press, 1959.

2. Cf. C. A. Anderson, "The Social Status of University Students in Relation to Type of Economy: An International Comparison," *Transactions of the Third World Congress of Sociology*, London, 1956, Vol. V, pp. 51–63; J. E. Floud, *Social Class and Educational Opportunity*, London: Heinemann, 1956; W. L. Warner, R. J. Havighurst, and M. B. Loeb, *Who Shall Be Educated?* New York: Harper, 1944.

3. Reference is made throughout the paper to "elite" and "masses." The generalizations, however, are intended to apply throughout the stratification continuum to relations between members of a given class and the class or classes above it. Statements about mobility are intended in general to apply to mobility from manual to middle-class levels, lower-middle to upper-middle class, and so on, as well as into the strictly elite groups. The simplified expressions avoid the repeated use of cumbersome and involved statements which might otherwise be required.

4. See, e.g., Lawrence Lipton, *The Holy Barbarians*, New York: Messner, 1959.

5. Cf. Albert K. Cohen, *Delinquent Boys: The Culture of the Gang*, Glencoe, Ill.: Free Press, 1955.

6. D. V. Glass, editor, *Social Mobility in Britain*, Glencoe, Ill.: Free Press, 1954, pp. 144–145, reports studies showing only small variations in intelligence between occupational levels.

7. Geoffrey Gorer, *The American People*, New York: Norton, 1948, pp. 172–187.

8. See, e.g., *Los Angeles Times*, May 4, 1959, Part I, p. 24.

9. The nature and operation of the "eleven plus" system are fully reviewed in a report by a committee of the British Psychological Society and in a report of extensive research into the adequacy of selection methods. See P. E. Vernon, editor, *Secondary School Selection: A British Psychological Inquiry*, London: Methuen, 1957; and Alfred Yates and D. A. Pidgeon, *Admission to Grammar Schools*, London: Newnes Educational Publishing Co., 1957.

10. J. E. Floud, A. H. Halsey, and F. M. Martin, *Social Class and Educational Opportunity*, London: Heinemann, 1956.

11. Cf. Lipset and Bendix, *op. cit.*, pp. 250 ff.

12. See, e.g., August B. Hollingshead and Frederick C. Redlich, *Social Class and Mental Illness*, New York: Wiley, 1958; W. Lloyd Warner and James C. Abegglen, *Big Business Leaders in America*, New York: Harper, 1955; Warner *et al.*, *Who Shall Be Educated?*, *op. cit.*; Peter M. Blau, "Social Mobility and Interpersonal Relations," *American Sociological Review*, 21 (June, 1956), pp. 290–300.

13. A. N. Oppenheim, "Social Status and Clique Formation among Grammar School Boys," *British Journal of Sociology*, 6 (September, 1955), pp. 228–245. Oppenheim's findings may be compared with A. B. Hollingshead, *Elmtown's Youth*, New York: Wiley, 1949, pp. 204–242. See also Joseph A. Kahl, *The American Class Structure*, New York: Rinehart, 1957, pp. 129–138.

14. Floud *et al.*, *op. cit.*, pp. 115 ff.

46 ■ David L. Featherman and Robert M. Hauser

A Refined Model of Occupational Mobility

In this [article] we describe and apply a log-linear model of the mobility table. . . . The model permits us to locate groups or clusters of cells in the classification that share similar chances of mobility or immobility, freed of the confounding influences of the relative numbers of men in each origin or destination category and of changes in those relative numbers between origin and destination distributions.

By modeling the mobility table in this way we obtain new insights into the process of mobility, changes in that process, and the interactions of the mobility process with changes in the occupational structure within one mobility classification or between two or more mobility classifications. For example, we take a fresh look at the differing tendencies toward immobility in the several occupational strata, at the existence of “class” boundaries limiting certain types of mobility, at differences in upward and downward exchanges between occupational strata, and at differences among strata in the dispersion of recruitment and supply. In these purposes our analysis parallels Blau and Duncan’s treatment of manpower flows (1967: Chap. 2; also, see Blau, 1965).

Several sociologists have recently drawn attention to relationships between occupational mobility and class formation, for example, Giddens (1973), Parkin (1971), and Westergaard and Resler (1975). Goldthorpe and Llewellyn (1977) have critically reviewed these and related works in light of British

mobility data collected in 1972. It would be easy to identify our present analytic interests with those of the class theorists, but we think such an inference unwarranted.

Although we are attempting a description of the mobility regime that is free of the distributions of occupational origins and destinations, we believe with Goldthorpe and Llewellyn that the class theorists are attempting to interpret what [might be] termed the gross flows of manpower. For the American case we have already described those flows [see Featherman and Hauser (1978: Chapter 3)], and our interest now centers on the net or underlying patterns of association in the mobility table.

We have approached the mobility table without strong theoretical presuppositions about affinities among occupational strata. Like Blau and Duncan, we have worked inductively, but our more refined analytic tools have led to substantively different conclusions than theirs about the major features of the mobility process in the United States.

Some readers may find the following discussion excessively technical, but we have tried to minimize the presentation of methodological detail. We have tried to avoid describing the methods by which empirical specifications of the mobility table may be explored, although we believe these are interesting in their own right. We have focused on the rationale and interpretation of our model, including comparisons with

other ways of looking at the mobility table that seem likely to elucidate the properties of the model.

Mobility Models

The record of sociological mobility studies is paralleled by a history of statistical analysis in which occupational mobility has often served as stimulus, object, or illustration of statistical ideas (for example, see Pearson, 1904; Chessa, 1911; Rogoff, 1953; Glass, 1954; Goodman, 1961, 1968, 1969a, 1972c; Tyree, 1973; White, 1963, 1970a; Singer and Spilerman, 1976). Indeed, it is consistent with the historical pattern that sociologists were introduced to the method of path analysis primarily by way of its successful application in studies of occupational mobility (Duncan and Hodge, 1963; Blau and Duncan, 1967). Devices for the statistical analysis of mobility data range from simple descriptive measures to complex analytic schemes. We make no systematic effort to review these measures and models, for there are several recent and comprehensive reviews (Boudon, 1973; Pullum, 1975; Bibby, 1975). We focus almost exclusively on multiplicative (loglinear) representations of the occupational mobility table. In so doing we do not intend to suggest that other methods and approaches are inferior, but to exploit features of the loglinear model that seem interesting and fruitful. . . .

In a series of papers, Goodman (1963, 1965, 1968, 1969a, 1969b, 1972c) developed and explicated methods for the analysis of contingency tables (including mobility tables) in which the significant interactions were localized in specified cells or sets of cells in the table (also, see Pullum, 1975). For example, in the case of highly aggregated (3×3 or 5×5) mobility tables Goodman showed that most of the interaction pertained to cells on or near the main diagonal (when the occupation categories of origin and destination were listed in order of increasing status). White (1963, 1970b) has made essentially the same sugges-

tion, but some aspects of his models and methods are less appealing. Goodman (1965, 1969a) proposed that the analyst ignore or “blank out” those cells where interaction was greatest (where frequencies were thought to be especially dense or especially sparse) and attempt to fit a modified model of statistical independence, termed “quasi-independence,” to the remaining frequencies in the table. In the case where only diagonal cells were blanked out in a mobility table, Goodman called the model one of “quasi-perfect mobility,” after the term “perfect mobility,” which had earlier been applied to the model of statistical independence in a mobility table. For an early application of this model to a large (17×17) table see Blau and Duncan (1967: 64–67). Goodman (1965, 1968, 1969a) noted that quasi-independence might hold over all cells in a table whose entries were not ignored, or it might hold within, but not between certain subsets of cells whose entries were not ignored. . . .

Models of quasi-independence have provided important insights into the structure of mobility tables. Aside from Goodman’s expository papers, they have been applied in cross-national, interurban, and cross-temporal analyses (Iutaka *et al.*, 1975; Featherman *et al.*, 1975; Pullum, 1975; Hauser *et al.*, 1975; Ramsøy, 1977; Goldthorpe *et al.*, 1978). Goodman (1969a) also has shown how related ideas may be supplied to test any specific hypothesis about the pattern of association in a mobility table.

At the same time the application of quasi-independence models in mobility analysis has been less than satisfying in some ways. Even where large numbers of cells are blocked, quasi-independence models do not fit large tables very well (Pullum, 1975; Hauser *et al.*, 1975). That is, when mobility data are not highly aggregated, it appears that association is not limited to the small number of cells on or near the main diagonal. The larger the number of entries blocked (or fitted exactly) before a good fit is obtained,

the less substantively appealing is the model of quasi-independence. Moreover, by treating departures from quasi-independence in the blocked or ignored cells as parameters or indices of mobility and departures in the unblocked cells as error, the quasi-independence model attaches too much theoretical importance to occupational inheritance (Hope, 1976). Of course, occupational inheritance is always defined by reference to a given classification of occupations, and the problem is exacerbated by the fact that the model of quasi-independence fits best when the mobility table is based on broad occupation groups. The model is of greatest validity in the measurement of immobility in classifications where the concept of occupational inheritance becomes vague.

The focus on fit on or near the main diagonal follows a traditional sociological interest in occupational inheritance, but it also draws our attention away from other aspects of association in the table. For example, one might hypothesize that certain types of mobility are as prevalent as other types of mobility or immobility. More generally, one might wish to construct a parametric model of mobility and immobility for the full table that would recognize the somewhat arbitrary character of occupational inheritance and the possible gradations of association throughout the table.

Goodman's (1972c) general multiplicative model of mobility tables and other cross-classifications substantially advanced the sophistication and precision of mobility analysis. For example, Goodman proposed and applied to the classic British and Danish mobility data a number of alternative specifications, all but two of which—the simple independence model and that of quasi-perfect mobility—assumed ordinality in the occupational categories. The models incorporated combinations of parameters for upward and downward mobility, for the number of boundaries crossed, and for barriers to crossing particular categorical boundaries. Many of

these models—as well as problems in comparing their goodness of fit—are reviewed by Bishop *et al.* (1975: Chaps. 5, 8, 9), and some of the same models are discussed by Haberman (1974: Chap. 6). Applying Goodman's (1972c) general model we take a slightly different approach in developing models of the mobility table. Elsewhere, Hauser (1978) has applied this approach in an analysis of the classic British mobility table, and Baron (1977) has used it in a re-analysis of Rogoff's (1953) Indianapolis data.

A Refined Multiplicative Model of the Mobility Table¹

Let x_{ij} be the observed frequency in the ij th cell of the classification of men by their own occupations ($j = 1, \dots, J$) and their own occupations or fathers' occupations at an earlier time ($i = 1, \dots, I$). In the context of mobility analysis the same categories will appear in rows and columns, and the table will be square with $I = J$. For $k = 1, \dots, K$, let H_k be a mutually exclusive and exhaustive partition of the pairs (i, j) in which

$$E[x_{ij}] = m_{ij} = \alpha\beta_i\gamma_j\delta_{ij}, \quad (1)$$

where $\delta_{ij} = \delta_k$ for $(i, j) \in H_k$, subject to the normalization $\prod_i \beta_i = \prod_j \gamma_j = \prod_i \prod_j \delta_{ij} = 1$. The normalization of parameters is a matter of convenience, and we choose the value of one so that it will hold. Note that, unlike the usual set-up, the interaction effects are not constrained within rows or columns although the marginal frequencies are fixed. The model says the expected frequencies are a product of an overall effect (α), a row effect (β_i), a column effect (γ_j), and an interaction effect (δ_{ij}). The row and column parameters represent conditions of occupational supply and demand; they reflect demographic replacement processes and past and present technologies and economic conditions. The cells (i, j) are assigned to K mutually exclusive and exhaustive levels, and each of those

TABLE 1
 Frequencies in a Classification of Mobility from Father's (or Other Family Head's) Occupation to Son's First Full-Time Civilian Occupation: U.S. Men Aged 20-64 in March 1973

| Father's occupation | Son's occupation | | | | | |
|---------------------|------------------|-----------------|--------------|--------------|-------|--------|
| | Upper nonmanual | Lower nonmanual | Upper manual | Lower manual | Other | Total |
| Upper nonmanual | 614 | 571 | 93 | 81 | 11 | 2,020 |
| Lower nonmanual | 124 | 571 | 141 | 150 | 18 | 2,255 |
| Upper manual | 156 | 676 | 864 | 678 | 106 | 4,066 |
| Lower manual | 14 | 14 | 14 | 14 | 14 | 504 |
| Other | 69 | 69 | 69 | 69 | 69 | 489 |
| Total | 107 | 1,071 | 1,071 | 1,071 | 1,071 | 14,911 |

**P/u from
Social Strat. 2e,
top of p. 329**

NOTE: Frequencies are based on observations using a complex design from simple random sampling (see Goodman 1972c). Occupational and industrial workers are divided into professional and kindred workers, clerical and kindred workers, and retail sales workers; upper manual into farm managers, nonfarm kindred workers, lower manual, service workers, operatives and kindred workers, and unskilled workers; and farm managers, farm laborers and foremen.

levels shares a common interaction parameter δ_k . Thus, aside from total, row, and column effects, each expected frequency is determined by only one parameter, which reflects the level of mobility or immobility in that cell relative to that in other cells in the table.

The interaction parameters of the model correspond directly to our notions of variations in the density of observations (White, 1963:26). Unlike several models fitted by Goodman (1972c), this model does not assume ordinal measurement of occupations. Of course, the assumption of ordinality may help us interpret results, or our findings may be used to explore the metric properties of our occupational classification. For the model to be informative, the distribution of levels across the cells of the table must form a meaningful pattern, and one in which the parameters are identified (Mason *et al.*, 1973; Haberman, 1974: 217). Furthermore, the number of levels (K) should be substantially less than the number of cells in the table. These latter properties are partly matters of substantive and statistical interpretation and judgment, rather than characteristics of the model or of the data. We have found it difficult to interpret models where the number of levels is much greater than the

number of categories recognized in the occupational classification. . . .

Mobility to First Jobs: An Illustration

Table 1 gives frequencies in a classification of son's first, full-time civilian occupation by father's (or other family head's) occupation at the son's sixteenth birthday among American men who were ages 20-64 in 1973 and were not currently enrolled in school.² Table 2 gives the design matrix of a model for the data of Table 1. Each numerical entry in the body of the table gives the level of H_k to which the corresponding entry in the frequency table was assigned. Formally, the entries are merely labels, but, for convenience in interpretation, the numerical values are inverse to the estimated density of mobility or immobility in the cells to which they refer.

On this understanding the design says that, aside from conditions of supply and demand, immobility is highest in farm occupations (Level 1) and next highest in the upper nonmanual category (Level 2). If we take the occupation groups as ranked from high to low in the order listed, we may say that there are zones of high and almost uniform density bordering the peaks at either end of the status distribution. There is one zone of high

TABLE 2

An asymmetric 5-Level Model of Mobility from Father's Occupation to First Full-Time Civilian Occupation

| Father's Occupation | Son's Occupation | | | | |
|---------------------|------------------|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| 1. Upper nonmanual | 2 | 1 | 3 | 4 | 5 |
| 2. Lower nonmanual | 1 | 4 | 3 | 5 | 2 |
| 3. Upper manual | 1 | 1 | 5 | 3 | 2 |
| 4. Lower and farm | 1 | 1 | 1 | 2 | 3 |
| 5. Farm | 1 | 1 | 1 | 1 | 1 |

**P/u from
Social Strat. 2e,
bottom of p. 329**

NOTE: Sons' occupations are: 1. upper nonmanual (white-collar), 2. lower nonmanual (blue-collar), 3. upper manual (white-collar), 4. lower manual (blue-collar), and 5. farm. Includes clerical, sales, and service workers, operatives and kindred workers, lower manual except farm, laborers and farm managers, laborers, laborers and foremen.

density that includes upward or downward movements between the two nonmanual groups and immobility in the lower nonmanual group. Mobility from lower to upper nonmanual occupations (Level 3) is more likely than the opposite movement, and the latter is as likely as stability in the lower nonmanual category (Level 4). Moreover, the densities of immobility in the lower nonmanual category and of downward mobility to it are identical to those in the second zone of relatively high density, which occurs at the lower end of the occupational hierarchy. The second zone includes movements from the farm to the lower manual group and back as well as immobility in the lower manual group. Last, there is a broad zone of relatively low density (Level 5) that includes immobility in the upper manual category, upward and downward mobility within the manual stratum, mobility between upper manual and farm groups, and all movements between nonmanual and either manual or farm groups. The design says that an upper manual worker's son is equally likely to be immobile or to move to the bottom or top of the occupational distribution; obversely, it says that an upper manual worker is equally likely to have been recruited from any location in the occupational hierarchy, including his

own. Also, it is worth noting that four of the five density levels recognized in the model occur along the main diagonal, and two of these (Levels 4 and 5) are assigned both to diagonal and off-diagonal cells.

With a single exception the design is symmetric. That is, the upward and downward flows between occupations are assigned to the same density levels, except mobility from lower to upper nonmanual strata (Level 3) exceeds that from upper to lower nonmanual strata (Level 4). This asymmetry in the design is striking because it suggests the power of upper white-collar families to block at least one type of status loss and because it is the *only* asymmetry in the design. For example, Blau and Duncan (1967: 58–67) suggest that there are semipermeable class boundaries separating white-collar, blue-collar, and farm occupations, which permit upward mobility but inhibit downward mobility. The only asymmetry in the present design occurs *within* one of the broad classes delineated by Blau and Duncan.

Overall, the design resembles a river valley in which two broad plains are joined by a narrow strip of land between two great peaks. The contours of the peaks differ in that the one forming one side of the valley is both taller and more nearly symmetric than that forming the other side. This representation appears in Figure 1.

In some respects, this design matrix parallels Levine's (1967: Chap. 4) description of the surface of the British mobility table as a saddle (also see Levine, 1972). However, our interpretation is more extreme, since the density reaches an absolute minimum in the center of the table, not merely a minimum among the diagonal cells. In this way our model for the American 5 × 5 table is closer to Goodman's (1969a: 38, 1969b: 846) conclusion that a British 5 × 5 table shows "status disinheritance" in the middle category. We show elsewhere (Featherman and Hauser, 1978) that Levine's interpretation of the British data is based on a confounding of

FIGURE 1

Volume of mobility from father's occupation to first full-time adult occupation, U.S. men aged 20-64 in March 1970. The base is a unit square, and the total volume under the surface is one. Length and breadth can be read as probabilities, and height is proportionate to probability. The vertical scale has been compressed by a factor of 10.

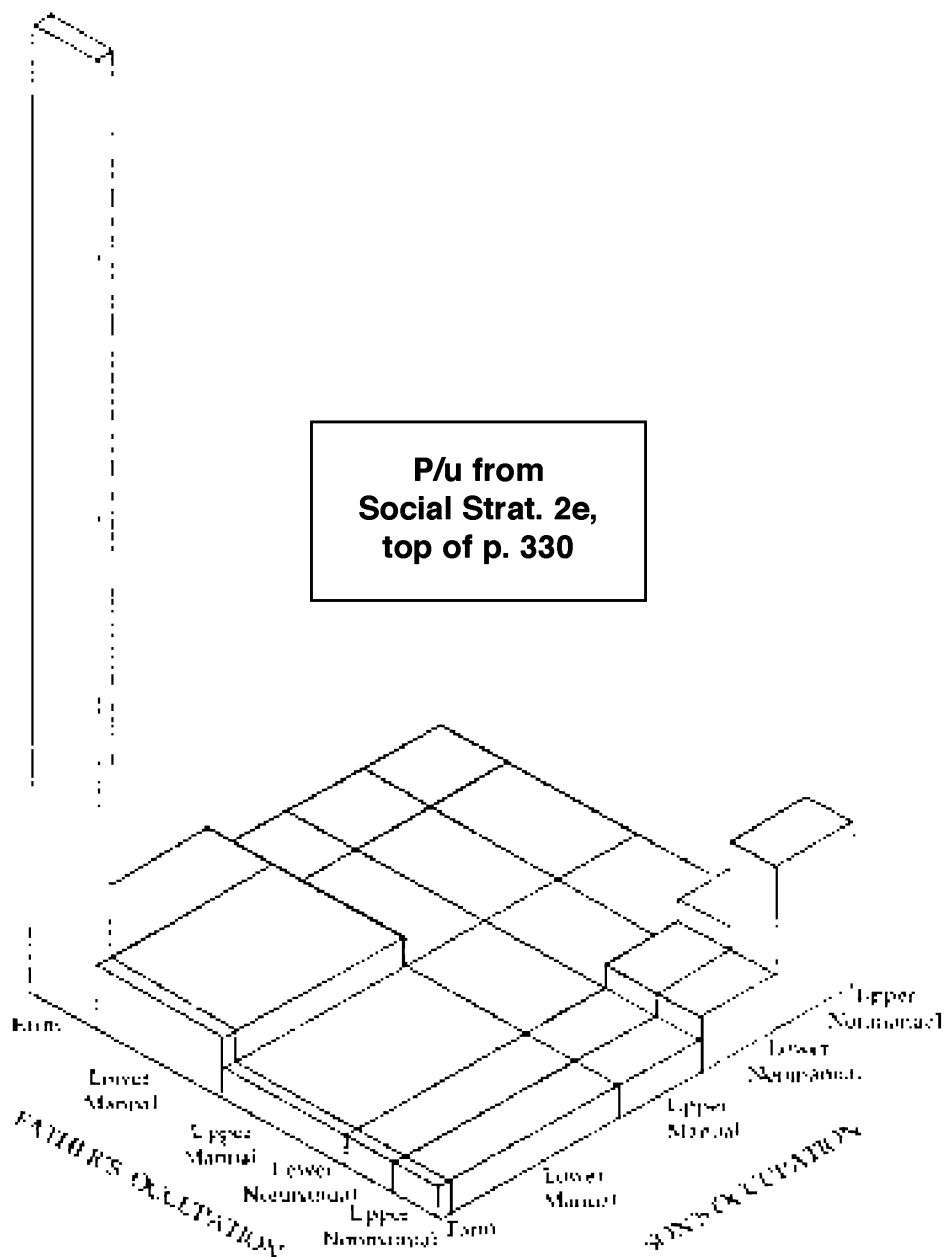


TABLE 3

Parameters and Residuals in Additive Model from Main, Row, and Column Effects in the Model of Table 2: Mobility from Father's (or Other Family Head's) Occupation to Son's First Full-Time Civilian Occupation, U.S. Men Aged 16-64 in March 1973

| A. Additive parameters | | Occupation of son, column of Table 2 | | | | |
|---|--------|--------------------------------------|------|-------|--------|------|
| Designation | | 1 | 2 | 3 | 4 | 5 |
| Father's occupation | | | | | | |
| Father's occupation | -1.068 | 1.41 | 1.83 | 1.70 | -1.13 | |
| Column (son's occupation) | 1.269 | 1.90 | 1.40 | 1.020 | -1.040 | |
| Level effects | | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| <div style="border: 1px solid black; padding: 5px; display: inline-block;"> P/u from Social Strat. 2e, p. 332 </div> | | | | | | |
| B. Exact partitioning of residuals | | New occupation | | | | |
| Father's occupation | | 1 | 2 | 3 | 4 | 5 |
| 1 Upper nonmanual | 1.231 | 1.78 | 1.40 | 1.26 | -1.13 | |
| 2 Lower nonmanual | 1.0 | 1.78 | 1.40 | 1.26 | -1.13 | |
| 3 Upper manual | -1.14 | 1.59 | 1.0 | 1.0 | 1.0 | |
| 4 Lower manual | 1.47 | 1.22 | 1.1 | 1.4 | 1.8 | |
| 5 Farm | 1.57 | 1.41 | 1.57 | 1.4 | 1.54 | |

NOTE: See text for explanation.

marginal effects and interactions which parallels that entailed in the use of mobility ratios, even though Levine did not use mobility ratios.

The model of Table 2 provides less than a complete description of the mobility data in Table 1. Under the model of statistical independence we obtain a likelihood-ratio statistic, $G^2 = 6167.7$, which is asymptotically distributed as χ^2 with 16 *df*. With the model of Table 2 as null hypothesis we obtain $G^2 = 66.5$ with 12 *df*; since we lose 4 *df* in creating the five categories of *H*. Clearly the model does not fit, if we take the probability associated with the test statistic as our only guide. On the other hand the model does account for 98.9% of the association in the data, that is, of the value of G^2 under independence. Given the extraordinarily large sample size we might expect small departures from frequencies predicted by the model to be statistically significant. . . .

The measures of fit we have examined have told us nothing about the several parameters of the model. That is, we have not shown that our suggested interpretation of the design matrix (Table 2) is substantively appeal-

ing, or even that the design correctly sorts the cells of the mobility table into zones of high and low density. Certainly, we want to look at the way in which the model fits and interprets the data as well as at deviations from fitted values.

The upper panel of Table 3 shows the row, column, and level parameters estimated under the model of Table 2 for mobility in the 1973 data from father's (or head's) occupation at son's sixteenth birthday to son's first full-time civilian occupation. The parameters are expressed in additive form, that is, they are effects on logs of frequencies under the model of Eq. (1). The row and column parameters clearly show an intergenerational shift out of farming and into white-collar or lower blue-collar occupations. Of course these parameters reflect a number of factors, including temporal shifts in the distribution of the labor force across occupations, differential fertility, and life cycle differences in occupational positions. The level parameters show very large differences in mobility and immobility across the several cells of the classification, and these differences closely follow our interpretation

of the design matrix. Differences between level parameters may readily be interpreted as differences in logs of frequencies, net of row and column effects. For example, the estimates say that immobility in farm occupations is $3.40 = 3.044 - (-.356)$ greater (in the metric of logged frequencies) than the estimated mobility or immobility in cells assigned to Level 5 in the design matrix. In multiplicative terms, immobility in farm occupations is $e^{3.40} = 29.96$ times greater than mobility or immobility at Level 5. It would be incorrect to attach too much importance to the signs of the level parameters as reported in Table 3, for they simply reflect our normalization rule that level parameters sum to zero (in the log-frequency metric) across the cells of the table. For example, while the parameters for Levels 4 and 5 each reflect relatively low densities, it is not clear that either parameter indicates “status disinheritance” in the diagonal cells to which it pertains (compare Goodman, 1969a, 1969b).

In any event the parameters do show a sharp density gradient across the levels of the design. The smallest difference, between Levels 3 and 4, indicates a relative density $e^{.549-.243} = e^{.306} = 1.36$ times as great at Level 3 as at Level 4. The heterogeneity of Level 5 is indicated by the fact that the difference in density between Levels 3 and 4 is about as large as the range of residuals within Level 5. Immobility in farm occupations and in upper nonmanual occupations is quite distinct from densities at other levels, but also immobility in the farm occupations is $e^{3.044-1.234} = e^{1.810} = 6.11$ times as great as in the upper nonmanual occupations.

We can write the sample counterpart of Eq. (1) as

$$\hat{m}_{ij} = \hat{\alpha}\hat{\beta}_i\hat{\gamma}_j\hat{\delta}_{ij} \tag{2}$$

Recalling that

$$e_{ij} = x_{ij} / \hat{m}_{ij}, \tag{3}$$

we substitute Eq. (2) into (3) and rearrange terms to obtain

$$x_{ij} = \hat{\alpha}\hat{\beta}_i\hat{\gamma}_j\hat{\delta}_{ij}e_{ij} \tag{4}$$

We divide both sides of Eq. (4) by the first three terms on the right-hand side to obtain

$$R_{ij}^* = \frac{x_{ij}}{\hat{\alpha}\hat{\beta}_i\hat{\gamma}_j} = \hat{\delta}_{ij}e_{ij} \tag{5}$$

We shall call R_{ij}^* the new mobility ratio, or, simply, the mobility ratio. In the case of diagonal cells R_{ij}^* is equivalent to the new immobility ratio proposed by Goodman (1969a,b, 1972c; also, see Pullum, 1975: 7–8), but we suggest the ratio be computed for all cells of the table as an aid both to substantive interpretation and to the evaluation of model design.

The lower panel of Table 3 gives logs of the new mobility ratios for the model of Table 2 fitted to the classification of mobility to first jobs. While the entries in this panel depend on our specification of the model, that specification need not rigidly govern our interpretation of the relative densities. Obviously, the pattern of relative densities does conform substantially to our earlier description of the design. The fit is good enough so there is no overlap in densities across levels recognized in the design, and all of the negative entries are neatly segregated in Level 5 of the design. If immobility among skilled workers—in cell (3, 3)—is high relative to mobility in other cells at Level 5, it is still clear that the immobility in that category is substantially less than the immobility in any other occupation group. . . .

Mobility Chances: A New Perspective

As an alternative to the Blau-Duncan interpretation, we think our multiplicative models yield a cogent and parsimonious description

of occupational mobility among American men. Unlike its precursor, our description does not reflect the shape of occupational distributions of origin or destination, but only the underlying patterns of immobility and exchange between occupational strata. It may be useful here to review the major features of this description that appear in mobility between generations and within the occupational career. In doing so, of course, we do some injustice to details reported in the preceding analysis.

First, there is great immobility at the top and at the bottom of the occupational hierarchy, here represented by upper nonmanual and by farm occupations, respectively. This immobility is far more extreme than has heretofore been supposed by most students of the mobility process; it may even be consistent with the beliefs of the more extreme critics of rigidity in the American class structure.

Second, surrounding the extremes of the occupational hierarchy at both top and bottom are transitional zones, within which there are relatively homogeneous chances of immobility and of exchange with adjacent extreme strata.

Third, taken as aggregates the extreme and transitional zones of the occupational hierarchy are relatively closed both to upward and to downward movements. That is, there are sharp breaks between the density of observations within the extreme and transitional zones and the density of mobility beyond those zones. In this sense (but not in others) we may say that the data suggest the existence of barriers to movement across class boundaries.

Fourth, once the boundaries of the transitional zones have been crossed, no social distance gradient seems to underlie variations in long-distance mobility chances. These are surprisingly uniform, and observed variations in them show no consistent pattern.

Fifth, if immobility is very great at the extremes of the occupational hierarchy, it is almost nonexistent in the middle of the hierarchy. Contrary to widespread belief, men of upper blue-collar origin are about as likely to end up anywhere higher or lower in the occupational hierarchy as in their stratum of origin. Obversely, upper blue-collar workers are about as likely to have originated anywhere higher or lower in the occupational hierarchy as in their stratum of destination. Those who would find their beliefs about “class” rigidity confirmed by our estimates of immobility at the extremes of the occupational hierarchy must reconcile these with our finding that between generations immobility in upper manual occupations is no more prevalent than most types of extreme, long-distance mobility. There is no evidence of “class” boundaries limiting the chances of movement to or from the skilled manual occupations.

Sixth, there is a rough equality in the propensities to move in one direction or the other between occupational strata. There are several exceptions to this symmetric mobility pattern, some of which may be quite important, but none suggests a dominant tendency toward upward relative to downward mobility across or within class boundaries.

Last, from a methodological perspective, our description of the mobility regime is extremely simple. In broad outline it might be fitted to a 5×5 table with the expenditure of as few as 2 *df*. None of our analyses of American mobility tables required the expenditure of more than 6 of the 16 *df* left unused by the model of simple statistical independence.

We reemphasize that the present description of relative mobility chances does not conflict in any way with our earlier description (Featherman and Hauser 1978: Chapter 3) of occupational inflow and outflow patterns. Rather, relative mobility chances are components of inflows and outflows, but the

latter are also affected by distributions of occupational origins and destinations. Major features of the inflow and outflow tables, like the pervasiveness of upward mobility between generations and within the occupational career, are absent from the present account because they are functions of changing distributions of occupational origins and destinations.

NOTES

1. We assume the familiarity of the reader with loglinear models for frequency data. Fienberg (1970a, 1977) and Goodman (1972a,b) give useful introductions, as does the comprehensive treatise by Bishop *et al.* (1975). We rely heavily on methods for the analysis of incomplete tables, which have been developed by Goodman (1963, 1965, 1968, 1969a,b, 1971, 1972c), Bishop and Fienberg (1969), Fienberg (1970b, 1972), and Mantel (1970); again, Bishop *et al.* (1975, especially pp. 206–211, 225–228, 282–309, 320–324) is valuable. Our model is a special case of Goodman's (1972c) general model.

2. The reported frequencies are based on a complex sampling design and have been weighted to estimate population counts while compensating for certain types of survey nonresponse. The estimated population counts have been scaled down to reflect underlying sample frequencies, and an additional downward adjustment was made to compensate for departures of the sampling design from simple random sampling (see Featherman and Hauser [1978: Appendix B]). The frequency estimates in Table 1 have been rounded to the nearest integer, but our computations have been based on unrounded figures. We treat the adjusted frequencies as if they had been obtained under simple random sampling.

BIBLIOGRAPHY

- Baron, James N. 1977. "The structure of intergenerational occupational mobility: Another look at the Indianapolis mobility data." Unpublished masters thesis, University of Wisconsin—Madison.
- Bibby, John. 1975. "Methods of measuring mobility." *Quality and Quantity* 9 (March): 107–136.
- Bishop, Yvonne M. M., and Stephen E. Fienberg. 1969. "Incomplete two-dimensional contingency tables." *Biometrics* 25:119–128.
- Bishop, Yvonne M. M., Stephen E. Fienberg, and Paul W. Holland. 1975. *Discrete Multivariate Analysis: Theory and Practice*. Cambridge: M.I.T. Press.
- Blau, Peter M. 1965. "The flow of occupational supply and recruitment." *American Sociological Review* 30 (August):475–490.
- Blau, Peter M., and Otis Dudley Duncan. 1967. *The American Occupational Structure*. New York: Wiley.
- Boudon, Raymond. 1973. *Mathematical Structure of Social Mobility*. San Francisco: Jossey-Bass.
- Chessa, Frederico. 1911. *La Trasmissione Ereditaria delle Professioni*. Torino: Fratelli Bocca.
- Duncan, Otis Dudley, and Robert W. Hodge. 1963. "Education and occupational mobility." *American Journal of Sociology* 68 (May): 629–644.
- Featherman, David L., and Robert M. Hauser. 1978. *Opportunity and Change*. New York: Academic Press.
- Featherman, David L., F. Lancaster Jones, and Robert M. Hauser. 1975. "Assumptions of social mobility research in the United States: the case of occupational status." *Social Science Research* 4:329–360.
- Fienberg, Stephen E. 1970a. "The analysis of multi-dimensional contingency tables." *Ecology* 51:419–433.
- . 1970b. "Quasi-independence and maximum likelihood estimation in incomplete contingency tables." *Journal of the American Statistical Association* 65:1610–1616.
- . 1972. "The analysis of incomplete multi-way contingency tables." *Biometrics* 23 (March):177–202.
- . 1977. *The Analysis of Cross-Classified Categorical Data*. Cambridge: M.I.T. Press.
- Giddens, Anthony. 1973. *The Class Structure of the Advanced Societies*. New York: Harper and Row.
- Glass, D. B. 1954. *Social Mobility in Britain*. London: Routledge and Kegan Paul.
- Goldthorpe, John W., and Catriona Llewellyn. 1977. "Class mobility in modern Britain: three theses examined." *Sociology* 11 (May): 257–287.
- Goldthorpe, John W., Clive Payne, and Catriona Llewellyn. 1978. "Trends in class mobility." *Sociology* 12 (September):441–468.
- Goodman, Leo A. 1961. "Statistical methods for the mover-stayer model." *Journal of the American Statistical Association* 56 (December):841–868.
- . 1963. "Statistical methods for the preliminary analysis of transaction flows." *Econometrica* 31 (January):197–208.

- . 1965. "On the statistical analysis of mobility tables." *American Journal of Sociology* 70 (March):564–585.
- . 1968. "The analysis of cross-classified data: independence, quasi-independence, and interaction in contingency tables with or without missing entries." *Journal of the American Statistical Association* 63 (December): 1091–1131.
- . 1969a. "How to ransack social mobility tables and other kinds of cross-classification tables." *American Journal of Sociology* 75 (July): 1–39.
- . 1969b. "On the measurement of social mobility: an index of status persistence." *American Sociological Review* 34 (December): 831–850.
- . 1971. "A simple simultaneous test procedure for quasi-independence in contingency tables." *Applied Statistics* 20:165–177.
- . 1972a. "A general model for the analysis of surveys." *American Journal of Sociology* 77 (May):1035–1086.
- . 1972b. "A modified multiple regression approach to the analysis of dichotomous variables." *American Sociological Review* 37 (February):28–46.
- . 1972c. "Some multiplicative models for the analysis of cross-classified data." Pp. 649–696 in *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley: University of California Press.
- Haberman, Shelby J. 1974. *The Analysis of Frequency Data*. Chicago: University of Chicago Press.
- Hauser, Robert M. 1978. "A structural model of the mobility table." *Social Forces* 56 (March): 919–953.
- Hauser, Robert M., Peter J. Dickinson, Harry P. Travis, and John M. Koffel. 1975. "Temporal change in occupational mobility: evidence for men in the United States." *American Sociological Review* 40 (June):279–297.
- Hope, Keith. 1976. Review of Thomas W. Pullum's *Measuring Occupational Inheritance*. *American Journal of Sociology* 82 (November):726–730.
- Iutaka, S., B. F. Bloomer, R. E. Burke, and O. Wolowyna. 1975. "Testing the quasi-perfect mobility model for intergenerational data: international comparison." *Economic and Social Review* 6:215–236.
- Levine, Joel Harvey. 1967. Measurement in the study of intergenerational status mobility. Unpublished doctoral dissertation. Department of Social Relations. Harvard University.
- . 1972. "A two-parameter model of interaction in father-son status mobility." *Behavioral Science* 17 (September):455–465.
- Mantel, Nathan. 1970. "Incomplete contingency tables." *Biometrics* 26:291–304.
- Mason, Karen Oppenheim, William M. Mason, Halliman H. Winsborough, and Kenneth W. Poole. 1973. "Some methodological issues in cohort analysis of archival data." *American Sociological Review* 38 (April):242–258.
- Parkin, Frank. 1971. *Class Inequality and Political Order: Social Stratification in Capitalist and Communist Societies*. New York: Praeger.
- Pearson, Karl. 1904. "On the theory of contingency and its relation to association and normal correlation." Reprinted, 1948, in *Karl Pearson's Early Papers*. Cambridge: Cambridge University Press.
- Pullum, Thomas. 1975. *Measuring Occupational Inheritance*. New York: Elsevier.
- Ramsøy, Natalie. 1977. *Social Mobilitet i Norge (Social Mobility in Norway)*. Oslo: Tiden Forlag.
- Rogoff, Natalie. 1953. *Recent Trends in Occupational Mobility*. Glencoe, Illinois: Free Press.
- Singer, Burton, and Seymour Spilerman. 1976. "The representation of social processes by Markov models." *American Journal of Sociology* 82 (July):1–54.
- Tyree, Andrea. 1973. "Mobility ratios and association in mobility tables." *Population Studies* 27 (July):577–588.
- Westergaard, John, and Henrietta Resler. 1975. *Class in a Capitalist Society: A Study of Contemporary Britain*. New York: Basic Books.
- White, Harrison C. 1963. "Cause and effect in social mobility tables." *Behavioral Science* 8:14–27.
- . 1970a. *Chains of Opportunity: System Models of Occupational Mobility in Organizations*. Cambridge: Harvard University Press.
- . 1970b. "Stayers and movers." *American Journal of Sociology* 76 (September): 307–324.

47 ■ Robert Erikson and John H. Goldthorpe

Trends in Class Mobility

The Post-War European Experience

Introduction

The issue of trends in class mobility in industrial societies is one characterised by a wide-ranging dissensus which, unfortunately, extends to matters of fact as well as of interpretation. We do not suppose that in this paper we will be able to resolve all the disagreements that are apparent. We do, however, believe that we can address the issue on the basis of comparative mobility data of a distinctively higher quality than those previously utilised, and that the results we report have significant consequences—positive or negative—for most of the rival positions that have been taken up.¹

From the 1960s onwards, perhaps the dominant view on mobility trends has been that derived from what we will refer to as the ‘liberal theory’ of industrialism, as developed by various American authors (Kerr *et al.*, 1960, 1973; Kerr, 1969, 1983; Dunlop *et al.*, 1975; cf. also Parsons, 1960: chs. 3 and 4; 1967: chs. 4 and 15; 1971). This theory is a functionalist one which aims at establishing the distinctive properties of industrial societies in terms of the essential prerequisites for, or necessary consequences of, the technical and economic rationality that is seen as their defining characteristic. What is implied so far as social mobility is concerned may be put in the form of the following three-part proposition.

In industrial societies, in comparison with preindustrial ones,

- (i) absolute rates of social mobility are generally high, and moreover upward mobility—i.e. from less to more advantaged positions—predominates over downward mobility;
- (ii) relative rates of mobility—or, that is, mobility opportunities—are more equal, in the sense that individuals of differing social origins compete on more equal terms to attain (or to avoid) particular destinations; and
- (iii) both the level of absolute rates of mobility and the degree of equality in relative rates tend to increase over time.

To explain *why* these contrasts between mobility in pre-industrial and industrial society should arise, a number of arguments are deployed which have, moreover, been elaborated and extended in the specialist literature by authors generally sympathetic to the liberal position (see esp. Blau and Duncan, 1967: ch. 12; Treiman, 1970). While all the arguments in question take on a functionalist form, one may usefully distinguish between those relating to three different kinds of effect—*structural*, *processual* and *compositional*.

First, it is held that within industrial society the dynamism of a rationally developed

This chapter, originally prepared for the European Research Conference (“European Society or European Societies,” Gausdal, Norway, 1991) relies on Goldthorpe, *The Constant Flux: A Study of Class Mobility in Industrial Societies*. Copyright © 1992 by Oxford University Press. Used by permission of Oxford University Press.

technology calls for continuous, and often rapid, change in the structure of the social division of labour, which also tends to become increasingly differentiated. High rates of mobility thus follow as from generation to generation, and in the course of individual lifetimes, the redistribution of the active population is required: that is, among economic sectors—first, from agriculture to manufacturing and then from manufacturing to services—and, in turn, among industries and among a growing diversity of occupations. Furthermore, the overall tendency is for advancing technology to *upgrade* levels of employment. Although some skills are rendered obsolete, new ones are created and the *net* effect is a reduction in the number of merely labouring and routine occupations and a rising demand for technically and professionally qualified personnel. At the same time, both the increasing scale of production, dictated by economic rationality, and the expansion of the services sector of the economy promote the growth of large bureaucratic organisations in which managerial and administrative positions also multiply. Industrial societies become increasingly ‘middleclass’ or at least ‘middle-mass’ societies. Consequently, upward mobility is more likely than downward in both intergenerational and worklife perspective. Under industrialism, the chances of ‘success’ are steadily improved for all.

Secondly, it is further claimed that as well as thus reshaping the objective structure of opportunity, industrialism transforms the processes through which particular individuals are allocated to different positions within the division of labour. Most fundamentally, rational procedures of social selection require a shift away from *ascription* and towards *achievement* as the leading criterion: what counts is increasingly what individuals can do, and not who they are. Moreover, the growing demand for highly qualified personnel promotes the expansion of education and training, and also the reform of educational

institutions so as to increase their accessibility to individuals of all social backgrounds. Human resources cannot be wasted; talent must be fully exploited wherever it is to be found. Thus, as within a society of widening educational provision ‘meritocratic’ selection comes to predominate, the association between individuals’ social origins and their eventual destinations tends steadily to weaken and the society takes on a more ‘open’ character. And at the same time various other features of industrialism also serve to reduce the influence of social origins on individuals’ future lives. For example, urbanisation and greater geographical mobility loosen ties of kinship and community; mass communications spread information, enlarge horizons and raise aspirations; and a greater equality of condition—that is, in incomes and living standards—means that the resources necessary for the realisation of ambition are more widely available.

Thirdly, it is argued that the foregoing effects interact with each other, in that the emphasis on achievement as the basis for social selection will be strongest within the expanding sectors of the economy—that is, the more technologically advanced manufacturing industries and services—and within the increasingly dominant form of large-scale bureaucratic organisation. Conversely, ascriptive tendencies will persist chiefly within declining sectors and organisational forms—for example, within agriculture or small-scale, family-based business enterprise. In other words, compositional effects on mobility occur in that, once a society begins to industrialise, the proportion of its population that is subject to the new ‘mobility regime’ characteristic of industrialism not only increases as that regime imposes itself, but further as those areas and modes of economic activity that are most resistant to it become in any event ever more marginal.

One reason that may then be suggested for the degree of dominance exerted by the liberal position is the coherent way in which the

underlying theory has been developed. Another is the manifest failure of the main attempt made directly to controvert it. That is, the revision and extension of the Marxist theory of proletarianisation (cf. Braverman, 1974; Carchedi, 1977; Wright and Singelmann, 1982; Crompton and Jones, 1984) which sought to show the necessity for the systematic ‘degrading’, rather than upgrading, of labour under the exigencies of late capitalism—with the consequence of large-scale *downward* mobility of a collective kind. This undertaking lacked from the start any secure empirical foundation, and the accumulation of results incompatible with the new theory resulted in its eventual abandonment even by those who had been among its most resourceful supporters (see e.g., Singelmann and Tienda, 1985; Wright and Martin, 1987). However, various other positions can still be identified that to a greater or lesser extent come into conflict with the liberal view and that continue to merit serious attention.

First of all, it should be noted that the theory of mobility in industrial society advanced by Lipset and Zetterberg (1956, 1959) and sometimes simply assimilated to the liberal theory (see e.g., Kerr, 1983: 53) does in fact differ from it in crucial respects. For example, Lipset and Zetterberg do not seek to argue that mobility steadily *increases* with industrial development: indeed, they remark that *among* industrial societies no association is apparent between mobility rates and rates of economic growth. What they propose (1959: 13) is, rather, some kind of ‘threshold’ effect: ‘our tentative interpretation is that the social mobility of societies becomes relatively high once their industrialization, and hence their economic expansion, reaches a certain level’. And although Lipset and Zetterberg’s claim that (absolute) mobility rates in industrial societies become *uniformly* high would now be generally regarded as empirically untenable, their suggestion that a historic upward shift in such rates tends to occur at some—perhaps quite early—stage in the in-

dustrialisation process has not been similarly disconfirmed. Again, it is not part of Lipset and Zetterberg’s case that the high mobility that they see as characteristic of industrial societies is the result of a tendency towards greater openness. Rather, they place the emphasis firmly on the effects of structural change, and in turn they are at pains to point out (1959: 27) that ‘the fact that one country contains a greater percentage of mobile individuals than another does *not* mean that that country approximates a model of equal opportunity more closely’.

Secondly, a yet more radical challenge to the liberal view may be derived from the pioneering work of Sorokin (1927/1959). Taking a synoptic view, as much dependent on historical and ethnographic evidence as on contemporary social research, Sorokin was led to the conclusion that in modern western societies mobility was at a relatively high level, and he was further ready to acknowledge the possibility that, from the eighteenth century onwards, mobility rates had in general shown a tendency to rise. However, he was at the same time much concerned to reject the idea that what was here manifested was in effect ‘the end of history’ and the start of a ‘perpetual and “eternal” increase of vertical mobility’. Rather, Sorokin argued, the present situation represented no more than a specific historical phase; in some societies in some periods mobility increased, while in other periods it declined. Overall, no ‘definite perpetual trend’ was to be seen towards either greater or less mobility, but only ‘trendless fluctuation’. Those who were impressed by the distinctiveness of the modern era knew too little about historical societies and their diversity: ‘What has been happening is only an alternation—the waves of greater mobility superseded by the cycles of greater immobility—and that is all’ (1959: 152–4).

It might from the foregoing appear that Sorokin’s position was merely negative. But, in fact, underlying his denial of developmental trends in mobility and his preference for a

cyclical view, at least the elements of a theory can be discerned. In arguing against the supposition that rates of mobility in the modern period are quite unprecedented, one of the points Sorokin most stresses is that while certain barriers to mobility have been largely removed—for example, juridical and religious ones—it is important to recognise that other barriers have become more severe or have been newly introduced: for example, those represented by systems of educational selection and occupational qualification (1959: 153–4, 169–79). This, moreover, is what must always be expected: the forms of social stratification which provide the context for mobility are themselves structures expressing differential power and advantage, and thus possess important self-maintaining properties. Those who hold privileged positions will not readily cede them and, in the nature of the case, can draw on superior resources in their defence. Indeed, Sorokin remarks that if he *had* to believe in the existence of a permanent trend in mobility, it would be in a declining one, since social strata are often observed to become more ‘closed’ over time as the cumulative result of those in superior positions using their power and advantage to restrict entry from below (1959: 158–60). However, this propensity for closure—which we may understand as being *endogenous* to all forms of stratification—is not the only influence on mobility rates. A further point that Sorokin several times makes (see e.g. 1959: 141–152, 466–72) is that in periods of both political and economic upheaval—associated, say, with revolution or war or with rapid commercial, industrial and technological change—marked surges in mobility are typically produced as the social structure as a whole, including the previously existing distribution of power and advantage, is disrupted. In other words, increased mobility here results from the impact of factors that are *exogenous* to the stratification order.

Thirdly and finally, one may note a more recently developed position which, however, has evident affinities with that of Sorokin.

Featherman, Jones and Hauser (1975) aim at presenting a reformulation of Lipset and Zetterberg’s hypothesis that across industrial societies rates of social mobility display a basic similarity. This hypothesis cannot stand if expressed in terms of absolute rates but, they argue, becomes far more plausible if applied, rather, to relative rates. When mobility is considered at the ‘phenotypical’ level of absolute rates cross-national similarity can scarcely be expected. This is because these rates are greatly influenced by the structural context of mobility and, in turn, by effects deriving from a range of economic, technological and demographic circumstances which are known to vary widely and which, so far as particular individuals and families are concerned, must be regarded as ‘exogenously determined’. When, however, mobility is considered *net of* all such effects, or that is, at the ‘genotypical’ level of relative rates, the likelihood of cross-national similarity being found is much greater. For at this level only those factors are involved that bear on the relative chances of individuals of differing social origin achieving or avoiding, in competition with each other, particular destination positions among those that are structurally given. And there is reason to suppose that in modern societies the conditions under which such ‘endogenous mobility regimes’ operate—for example, the degree of differentiation in occupational hierarchies and in job rewards and requirements—may not be subject to substantial variation.

For present purposes, then, the chief significance of the FJH hypothesis lies in the rather comprehensive challenge that it poses to the claims of liberal theorists. On the one hand, so far as absolute mobility rates are concerned, it implies a basic scepticism, essentially akin to that of Sorokin, about the possibility of *any* long-term, developmentally-driven trend; while, on the other hand, it stands directly opposed to the proposition that under industrialism a steady increase occurs in the equality of mobility chances. Although some initial developmental effect in

this direction early in the industrialisation process might be compatible with the hypothesis, any continuing change in relative mobility rates is clearly precluded (cf. Grusky and Hauser, 1984: 20). Once societies can be deemed to have become industrial, their mobility regimes should stabilise in some approximation to the common pattern that the FJH hypothesis proposes, and should not thereafter reveal any specific or persistent tendencies, whether towards convergence on greater openness or otherwise. No forces are recognised as inherent in the functional dynamics of industrialism that work systematically to expand mobility opportunities.

The divergent arguments concerning mobility trends that we have reviewed in this section will then provide the context within which we present our empirical analyses of data for European nations. First, though, we must say something about what we take to be the particular relevance to evaluating these arguments of the European experience over the post-war years.

The Relevance of the European Experience

It is not difficult to detect within the liberal theory of industrialism a degree of American or, more accurately perhaps, of Anglo-American ethnocentricity. Historically, the origins of modern industrial society are traced back to late eighteenth- and early nineteenth-century England; and other western nations, including the USA, are then seen as having successively followed England's lead in breaking free of the constraints of a traditional social order and entering the industrial world.² Contemporaneously, it is the USA rather than England that is recognised as the vanguard nation; and with industrialisation now on the global agenda, the major differences between industrial and pre-industrial, or modern and traditional, society are seen as best revealed through explicit or implicit USA–Third World comparisons. Within these perspectives, therefore,

the experience of industrialisation of the mainland European nations is viewed in only a rather restrictive way. It tends either to be taken for granted, as fitting unproblematically into the trajectories defined by the two paradigm cases, or alternatively as providing interesting instances of 'deviations', over which, however, the logic of industrialism has eventually to prevail.³

This schematic background to the liberal theory must be regarded as excessively simplified, and possibly misleading, in at least two respects. On the one hand, while England did indeed industrialise early, the supposition that other western nations then followed along the same path, being differentiated only by the degree of their 'retardation', is one that has no sound historiographic basis. What is chiefly significant about the process of industrialisation in England is that—in part *because* of England's priority, but for other reasons too—it took on a quite distinctive character which subsequent cases could scarcely reflect (cf. Kemp, 1978: ch. 1 esp.). Rather than having simply followed in England's wake, other European nations do in fact display in their recent economic histories a great diversity of developmental paths; and it is, furthermore, important to recognise that later industrialisation and economic retardation should not always be equated—as, for example, the French case can well illustrate (cf. O'Brien and Keyder, 1978).

On the other hand, it would also be mistaken to suppose that by the end of the nineteenth century the industrialisation of Europe was essentially completed. This would be to neglect the great economic and social importance that agriculture and also artisanal and other 'pre-industrial' forms of production continued to have throughout the nineteenth, and for well into the twentieth, century—and in many of the more advanced European nations as well as in those on the 'periphery'. It was in fact, as Bell has remarked (1980: 233), only in the period *after the second world war* that Europe as a whole

became an industrial society. Indeed, various interpretations of the ‘long boom’ of this period have seen it as reflecting aspects of this culmination—for example, as being driven by the final phase of the supply of surplus rural labour (Kindleberger, 1967) or as marking the ultimate overcoming of the dualism of traditional and modern sectors within European economies (Lutz, 1984).⁴

The fact, then, that we here concentrate on the experience of European nations by no means implies that we will be treating questions of mobility trends within an unduly limited context. To the contrary, we have the advantage that while these nations can supply us with high quality data (far better, for example, than those usually available from Third World nations), they do also display a remarkably wide range of variation in their levels and patterns of industrial development—and even if we consider only that time-span to which our mobility data have some reference: that is, from the 1970s back to the first two decades of the century, in which the oldest respondents within our national samples were born. (For details of the surveys utilised, see Appendix Table 1). . . .

It will from the foregoing have become evident that, in seeking to exploit the historical richness of our European data, we are prepared to make a large assumption: namely, that valid inferences about the presence or absence of mobility trends can be drawn from the data of single inquiries. What we are in effect proposing is that age-groups distinguished within our samples from the 1970s can be treated as successive birth-cohorts, and that the mobility experience of their members can then be taken as indicative of whether or not change over time has occurred in mobility rates and patterns. In such an approach certain well-known difficulties arise, above all in treating intergenerational mobility, and in the conclusion of this section we should therefore give these some attention.

To begin with, we must recognise that we are not in fact dealing with true birth cohorts

within the nations we consider but only with what might better be called ‘quasi-cohorts’: that is, with the survivors of true cohorts, following on losses due to mortality and emigration, to whom immigrants will then be added. In such cases as those of the FRG, Ireland and Poland, the numbers here involved will obviously be substantial. This, however, is a situation that we can do little to remedy; we can only trust that no serious distortions will be introduced into our data of a kind that might affect our conclusions regarding trends.⁵

Further, there is the so-called ‘identification problem’. If for the members of a national sample one compares their present class position (i.e. at time of inquiry) with their class of origin (i.e. father’s class), the mobility experience of the individuals within successive ‘quasi-cohorts’ will be likely to reflect several different effects: not only those of the historical period through which they have lived but also those of their age and of their cohort membership *per se*. Thus, the problem is that of how we can assess ‘period’ effects—which are those relevant to questions of mobility trends—separately from effects of the other kinds. No clear-cut solution is, or can be, available (cf. Glenn, 1977), since birth-cohorts and age-groups are inescapably ‘embedded’ in historical time. However, several considerations would lead us to believe that in pursuing our present purposes we need not in fact be at so great a disadvantage in this respect as might initially appear.

First, it would seem empirically defensible to regard men of around 30–35 years of age as having reached a stage of ‘occupational maturity’, beyond which further major changes in their class positions become relatively unlikely (Goldthorpe, 1980, 1987: ch. 3; cf. also Blossfeld, 1986). Thus, we may take results for cohorts of this age or older as giving a reasonably reliable indication of the ‘completed’ pattern of the collective class mobility of their members.

Secondly, for all of our European nations except one, the Federal Republic of Germany, we have information on individuals' experience of mobility from their class of origin to their class of *first* employment. For this transition, therefore, age effects at least will obviously be much reduced, since attention is focused on a fairly well defined life-cycle stage. We would not wish to regard data on this transition as a very satisfactory basis for cross-national comparisons of intergenerational mobility, on grounds that we discuss elsewhere (Erikson and Goldthorpe, 1992: ch. 8). None the less, we are thus provided with the possibility of checking whether or not the conclusions that we reach on trends, or their absence, in mobility from class of origin to present class are consistent with ones that pertain to mobility rates of a more age-specific kind.⁶

Thirdly, it is important for us to emphasise that in the analyses that follow our concern will be not so much with the actual empirical description of mobility trends as with the evaluation of particular claims about such trends. What, therefore, we can always consider is whether, if we were to suppose some confounding of effects in our results, these would be of a kind that would tend unduly to favour or disfavour a given position. Thus, for example, in the case of the liberal claim that within industrial nations mobility and openness tend steadily to increase, it is difficult to see why any confounding of period effects by age effects should produce unfairly *negative* results: that is to say, it would appear unlikely that an actual increase in openness and mobility among the more recent cohorts within our national samples would be concealed by the fact that these cohorts are made up of young persons. If anything, one might expect the contrary, since younger persons will have benefited more widely from the expansion of educational provision which, according to the liberal theory, is one of the major sources of greater mobility and equality of opportunity. Likewise, there would

seem no reason why age effects should obscure any trends within our data for the mobility rates of different nations to converge—as would be expected under the liberal theory as differences between nations' levels of industrial development are reduced. For if, as the theory maintains, the determinants and processes of mobility become increasingly standardised through the logic of industrialism, then convergence in mobility rates should, presumably, be *more* apparent among the younger than the older age groups in our samples (for an elaboration of this point, see Erikson, Goldthorpe and Portocarero, 1983: 307–10 and Figure 1).

As Glenn has observed (1977: 17), cohort analysis should never be a mechanical exercise, uninformed by theory and by additional 'external' evidence; and this point obviously applies *a fortiori* in the case of analyses, such as those we shall present, which rest only on 'quasi-cohorts'. But since we do have some knowledge about both the historical setting of the mobility that we consider and its life-course phasing, and since we are addressing a number of more or less specific and theoretically grounded hypotheses rather than proceeding quite empirically, our strategy is, we believe, one capable of producing results that can be interpreted in a reasonably reliable and consequential way. It is to these results that we now turn.

Absolute Rates

In seeking to assess the arguments that we have earlier reviewed, we start with evidence on intergenerational class mobility in the form of absolute rates: that is, rates based on differing versions of our class schema (see Appendix Table 2) and expressed in simple percentage terms. So as to avoid marked age effects in considering the transition from class of origin to present class, we restrict our attention to men in our national samples who were over age 30 at the time of inquiry (i.e. at some point in the early or mid-1970s; cf. Appendix

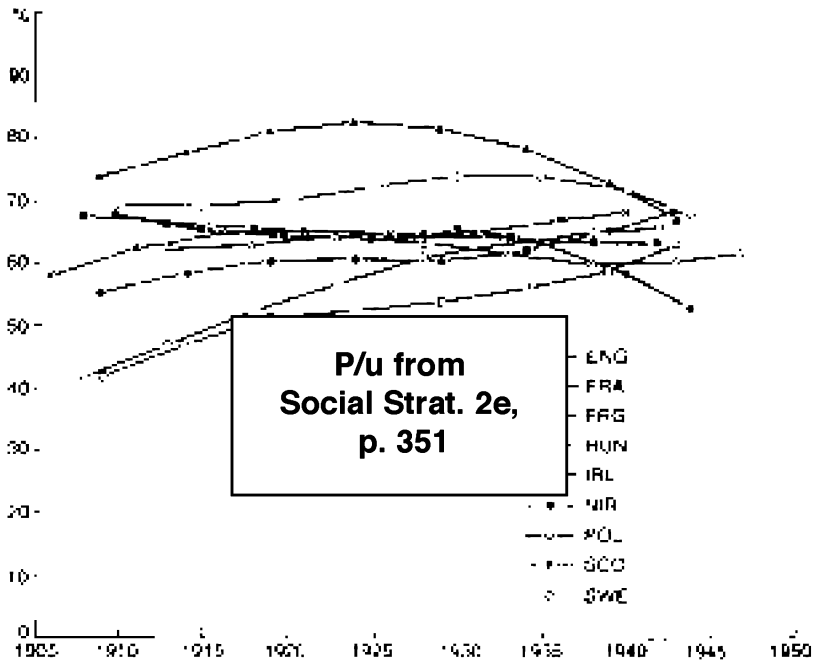


Figure 1. Total Mobility Rates for Men in the Nine Nations by Birth Year

Table 1). These men, we suppose, would be approaching, or would have attained, a stage of relative occupational maturity. The maximum age-limit that we apply here—and in all subsequent analyses—is 64.⁷

First of all, we consider *total* mobility rates. That is, the percentage of all men in our national samples found in cells off the main diagonal of the intergenerational mobility table based on the sevenfold version of the class schema; or, in other words, the percentage of all men whose 'present', or destination, class was different to their class of origin—the latter being indexed by the respondent's *father's* class at the time of the respondent's early adolescence.⁸ In Figure 1 we seek to plot the course followed by the total mobility rate in each of our nine European nations on the basis of moving weighted averages of this rate for men *in successive birth years*, using a method of graduation that has been developed by Hoem and Linneman (1987).⁹

From inspection of Figure 1, the general impression gained must be one of support for the contention that absolute mobility

rates display merely trendless change. It would, at all events, be difficult to ally the data here presented with the idea of mobility increasing steadily as industrialism advances. No regular tendency is apparent for the mobility of older respondents—of men born, say, in the first two decades of the century—to be exceeded by that of respondents who were born some twenty years later, and who would have reached occupational maturity during the long boom of the post-war years. And we may add that no essentially different picture emerges if, using the same technique, we plot total mobility rates from class of origin to class of *first* employment.¹⁰ It would thus seem improbable that the failure of the graphs of Figure 1 to move upwards to the right, as would be expected from the liberal theory, can be explained simply in terms of the confounding of period by age effects.

The one possible pattern that might be discerned in Figure 1 is some tendency for total mobility rates to converge—even if not while steadily rising. That is to say, some narrowing down could be claimed in the cross-

national range of mobility levels as between those displayed by the oldest and the youngest cohorts in our samples. For the former, as can be seen, the range of total mobility rates is from around 40 to over 70 per cent, while for the latter it is from 50 to under 70 per cent—and, one might add, would be some ten percentage points narrower still if the one case of Scotland were to be discounted.

However, it is important to note how this convergence comes about. It is in fact to a large extent the result of an increase in total mobility in two nations, Ireland and Poland, in which the rate among older cohorts, at around the 40 per cent mark, was substantially lower than in other nations. Ireland and Poland are—together with Hungary—those nations in our sample in which, as we have earlier indicated, industrialisation was most delayed. An alternative interpretation of Figure 1 would therefore be that instead of revealing a general tendency towards convergence in mobility rates, it rather supports a hypothesis of the kind suggested by Lipset and Zetterberg: that is, of a specific upturn in mobility occurring at a stage relatively early in the industrialisation process when the first major impact of structural change is felt.

That Hungary would then appear as deviant, in showing a high total mobility rate even among the oldest men considered, need not be found surprising. This could be seen as the result of the quite exceptional amount of mobility imposed upon the Hungarian agricultural workforce through direct political intervention, and which is thus reflected across the experience of all age-groups alike. In the period immediately following the second world war, the land reforms of the provisional government created over half-a-million new peasant proprietors; but then, under the subsequent state socialist regime, agriculture was within a decade almost entirely collectivised through the establishment of co-operatives or state farms (Kulcsár,

1984: 78–84, 96–100; Brus, 1986a, 1986b). Thus, respondents to our 1973 survey who were the sons of agricultural proprietors—over a quarter of the total—held different class positions to their fathers more or less of necessity, and even in fact where they continued to work the same land.¹¹

These findings on total mobility do, we believe, carry significant implications, to which we shall wish to return. It is, however, of further interest here to try to obtain a somewhat more detailed picture of tendencies in absolute rates by considering also intergenerational *outflow* rates. Unfortunately, the relative smallness of the sizes of certain of our national samples means that we cannot reliably base our examination of such rates on the seven-class version of our schema but must, for the most part, resort to the three-class version (cf. Appendix Table 2) which distinguishes simply between nonmanual, manual and farm classes.¹²

Figures 2 to 6, which are produced via the same procedures as Figure 1, show the course followed in each of our nations by five different outflow rates calculated from 3×3 intergenerational mobility tables (again for men aged 30–64). As is indicated, the rates in question are those for intergenerational immobility within the farm class, for mobility from farm origins to manual and to nonmanual destinations, and for mobility from manual origins to nonmanual destinations and *vice versa*. Of the other transitions possible within the 3×3 tables, those from manual and nonmanual origins to farm positions were generally followed by too few individuals to allow any reliable rates to be established; and the fact that the numbers involved here are more or less negligible means in turn that trends in the remaining rates—that is, rates of immobility within the manual and nonmanual classes—need scarcely be plotted separately, since they will be essentially the complements of those already examined of mobility between these two classes.

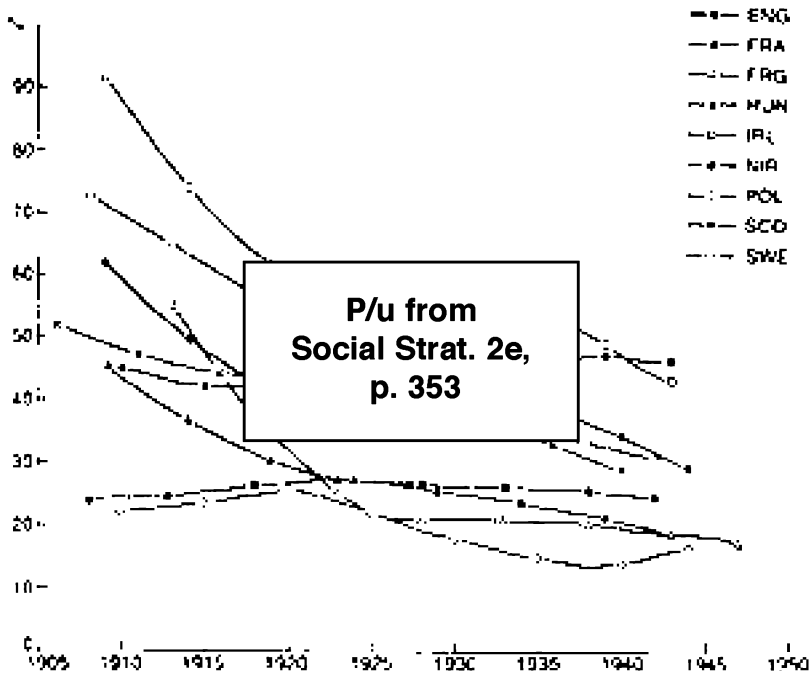


Figure 2. Outflow Rates from Farm Origins to Farm Destinations for Men in Nine Nations by Birth Year

Figure 2 displays the changing proportions of men across birth-cohorts in our nine nations who were of farm origins and who were themselves found in farm work. A broad tendency is apparent for such intergenerational immobility to decline, which might be expected in consequence of the general contraction of agricultural employment. The decline in the cases of Ireland and Poland from farm immobility rates of upwards of 70 per cent in the oldest cohorts is of particular interest in view of the interpretation we have suggested of the increases in total mobility in these nations revealed in Figure 1. By resorting to the raw data, we can in fact show that changes within the farm sector here played a crucial part. Thus, the contribution of this sector to the total *immobility* rate (i.e. the proportion of all cases in the mobility table found in cells on the main diagonal) fell in the Irish case from 69 per cent for men born before 1925 to only 27 per cent for those born after 1940, while in the Polish

case the corresponding decline was from 77 to 35 per cent.

In two other nations, France and Hungary, the decline in farm immobility is also more or less continuous over the period to which our data refer. However, the cases of England, Scotland and Northern Ireland and likewise that of the FRG would suggest that, once the decline of agriculture has reached a certain point, rates of farm immobility tend to level out or to become rather variable. And Sweden appears quite distinctive in that farm immobility is shown at a low level—never more than 25 per cent—throughout the decades in which agricultural employment was falling. Here, though, we do have evidence to suggest some distortion in our results. Our corresponding plot for the transition to class of first employment indicates a strong decline in farm immobility; but, on account perhaps of the very rapidity of agricultural contraction in the post-war years, it would seem that many men also left the farm workforce at a quite late age, thus obscuring

the downward trend when the transition to present class is considered.

Finally, it may be observed that in Figure 2, as in Figure 1, any impression of converging rates is created essentially by the rather dramatic Irish and Polish graphs. If these are disregarded, the cross-national range in rates of farm immobility merely fluctuates, being, for example, no narrower—at around 15 to 50 per cent—for men born from the mid-1930s onwards than it was for men born around 1920.

Figures 3 and 4 then display the course of outflow rates from farm origins to manual and nonmanual destinations respectively. Figure 3 would suggest that in those cases where declining trends in intergenerational immobility in farming were revealed in Figure 2, their counterpart has been increased outflows from farm origins into manual wage-earning positions in industry. France, Hungary, Ireland and Poland all show such increases of a continuous kind. In the remaining nations, however, trends are less readily discerned. In the cases of the FRG and Sweden, increasing proportions of men in the cohorts born up to about 1930 moved from farm origins into manual work—following what might perhaps be taken as a characteristic tendency of the drive to ‘mature’ industrialism. But in later cohorts this tendency is clearly not sustained, although for Sweden there is probably some underestimation of the rate in question as the converse of the distortion noted in regard to farm immobility. For England, Northern Ireland and Scotland, the graphs undulate in no readily interpretable way. Turning to the outflow rates from farm origins to nonmanual destinations presented in Figure 4, we find that trendless change is here still more manifest. The most remarkable feature of the graphs displayed is indeed their flatness, apart from the early rise from a near-zero level in the Irish case.¹³

The remaining point to be observed from Figures 3 and 4 together is that we find little

indication at all of national mobility rates converging. Over the period covered, the cross-national range for farm-to-manual outflows shifts upwards, but with little narrowing, from around 10 to 55 per cent to 30 to 70 per cent; while the rates for farm-to-nonmanual outflows are notable for being almost entirely confined within a range of 10 to 25 per cent.

The last two Figures in the series, 5 and 6, show changes in rates of intergenerational mobility between the broad manual and nonmanual classes that we distinguish. From inspection of the graphs, it would once again seem difficult to avoid the conclusion that no clear trends emerge. Although some impression may perhaps be given that, overall, mobility from manual origins to nonmanual destinations has decreased while that in the reverse direction has increased, it is in fact only in the Polish case that monotonic trends in these directions can be found. In general, fluctuating rates are displayed, and the graphs for different nations frequently cross. Moreover, as earlier remarked, the negligible volume of outflows from both manual and nonmanual origins to farm destinations means that the plots of Figures 5 and 6 can be taken as essentially the obverses of those relating to *immobility* within our manual and nonmanual classes; so in the case of these rates too an absence of trends may be claimed.

Finally, Figures 5 and 6 again fail to provide evidence of cross-national convergence in mobility rates. Over the period covered, the cross-national range for rates of mobility from manual origins to nonmanual destinations narrows only slightly as it falls from around 30 to 55 per cent for the oldest cohorts down to 20 to 40 per cent for the youngest; and the range for mobility in the reverse direction shows no narrowing at all in moving from 20 to 45 up to 30 to 55 per cent.¹⁴

It is the results contained in these last two figures—and also in Figure 4—that may occasion the most surprise among those presented

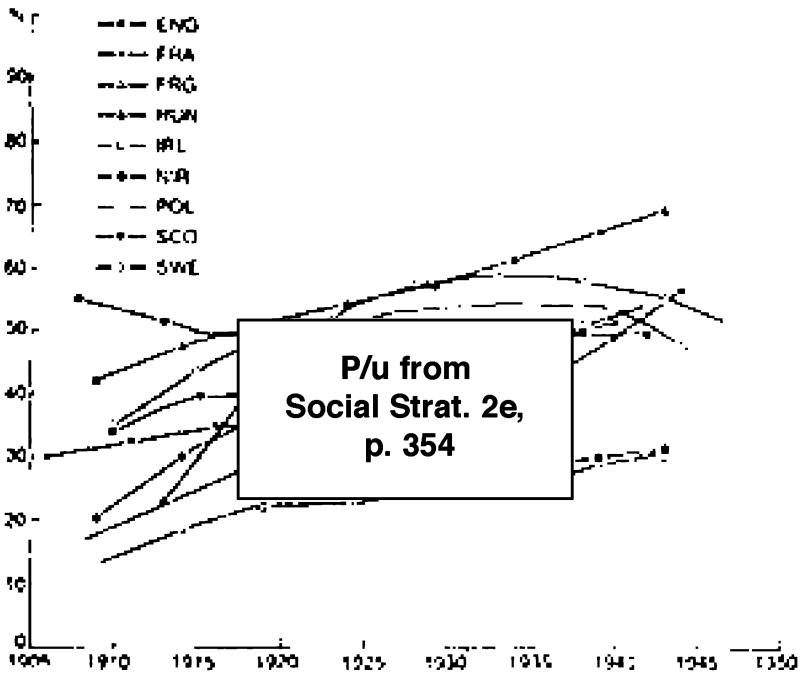


Figure 3. Outflow Rates from Farm Origins to Manual Destinations for Men in Nine Nations by Birth Year

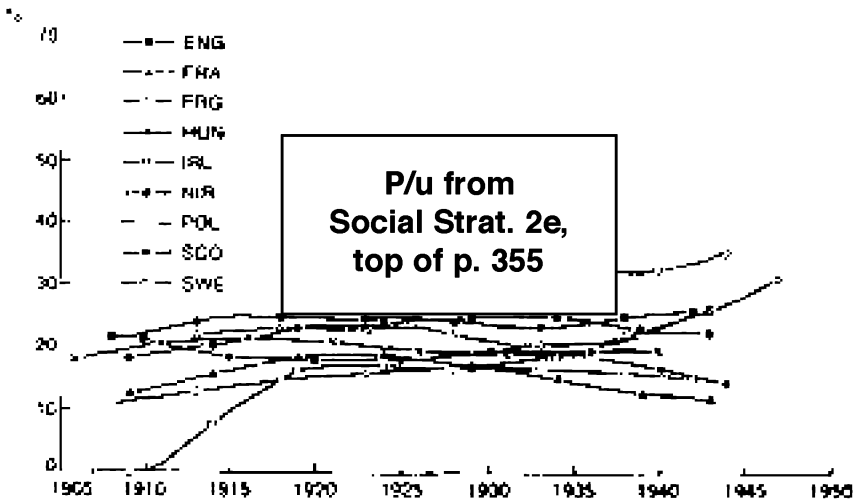


Figure 4. Outflow Rates from Farm Origins to Non-manual Destinations for Men in Nine Nations by Birth Year

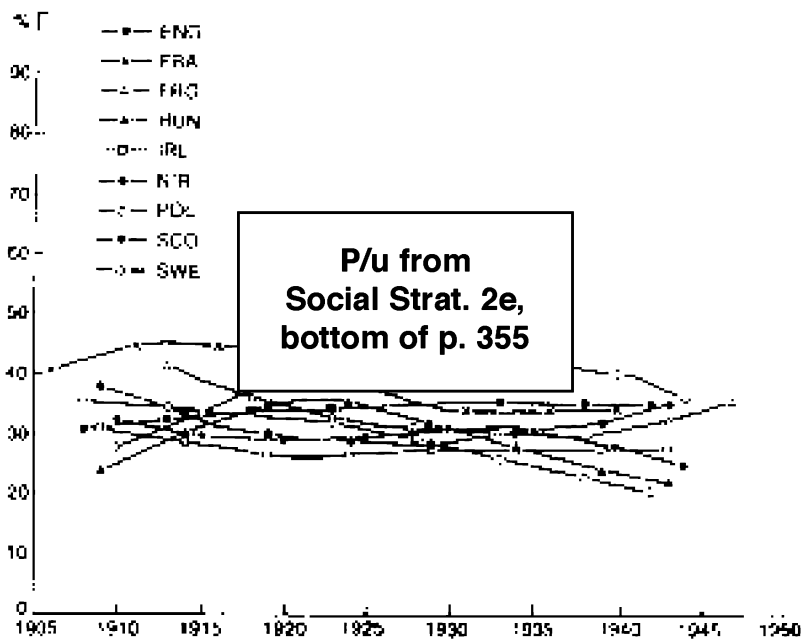


Figure 5. Outflow Rates from Manual Origins to Non-manual Destinations for Men in Nine Nations by Birth Year

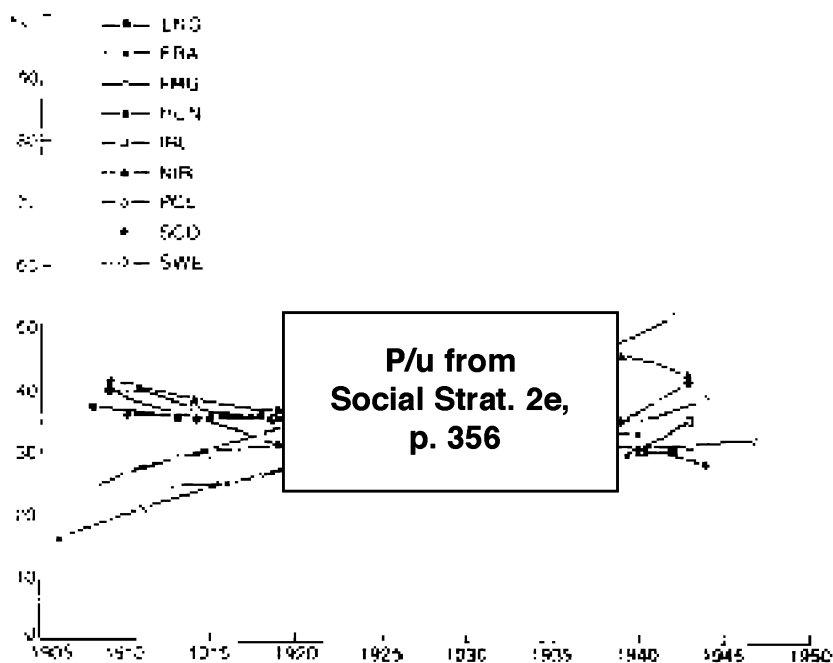


Figure 6. Outflow Rates from Non-Manual Origins to Manual Destinations for Men in Nine Nations by Birth Year

so far. It would be generally accepted that nonmanual work tends to grow and manual work to contract as industrial societies reach the more advanced stages of their development—regardless of whether this is seen as contributing to a net degrading or upgrading of the employment structure overall. And thus, within the context of the three-class version of our schema, increasing mobility into nonmanual destinations from farm and manual origins alike should be ‘structurally’ favoured. Yet, in our data, no consistent indication of such tendencies is to be found, even within the more advanced nations or among the younger cohorts.

However, it must in this connection be noted that our nonmanual class is very widely defined. It includes some groupings, such as routine nonmanual employees in administration, commerce and services, which have grown primarily through the greater workforce participation of women; and others, such as small proprietors and other self-employed workers, which, over the period that our data cover, were typically in decline. If, then, we wish to consider mobility flows into the nonmanual groupings that have most clearly expanded among the male workforce, which are in fact mostly at the higher levels of the white-collar range, we need to draw on data from mobility tables of a more elaborated kind. Although, as we have explained, we cannot go over entirely to the seven-class version of our schema, we can, for present purposes, make a useful compromise: we can construct mobility tables that apply the seven-class version to destinations while retaining the three-class version for origins.

In Figures 7 and 8 we show the course followed by two outflow rates derived from such 3×7 tables: that is, outflow rates from farm and from manual origins respectively into Class I+II of the seven-class version—the service class of primarily professional, higher technical, administrative and managerial employees. In other words, we here focus

on subsets of the rates presented in Figures 4 and 5 where the mobility in question is into types of employment that *have* been in general expansion. Furthermore, we can also in this way examine changes in mobility flows which, in the light of the hierarchical divisions that we make within our class schema, could be regarded as representing mobility *upwards* from less to more advantaged class positions.¹⁵

A preliminary point to be noted about Figures 7 and 8 is that, because the rates we are here concerned with are generally lower and less differentiated than those presented in previous figures, we have doubled the vertical scale, thus of course ‘enlarging’ the changes that are depicted. Even so, they do not appear as highly dramatic.

In Figure 7, some increase in mobility from farm origins into the service class is shown among younger cohorts in several nations—that is, in England, the FRG and Sweden and, more weakly, in Ireland. But in the remainder any increase that can be detected occurs among older cohorts, and the trend then fades out—so that cross-national differences in rates in fact *widen*. In Figure 8, rates of mobility from manual origins into the service class likewise display an increase in England and Sweden which is at least held, and in the FRG they turn sharply upwards among the youngest cohorts. But again, too, the other nations differ, with the most common tendency being for rates first to rise but then at some later point to decline. It is of interest to note that this tendency is most marked in Figure 8, and also appears in Figure 7, in the graphs for our two eastern European nations. Upward mobility into the service class among Hungarian and Polish men born in the early 1920s—who would reach occupational maturity in the period of post-war ‘socialist reconstruction’—rose to a level that later cohorts quite fail to match.

Where, among younger cohorts, the graphs of Figures 7 and 8 level out or turn down, we

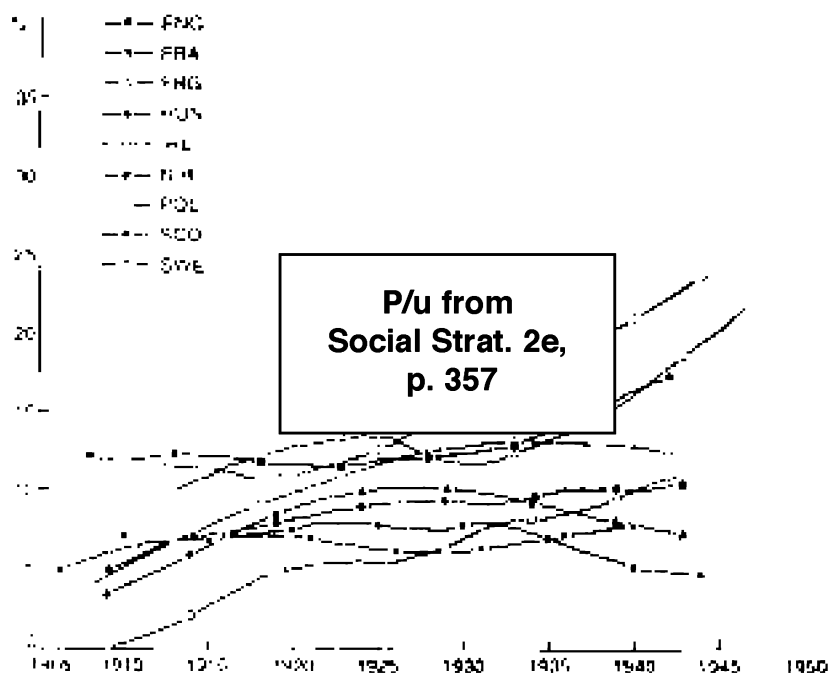


Figure 7. Outflow Rates from Farm Origins to Service Class Destinations for Men in Nine Nations by Birth Year

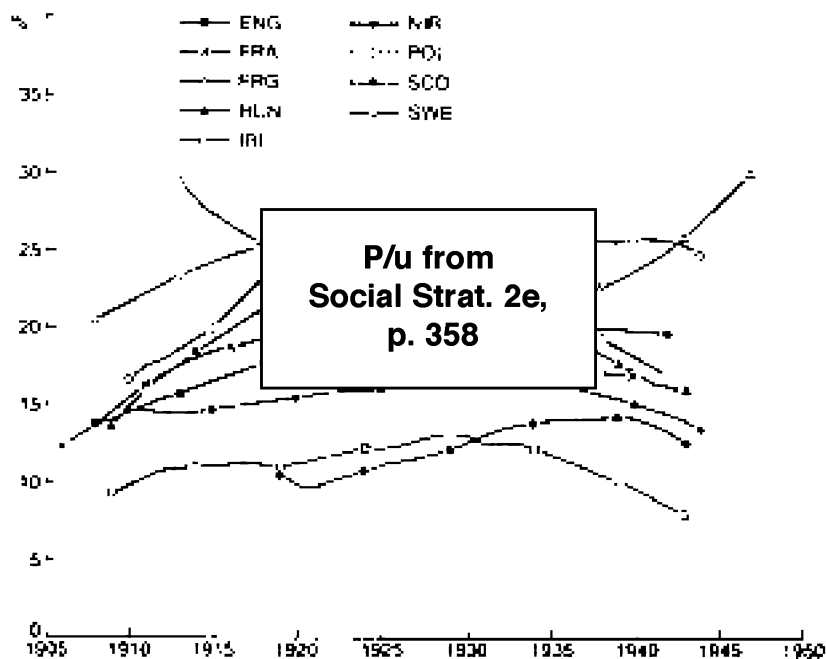


Figure 8. Outflow Rates from Manual Origins to Service Class Destinations for Men in Nine Nations by Birth Year

cannot preclude the possibility that this is in some part the result of age effects—so that in fact, as the men in these cohorts become older, a larger proportion will enter into service-class positions. However, what we would doubt is that such effects are likely to an extent that would make the graphs seriously misleading. There are good empirical grounds for supposing—and the liberal theory would certainly predict—that the younger the men in our samples, the more probable it is that they will achieve upward mobility through education, so that this mobility will be apparent at a relatively early stage in their working lives. And in this connection, it is then further relevant to emphasise that the graphs corresponding to those of Figures 7 and 8 which depict mobility to class of first employment similarly fail to reveal consistently rising trends. It emerges rather that even if one considers only men born from the 1920s onwards, a steady increase in mobility into the service class from farm and manual origins alike is found in just one nation—namely, Sweden.

What, we believe, should be emphasised here is that although the service class does show a general tendency to expand within modern societies, it does not do so at a steady pace simply in response to the exigencies of industrial development. Periods of relatively rapid growth and of stasis will alternate under the influence of other factors, not least political ones, thus producing rises and falls in rates of upward mobility into this class of a kind which the Hungarian and Polish cases do no more perhaps than exemplify at their most striking.¹⁶

How, then, can we best sum up the foregoing findings in regard to trends in absolute rates? To begin with, we can say that our investigation of outflow rates points to certain trends which seem likely to have occurred in most, if not all, of the nations we consider at *some* stage of their industrial development:

- (i) a decline in intergenerational immobility within the farm sector;

- (ii) an increase in mobility from farm origins into manual employment in industry; and
- (iii) some upturn in mobility from farm and manual origins into service-class positions.

To this extent, therefore, we might stop short of an extreme antinomianism of the kind that Sorokin could be taken to represent, and recognise that industrialism does carry typical implications for the direction of several broad mobility flows—as indeed Carlsson (1963) argued some time ago in a critique of Sorokin's position.

However, we must then also say that the trends that we are able to identify have rarely appeared as continuous over the period to which our data refer; that from nation to nation their phasing within the developmental process evidently differs; and moreover that changes in all other rates that we have examined would appear to be essentially directionless. In turn, therefore, we can produce little evidence that within our European nations mobility rates overall are moving steadily towards some relatively well defined 'industrial' pattern in the way that the liberal theory would suggest—and certainly not towards one that is characterised by steadily rising mobility. The outflow rates that we have considered show no clear tendency to converge; and, whatever course they may follow, they are not associated with any consistent upward trend in total mobility rates. In the light of the evidence presented, and in particular of that of Figure 1, it could not be claimed that men in our national samples whose working lives began, say, after the second world war have generally lived in more mobile societies than those who first entered employment in the 1920s.

The one possible qualification that might here be made is suggested by our findings for Ireland and Poland. These could be taken as meeting the expectation that an upward shift in the level of total mobility will occur in the

course of industrialisation where a rapid decline in agricultural employment or, more specifically perhaps, the break-up of a predominantly peasant economy goes together with a rising demand for industrial labour. But even if, as Lipset and Zetterberg would hold, a shift of this kind can be regarded as a general characteristic of the emergence of industrial societies, at least two further points would still, on our evidence, need to be made. First, the upturn in mobility has to be seen as being of a delimited, once-and-for-all kind; and secondly, it still leaves ample room for subsequent variation in absolute mobility rates and patterns as nations proceed to the further stages of their development.¹⁷

Relative Rates

We move on now to consider trends in inter-generational class mobility from the standpoint of relative rates, which we shall treat in terms of odds ratios. That is, ratios which show the relative odds of individuals in two different classes of origin being found in one rather than another of two different classes of destination; or, alternatively, one could say, which show the degree of *net* association that exists between the classes of origin and of destination involved. Two further points concerning such ratios may here be noted.

First, the total set of odds ratios that can be calculated within a mobility table may be taken as representing the 'endogenous mobility regime' or 'pattern of social fluidity' that the table embodies. And since odds ratios are 'margin insensitive' measures, it is then possible for this underlying regime or pattern to remain unaltered as between two or more mobility tables even though their marginal distributions differ and, thus, all absolute rates that can be derived from them. Further, we might add, it would be possible for relative rates as expressed by odds ratios to differ from table to table in some systematic way without this being readily apparent from an inspection of absolute rates.

Secondly, odds ratios are the elements of loglinear models; and, thus, where these ratios are taken as the measure of relative mobility rates, hypotheses about the latter can be presented and formally tested through the application of such models. This is the approach that we shall here follow.

In examining relative rates, we shall attempt, in the same way as we did with absolute rates, to make inferences about the extent of changes over time from the mobility experience of successive birth-cohorts within our national samples. However, instead of here working with yearly cohorts, we distinguish four ten-year birth-cohorts, which can then be regarded, given the closeness of the dates of the inquiries from which the samples derive, as having more or less comparable locations within the broad sweep of recent European economic history.

The first—that is, the earliest—of these cohorts comprises men aged 55 to 64, who were thus mostly born in the first two decades of the century, and who entered employment before or during the inter-war depression years. The second is of men aged 45 to 54, the majority of whom were born in the 1920s and entered employment in the later 1930s or the war years. The third is that of men aged 35 to 44, who were born between the late 1920s and the early 1940s, and whose working lives have fallen very largely within the post-war period. And finally the fourth cohort is that of men aged 25 to 34, who were born from the end of the 1930s onwards and who mostly entered employment while the long boom was in train. This last cohort is clearly made up of respondents who could not be generally assumed to have reached a stage of occupational maturity, and this we must recall where it might be relevant to the interpretation of our results.

When we thus divide our national samples into cohorts, we again have a potential problem of unduly low cell counts in the mobility tables for each cohort. To obviate this, we base

our analyses throughout on the five-class version of our schema (cf. Appendix Table 2).

As regards relative rates, there is, as we have noted, an obvious opposition between the expectations that follow from the liberal theory and those that may be derived from the FJH hypothesis. According to the former, a tendency should be found in the course of the development of industrial societies for relative rates to become more equal—or, one could say, for all odds ratios to move closer to the value of 1, which signifies the complete independence of class origins and destinations or ‘perfect mobility.’ According to the latter, relative rates will be basically the same across all societies that have market economies and (at least) nuclear family systems, whatever stage their level of industrial development may have reached; and thus, when examined over time within particular industrial societies, relative rates should reveal little change at all.

We may then start off from an attempt at evaluating these rival positions, and, to this end, we introduce a rather simple loglinear model which is, however, able to provide a direct representation of expectations under the FJH hypothesis, at least if this is taken *stricto sensu*. This model, which we have earlier labelled the ‘constant social fluidity’ (CnSF) model (Goldthorpe, 1980, 1987: ch. 3; Erikson, Goldthorpe and Portocarero, 1983) may, for present purposes, be written as

$$\log F_{ijk} = \mu + \lambda_i^O + \lambda_j^D + \lambda_k^C + \lambda_{ik}^{OC} + \lambda_{jk}^{DC} + \lambda_{ij}^{OD} \quad (1)$$

where F_{ijk} is the expected frequency in cell ijk of a three-way table comprising class of origin (O), class of destination (D) and cohort (C) and, on the right-hand side of the equation, μ is a scale factor, λ_i^O , λ_j^D and λ_k^C represent the ‘main’ effects of the distribution of individuals over origins, destinations and cohorts respectively, and the remaining terms represent the effects for the three possible two-way associations in Appendix Table 2.

Thus, the model entails a number of substantive propositions, most of which are unproblematic: for example, that an association exists between class of origin and class of destination; and that further associations exist between class of origin and cohort and between class of destination and cohort—in other words, men in different cohorts have different origin and destination distributions. It is, however, a further proposition that is critical. Since no *three-way* association is provided for in the model (the λ_{ijk}^{ODC} term does not appear), it is also entailed that the level of association between class of origin and of destination is *constant across cohorts* or, one could alternatively say, that over the mobility tables for successive cohorts all corresponding relative rates, as measured by odds ratios, are identical.

We can then consider our nine nations separately, and in each case fit the above CnSF model to a three-way table which comprises the five classes of origin, five classes of destination and four cohorts that we propose to distinguish. The results of so doing are presented in Table 1.

In this table we also report the results of applying a model that represents the hypothesis of the (conditional) independence of class origins and destinations; that is, the CnSF model minus the λ_{ij}^{OD} term. We do not expect this independence model to fit the data—and, as can be seen, in no case does it—but it serves as a useful baseline, by reference to which we can assess, through the rG^2 statistic in the fourth column of the table, how much of the total association between class of origin and class of destination the CnSF model is able to account for.¹⁸

Also in addition to the more usual ‘goodness of fit’ statistics, we give in the last column of Table 1 values for the statistic $G^2(S)$. This we introduce here to attempt to deal with a difficulty arising from the large variation in the sizes of our national samples. In consequence of this variation, our national mobility tables differ quite widely in their

Table 1. Results of Fitting the CnSF Model to Intergenerational Class Mobility for Four Birth-Cohorts

| Model ^a | G ² | df | p | (G ²) | Δ ^b | G ² (N) (1,746) |
|--------------------|----------------|----|-------|-------------------|----------------|-------------------------------|
| ENG | | | | | | |
| (N = 8,343) | | | | | | |
| OC DC (cont. ind.) | 1,095.0 | 64 | 0.00 | — | 16.1 | 405 |
| OC DC OD (CnSF) | 53.1 | 48 | 0.28 | 96.9 | 2.6 | 49 |
| FRA | | | | | | |
| (N = 16,431) | | | | | | |
| OC DC | 8,370.6 | 64 | 0.00 | — | 24.7 | 734 |
| OC DC OD | 96.7 | 48 | 0.00 | 98.5 | 2.0 | 53 |
| FRG | | | | | | |
| (N = 3,570) | | | | | | |
| OC DC | 1,092.0 | 64 | 0.00 | — | 21.2 | 567 |
| OC DC OD | 81.9 | 48 | 0.00 | 97.5 | 4.4 | 65 |
| HUN | | | | | | |
| (N = 10,319) | | | | | | |
| OC DC | — | — | — | — | 19.2 | 457 |
| OC DC OD | — | — | — | — | 2.4 | 52 |
| IRL | | | | | | |
| (N = 1,746) | | | | | | |
| OC DC | 902.3 | 64 | 0.00 | — | 20.2 | 902 |
| OC DC OD | 60.2 | 48 | 0.11 | 93.3 | 5.2 | 69 |
| NIR | | | | | | |
| (N = 1,948) | | | | | | |
| OC DC | 780.6 | 64 | 0.00 | — | 25.5 | 756 |
| OC DC OD | 44.5 | 48 | >0.50 | 94.3 | 5.0 | 45 |
| NOI | | | | | | |
| (N = 27,993) | | | | | | |
| OC DC | 7,357.7 | 64 | 0.00 | — | 19.6 | 519 |
| OC DC OD | 66.7 | 48 | 0.04 | 99.3 | 1.4 | 49 |
| SCG | | | | | | |
| (N = 3,985) | | | | | | |
| OC DC | 1,146.6 | 64 | 0.00 | — | 18.1 | 538 |
| OC DC OD | 66.3 | 48 | 0.04 | 94.2 | 4.4 | 56 |
| SWI | | | | | | |
| (N = 1,882) | | | | | | |
| OC DC | 405.9 | 64 | 0.00 | — | 17.3 | 379 |
| OC DC OD | 45.2 | 48 | >0.50 | 88.8 | 5.3 | 45 |

Notes

^a O = origin class, D = destination class, C = cohort

^b (G²) shows the percentage reduction in the G² for a model taken as baseline (here the conditional independence model) that is achieved by a more complex model (here the CnSF model). For further discussion of this statistic, see p. 18.

^c Δ is the dissimilarity index, showing the percentage of all cases in the table analysed that are misclassified—that is, allocated to the wrong cell—by a particular model.

P/u from
Social Strat. 2e,
p. 362

capacity to show up as statistically significant relatively small deviations from models that we fit to them. It is as if we were looking at slides through microscopes of greatly differing power: we have the possibility of seeing far more detail in some cases than in others. Thus, there is the evident danger that we do not evaluate a model in an evenhanded way from nation to nation. We could, for example, be led to reject a model in the Polish case on account of deviations which, were they present also in the case of, say, Ireland, we would simply not observe. Thus, we evidently need some measure of goodness of fit that is standardised by sample size. One possibility would be to take G^2/N . However, we prefer, as a more refined measure, Schwartz's suggestion of $G^2(S)$ which is given by $((G^2 - df)/N) \times K + df$, where K is the sample size which is to be taken as standard.¹⁹ We will follow the conservative practice of setting K equal to the size of the *smallest* of the national samples with which we are concerned—and thus, in Table 1, at 1746. To help remind the reader of the hypothetical nature of $G^2(S)$ —that it is the χ^2 value that we would expect from a sample of size K , all other things being equal—we report it only to the nearest integer.

What, then, can we learn from the content of Table 1? It would in fact appear that the CnSF model performs fairly well. It is true that the p values reported indicate that in only four of the nine nations—England, Ireland, Northern Ireland and Sweden—would one retain this model, taken as the null hypothesis, according to the conventional 0.05 criterion. However, it is also evident that much of the variation in the G^2 and p values returned is attributable to differences in sample size. When one examines the $G^2(S)$ values in the final column of the table, one finds that these are in fact contained within a rather narrow range, and moreover that in no case do they exceed the 65 mark, thus implying, with $df = 48$, p values of less than 0.05. In other words, if we were restricted throughout

to a sample size of 1746, such as that we have for Ireland, we would find it difficult to reject the CnSF model for any nation—although the FRG would have to be regarded as a borderline case.

What could therefore be claimed on the basis of the foregoing is that while significant deviations from the CnSF model do have to be recognised in at least some of our nations, such deviations would not appear to be at all substantial. In this connection, it is of further relevance to note that in all cases but one the CnSF model accounts for more than 90 per cent of the total association existing between class of origin and of destination—the exception being Sweden, where the independence model fits least badly; and again, that within the different national mobility tables the CnSF model leads to the misclassification of, at most, only a little over 5 per cent of all cases.²⁰ . . .

Conclusions

We have sought in this chapter to use data from European nations in order to evaluate various arguments concerning mobility trends within industrial societies. The major outcome, it might be said, has been a negative one: that is, considerable doubt has been thrown on claims associated with what we have called the liberal theory of industrialism. We have found no evidence of general and abiding trends towards either higher levels of total mobility or of social fluidity within the nations we have considered; nor evidence that mobility rates, whether absolute or relative, are changing in any other consistent direction; nor again evidence that such rates show a tendency over time to become cross-nationally more similar. The most that could be said on the side of arguments proposing some linkage between industrial development and increased and more standardised mobility rates would be that structural changes—most importantly, the decline in agriculture—appear likely to

generate upturns in total, and also perhaps in certain outflow rates over periods of limited duration and of very variable phasing.

Such results are all the more damaging to the liberal theory since, as we earlier emphasised, Europe over the middle decades of the twentieth century, and above all in the post-war era, provides a context in which the theory should have every chance of showing its force. Furthermore, we may reiterate the point that any distortions in our findings that derive from our reliance on (quasi-) cohort analysis, and in particular from the confounding of age and period effects, are unlikely to be ones that tell unfairly against liberal claims: if anything, the contrary should be supposed.

We would therefore believe that the attempt to represent changes in mobility rates in modern societies as displaying regular developmental patterns, driven by a functional logic of industrialism, is one that faces serious empirical difficulties; and, in turn, we would argue that the need must be recognised to search for ways in which these changes might be more satisfactorily understood. In this connection, there is then one further outcome of the analyses of the present chapter which should, in our judgment, be seen as having major significance: namely, that while it appears that liberal expectations of directional tendencies in absolute and relative mobility rates are in both respects largely unfounded, *the nature of the contrary evidence is quite different from one case to the other: with absolute rates it is evidence of trendless, though often quite wide, fluctuation, but with relative rates it is evidence of considerable stability.* What is thus suggested is that in attempts to go beyond the liberal theory, the treatment of absolute and of relative rates is likely to set quite different kinds of problem and of analytical task.

As regards absolute rates, liberal expectations most obviously fail, we would suggest, because changes in structural influences on mobility do not themselves have the regular-

ity that liberal theorists have been wont to suppose. Analyses of economic growth advanced in the 1950s and 1960s by authors such as Clark (1957), Rostow (1960) and Kuznets (1966) were taken as demonstrating clear sequences of change in the sectoral and occupational, and hence in the class, composition of labour forces. However, it would by now be widely accepted that, whatever theoretical insights the work of these authors may provide, it does *not* allow one to think, at a historical level, in terms of a well-defined series of developmental stages through which the structure of the labour forces of different nations will pass in turn as industrialisation proceeds. While certain very general tendencies of change may in this respect be identified, considerable variation still prevails from case to case in relation to the speed with which change occurs and the extent to which different aspects of change are separated in time or overlap (Singelmann, 1978; Gagliani, 1985).

The European experience of industrialisation, which has provided the setting for our analyses in this paper, itself well illustrates the variety of paths that the development of labour forces may follow; and it does, moreover, bring out the diversity of the causal factors at work here—by no means all of which can be plausibly seen as part of some englobing developmental process. Thus, the historical formation of national class structures has to be seen as reflecting not only early or late industrialisation but, in addition, important influences stemming, on the one hand, from the international political economy and, on the other, from the various strategies pursued by national governments in response to both external and internal pressures. To take but one example here, the contraction of agriculture—which we have found to play a major part in the pattern of change in absolute class mobility rates—cannot be understood, as it has occurred in particular cases, simply in terms of the shifting marginal productivity of sectors and differences in the elasticities of demand for their products

that the theory of economic growth would emphasise. As the agrarian histories of our nations can amply show (see, e.g., Priebe, 1976), the pace and timing of agricultural contraction also—and often far more decisively—reflects whether nations were at the centre or on the periphery of international trading relations; in a position of economic dominance or dependence; and, further, the policies that their governments adopted towards agriculture both in regard to its social organisation and its protection against or exposure to market forces.²¹

Once, therefore, the variability and complexity of the determination of the structural contexts of mobility is appreciated, the extent to which the movement of absolute rates over time appears as merely trendless can no longer be found especially surprising. If changes in such rates do largely express the shifting conjunctures of a diversity of exogenous effects, then ‘trendlessness’ as suggested by Sorokin is indeed what must be expected. It is noteworthy that it is essentially an argument on these lines that has been pursued by the several European economic and social historians who have sought to join in the sociological debate. In rejecting ‘the idea of a sustained growth in social mobility during industrialization’, these authors have emphasised the ‘multitude of factors’ which affect mobility levels; and, in place of developmental stages, they have sought rather to establish empirically a number of different ‘eras’ or ‘phases’ of both rising *and falling* mobility within the period in which European industrialisation has occurred (see esp. Kaelble, 1984: 490; also Kaelble, 1981; Mendels, 1976; Kocka, 1980).

Thus, we would maintain, the crucial issue that arises so far as absolute rates of mobility are concerned is that of whether, or how far the course of change they follow is in fact a phenomenon open to explanation in macrosociological terms. Investigators who have been impressed by the degree of temporal variation in absolute, as compared with relative, rates have gone on to conclude that the

dynamism of the former must lie primarily in structural effects, and in turn they have urged that these should not be treated as merely a ‘nuisance factor’ but should become themselves the focus of inquiry (e.g. Hauser *et al.*, 1975; Grusky and Hauser, 1984; Goldthorpe, 1985). However, while this argument has an evident logic, it does leave quite undecided the question of just what *kind* of understanding of structural effects—and thence of change in absolute rates—it might be possible to achieve. In so far as generalisations about such effects can be made, will they prove to be of any great explanatory value when applied to particular instances? Or may one be in this respect forced back willy-nilly to a reliance largely on specific historical descriptions—as a position such as that of Sorokin would in effect imply? Or again are there perhaps intermediate possibilities?

Turning now to relative rates, we meet with a very different situation. In this case, the liberal theory is undermined because, instead of the anticipated trend of change, in the direction of greater equality, we find evidence of an essential stability. Although shifts in relative rates can in some cases be detected (see Erikson and Goldthorpe 1992: 90–101), these are not only ones which go in various directions but, more importantly, ones which, as against those observed in absolute rates, are of very limited magnitude—so that one might wish to speak more of ‘oscillation’ than of fluctuation. In other words, the liberal theory would here appear to fail because the logic of industrialism has not in fact automatically generated the changes within processes of social selection which were expected of it, and through which a steady increase in fluidity and openness would be promoted.

It is in this connection of interest to note that of late exponents of the liberal theory appear to have modified their position in regard to relative rates quite significantly. Thus, for example, Treiman initially sought to provide the hypothesis of a trend towards greater openness with a rationale largely in terms of the

functional exigencies of industrialism (1970: 218). However, in a recent paper (with Yip), he puts much stronger emphasis on the part that is played in creating greater openness and equality of opportunity by the more proximate factor of greater *equality of condition*—that is, by a greater equality in the economic, cultural and social resources that families possess. And while it is still maintained that this increase in equality of condition itself ultimately derives from the development of industrialism, it is at the same time accepted that ‘industrialization and inequality do not move in perfect concert’ and, further, that *other* factors, especially political ones—for example, whether a nation has a socialist regime—may also affect the degree of inequality that exists (Treiman and Yip, 1989: 376–7). That is to say, it would here seem to be recognised that even in cases where a trend towards greater fluidity may be empirically established, this cannot be regarded as simply a matter of developmental necessity but must rather be explained as the contingent outcome of quite complex patterns of social action (cf. also Ganzeboom, Luijkx and Treiman, 1989; Simkus *et al.*, 1990). And conversely, this revised, and evidently much weaker, position is then of course able to accommodate the alternative possibility that, in particular instances, no trend of this kind is observed—because countervailing forces have in fact proved too strong.

The stability in relative rates that we have shown gains in significance, we may add, not only because the period that our data cover comprised decades of unprecedented economic growth but also because it was, of course, one of major political upheavals, in which, in the train of war and revolution, national frontiers were redrawn and massive shifts of population occurred.²² The fact that the relative rates underlying the mobility experience of cohorts within our national samples should then reveal so little change—whether directional or otherwise—becomes all the more remarkable. While we have not been able to support the

claim of a sustained developmental trend, we have, it appears, found indications of something of no less sociological interest: that is, of a constancy in social process prevailing within our several nations over decades that would in general have to be characterised in terms of the transformation and turbulence that they witnessed.

Furthermore, this finding is, as we have indicated, one which may be related to a larger sociological argument, namely, that represented by the FJH hypothesis. We have presented analyses which indicate that *some* variation in fluidity patterns does in fact occur among nations—indeed, more than within nations over time—and also that this variation shows no tendency to diminish (see Erikson and Goldthorpe 1992: 90–101). Thus, expectations of convergence are not met. However, neither would cross-national variation appear to be increasing; and, more importantly, as we have elsewhere sought to show (Erikson and Goldthorpe 1987, 1992: ch. 5), it could not be reckoned as sufficiently wide to rule out the possibility that the ‘basic’ similarity in relative rates that the FJH hypothesis claims is the major source of the temporal stability that we have observed; or, that is, the possibility that *constancy* above all reflects *commonality*.

In other words, in so far as the degree of similarity proposed by the FJH hypothesis is established, we may think of temporal shifts in fluidity within nations as being no more than oscillations occurring around the standard pattern that the hypothesis implies or, at all events, as being restricted in their frequency and extent by whatever set of effects it is that generates this pattern. And in this regard, then, the ultimate task becomes that of understanding these effects; or, that is, of seeking to explain not variance, to which, as Lieberman (1987) has observed, analytical strategies within macrosociology have thus far been chiefly oriented, but rather a *lack* of variance—for which, unfortunately, appropriate strategies remain largely to be devised.

Appendix Table 1. National Inquiries Used as Data Sources

| | Survey | Date | Reference for survey details |
|-----------------------------------|---|-----------|-------------------------------------|
| England & Wales (UK) | Oxford National Occupational Mobility Survey | 1971 | Goldthorpe (1992) |
| France (FR) | INSEE Enquête Française sur les Conditions Professionnelles | 1971 | Taylor, Theka and Janssen (1975) |
| Federal Republic of Germany (FRG) | ZfM | 1978 | Trampusch <i>et al.</i> (1989) |
| Hungary (HU) | Social Class | | Andriessen and Zagzecki (1980) |
| Irish Republic (IR) | Deer | 1971 | O'Malley, Scully and Wiggins (1977) |
| Netherlands (NL) | De Ontwikkeling van de Occupatie en de Inkomens | 1975-1981 | O'Malley, Scully and Wiggins (1977) |
| Poland (PL) | Change — The Social Experiment in Structure | 1977 | Zagzecki (1977-8) |
| Scotland (SC) | Scottish Mobility Survey | 1974-1975 | Parton (1987) |
| Sweden (SW) | Level of Living Survey | 1975 | Andriessen (1980) |

P/u from
Social Strat. 2e,
p. 366

NOTES

1. This paper is based on chapter 3 of Robert Erikson and John H. Goldthorpe, *The Constant Flux: A Study of Class Mobility in Industrial Societies*, The Clarendon Press, Oxford, 1992. The research on which this book reports was carried out under the auspices of the CASMIN-Projekt, based at the Institut für Sozialwissenschaften of the University of Mannheim and funded by grants from the Stiftung Volkswagenwerk, Hanover. Readers are referred to the above work (chapter 2 esp.) for full details of the comparative methodology followed in research.

2. The influence here of the 'stages-of-growth' model of Rostow (1960: see esp. Chart 1) would seem to be of particular importance and also, perhaps—though the evidence is indirect—the interpretation of European industrial development provided by Landes (1957, 1965, 1972), which places major emphasis upon the rate and pattern of diffusion of techniques of production from Britain to the more 'backward' economies of the European mainland.

3. Thus, for example, in Kerr *et al.* (1960) discussion of France and Italy is largely concerned with the impediments to industrial development that result from the persisting importance of 'family-dominated enterprises' with 'patrimonial management' (see e.g. pp. 80, 141-2; and cf. Landes, 1957); and discussion of Germany, with difficulties of social rigidity and authoritarianism, following from the promotion of industrialisation by a dynastic elite (e.g. pp. 54-5, 150-1).

4. It is of interest that Bell should refer to the situation on which he comments as one 'that has gone relatively unexamined'. This statement may well be

true for American theorists of industrialism, but it can scarcely hold in the case of European economic and social historians. See, for example, the discussion of issues central to the 'reperiodisation' of the development of industrial society in Europe that are found in Wrigley (1972) and Mayer (1981).

5. So far as emigration is concerned, a detailed review of the possible and likely effects on mobility rates and propensities is provided in Hout (1989), with special reference to the Irish case.

6. For our present purposes, it is the confounding of period by age effects that is most likely to create problems. To the extent that cohort effects are present in the data, this may be regarded as valid evidence against the occurrence of secular trends.

7. This is in fact the highest maximum age that we could apply across all nine of our national samples.

8. The wording of the questions from which this information was derived varied somewhat from one national inquiry to another but not, we believe, in ways likely to have any significant effects on the comparability of data. In this and all similar instances full details of question wording, construction of variables, etc. are to be found in the documentation to the CASMIN International Social Mobility Superfile (Erikson *et al.*, 1988).

9. We are greatly indebted to Jan Hoem for his most generous help in this aspect of our work.

10. These plots are not shown but in what follows it may be assumed that where no reference is made to rates of mobility from class origins to class of first employment, our findings in this respect would not lead us seriously to qualify those we have obtained for rates from class origins to present class.

Appendix Table 2. The Class Schema

| Full version | | Condensed version | | Three class | |
|--------------|--|-------------------|--|-------------|----------------------|
| | | Seven-class | | Five class | Three class |
| I | Higher grade professionals, administrators, and officials, managers in large industrial establishments, large proprietors | I+II | Service class, professionals, administrators and managers; higher grade technicians, supervisors of non-manual workers | I-III | White-collar workers |
| II | Lower grade professionals, administrators, and officials, higher grade technicians, managers in small industrial establishments, supervisors of non-manual employees | | | | |
| IIIa | Higher grade manual employees, higher grade craftsmen, higher grade employees | III | Non-manual manual workers | | Non-manual workers |
| IIIb | Lower grade manual employees, lower grade trades and services | | | | |
| IVa | Small proprietors, artisans, etc. with employees | | | | |
| IVb | Small proprietors, artisans, etc. without employees | IVa+IVb | P/u from Social Strat. 2e, p. 367 workers without employers | | Partly bourgeoisie |
| IVc | Farmers and smallholders, other self-employed workers in primary production | | | | |
| V | Lower grade technicians, supervisors of manual workers | V+VI | Skilled workers, lower grade technicians, supervisors of manual workers, skilled manual workers | V+VI | Skilled workers |
| VI | Skilled manual workers | | | | |
| VIIa | Semi-skilled and unskilled manual workers of non-agriculture, etc. | VIIa | Semi-skilled workers, semi- and unskilled manual workers (non-agriculture, etc.) | VIIa | Non-skilled workers |
| VIIb | Agricultural and other workers in primary production | | | | |
| VIII | Agricultural and other workers in primary production | VIII | Agricultural, fishermen, agricultural and other workers in primary production | | Farm workers |

11. The results that we report here for Hungary do of course depend on our treating workers on agricultural co-operatives or state farms as having a different class position (VIIb) from that of peasant proprietors (IVc). Some analysts of mobility in Hungary have not made this distinction; but we would argue the desirability of so doing, wherever it is practically feasible. It was, after all, precisely the aim both of the immediate post-war land reform and of the subsequent collectivisation programme to *change* agrarian class relations. In the Polish case, it should be noted, the attempt to collectivise agriculture that the regime launched at the end of the 1940s met with fierce peasant opposition and was finally abandoned in 1956 (cf. Lewis, 1973).

12. Although, then, we are here forced back to the obviously rather crude three-class basis of much earlier comparative research, we must stress that we still do

achieve a much higher standard of data comparability. As a result of our systematic recoding of the original unit-record data (see Erikson and Goldthorpe, 1992: ch. 2), we have a reasonable assurance that the categories of 'nonmanual', 'manual' and 'farm' are being applied in a consistent manner from nation to nation, rather than providing comparability of a merely nominal kind.

13. It may be noted that in Figure 4 the left tail of the curve for Sweden has been deleted. This is on account of its unreliability, as determined by a test developed by Hoem (see Erikson and Goldthorpe, 1992: ch. 3, Annex). For the same reason, we have also deleted the left tail of the curve for Northern Ireland in Figure 8.

14. These results are of direct relevance to the Lipset-Zetterberg hypothesis of cross-national similarity in

absolute rates, since this was in fact formulated in terms of outflow rates from nonmanual to manual positions and *vice versa*. We do not take up this issue here (but see further Erikson and Goldthorpe, 1992: ch. 6).

15. Following the hierarchical levels that we propose, a further upward flow—that our 3×7 tables do not enable us to distinguish—would be represented by men entering Class I + II positions from Class III origins.

16. Thus, for example, in the English case the more or less continuous rise in upward mobility into the service class across the cohorts we distinguish can be related to a corresponding steady expansion of this class from a time somewhere between 1931 and 1951 (there was no 1941 Census)—following, however, on several decades in which it grew scarcely at all (see Goldthorpe, 1980, 1987: ch. 2 esp.). As regards socialist societies, it may further be noted that evidence of a ‘parabolic’ curve for upward mobility, similar to that we record in Hungary and Poland, is also found for post-war Czechoslovakia in data from a survey conducted in 1984 (personal communication from Marek Boguszak and cf. Boguszak, 1990).

17. It would, moreover, be mistaken simply to equate a peasant economy—or society—with a ‘traditional’ one. Thus, while one may with justification speak of a peasant economy existing in substantial areas of Ireland at least up to the 1940s, many of its key institutional features—most importantly, perhaps, non-partible inheritance—were relatively new (cf. Hannan, 1979). The Irish peasant community, as classically depicted by Arensberg and Kimball (1940, 1968), has in fact to be seen as the historical product of economic and social conditions in Ireland following the Great Famine of 1846–9 and then of the land reform legislation introduced between 1870 and the First World War.

18. It is important that rG^2 , referred to by Goodman (1972) as the ‘coefficient of multiple determination’, should be interpreted within the particular context of loglinear modelling, rather than being taken as the equivalent of the perhaps more familiar R^2 of regression analysis. As Schwartz has pointed out (1985), the fact that R^2 s are typically much lower than rG^2 s reflects the fact that in regression the units of analysis are usually individuals while in loglinear modelling they are the cells of cross-tabulations and the scores are the numbers of individuals in a cell. Such aggregate data must then be expected to reveal stronger regularities than individual-level data. Schwartz’s summary (1985: 2–3) is apt: rG^2 ‘measures how adequately a model accounts for the observed *associations* among a pre-specified set of variables while R^2 and Eta^2 measure the amount of *variation* in one variable that can be ac-

counted for by its (linear) association with specified independent variables’. The point may be added that the substantive meaning of rG^2 will of course depend on the model that is chosen as baseline.

19. This suggestion was made to us by Joseph E. Schwartz in a personal communication, for which we are duly grateful.

20. We may add that results from equivalent analyses of data referring to mobility from class of origin to class of first employment are essentially similar. In only one case, that of Ireland, would the CnSF model be rejected on the basis of the $G^2(S)$ statistic; and again only in the Swedish case does the model not account for at least 90 per cent of the total origin-destination association, while at most only a little over 5 per cent of all cases are misclassified. It should, however, be recalled that we cannot undertake an analysis of the kind in question for the FRG, owing to lack of information on first employment.

21. Moreover, while we would believe that ‘demand side’ factors are generally of major importance in promoting structural change, ‘supply side’ ones may also have to be taken into account—for example, the effects of demographic change, including in- and out-migration, and of changes in the workforce participation rates of women and of different age-groups. And in these respects too political intervention may obviously play a crucial role.

22. Most importantly, in the aftermath of World War II the FRG was created out of the division of the Third Reich, and Poland’s frontiers were moved some 150–200 miles to the west—both changes being accompanied by large population movements. In addition, one may note the truncation of Hungary in 1920 (with the loss of almost 70 per cent of its area and 60 per cent of its population); and the partition of Ireland in 1920–2, following the War of Independence and the Civil War, so as to create the Irish Free State (which became the Irish Republic in 1949) and the six counties of Northern Ireland, a constituent element of the United Kingdom with, up to 1973, its own parliament and executive.

BIBLIOGRAPHY

- Andersson, L. (1987): ‘Appendix A: Sampling and Data Collection’ in R. Erikson and R. Åberg eds., *Welfare in Transition: A Survey of Living Conditions in Sweden 1968–1981*. Oxford: Clarendon Press.
- Andorka, R. and K. Zagórski (1980): *Socio-Occupational Mobility in Hungary and Poland*. Warsaw: Polish Academy of Sciences.

- Arensberg, C.M. and S.T. Kimball (1940, 2nd ed. 1968): *Family and Community in Ireland*. Cambridge, Mass.: Harvard University Press.
- Bell, D. (1980): 'Liberalism in the Post-Industrial Society' in *Sociological Journeys*. London: Heinemann.
- Blau, P.M. and O.D. Duncan (1967): *The American Occupational Structure*. New York: Wiley.
- Blossfeld, H.-P. (1986): 'Career Opportunities in the Federal Republic of Germany'. *European Sociological Review* 2.
- Boguszak, M. (1990): 'Transition to Socialism and Intergenerational Class Mobility: The Model of Core Social Fluidity Applied to Czechoslovakia' in M. Haller, ed., *Class Structure in Europe*. Armonk, N.Y.: Sharpe.
- Boyle, J.F. (1976): 'Analysis of the Irish Occupational Index'. Department of Social Studies, The Queen's University, Belfast.
- Braverman, H. (1974): *Labor and Monopoly Capitalism*. New York: Monthly Review Press.
- Brus, W. (1986a): 'Postwar Reconstruction and Socio-Economic Transformation' in M.C. Kaser and E.A. Radice, eds., *The Economic History of Eastern Europe, 1919–1975*, vol. 2. Oxford: Clarendon Press.
- Brus, W. (1986b): '1950 to 1953: The Peak of Stalinism' in M.C. Kaser, ed., *The Economic History of Eastern Europe, 1919–1975*, vol. 3. Oxford: Clarendon Press.
- Carchedi, G. (1977): *On the Economic Identification of Classes*. London: Routledge.
- Carlsson, G. (1963): 'Sorokin's Theory of Social Mobility' in P.J. Allen, ed., *Pitirim A. Sorokin in Review*. Durham, N.C.: Duke University Press.
- Clark, C. (3rd ed., 1957): *The Conditions of Economic Progress*. London: Macmillan.
- Crompton, R. and G. Jones (1984): *White-Collar Proletariat: Deskilling and Gender in Clerical Work*. London: Macmillan.
- Duncan, O.D. (1961): 'A Socioeconomic Index for All Occupations' in A.J. Reiss, ed., *Occupations and Social Status*. New York: Free Press.
- Dunlop, J.T., F.H. Harbison, C. Kerr and C.A. Myers (1975): *Industrialism and Industrial Man Reconsidered*. Princeton: Inter-University Study of Human Resources in National Development.
- Erikson, R. and J.H. Goldthorpe (1987): 'Commonality and Variation in Social Fluidity in Industrial Nations. Part I: A Model for Evaluating the "FJH Hypothesis"; Part II: The Model of Core Social Fluidity Applied'. *European Sociological Review* 3.
- Erikson, R. and J.H. Goldthorpe (1992): *The Constant Flux: A Study of Class Mobility in Industrial Societies*. Oxford: Clarendon Press.
- Erikson, R., J.H. Goldthorpe and L. Portocarero (1983): 'Intergenerational Class Mobility and the Convergence Thesis'. *British Journal of Sociology* 34.
- Erikson, R., J.H. Goldthorpe, W. König, P. Lüttinger and W. Müller (1988): 'CASMIN International Mobility Superfile: Documentation'. Mannheim: Institut für Sozialwissenschaften, University of Mannheim.
- Featherman, D.L., F.L. Jones and R.M. Hauser (1975): 'Assumptions of Social Mobility Research in the US: The Case of Occupational Status'. *Social Science Research* 4.
- Gagliani, G. (1985): 'Long-Term Changes in the Occupational Structure'. *European Sociological Review* 1.
- Ganzeboom, H., R. Luijkx and D.J. Treiman (1989): 'Intergenerational Class Mobility in Comparative Perspective'. *Research in Social Stratification and Mobility* 8.
- Glenn, N.D. (1977): *Cohort Analysis*. Beverly Hills: Sage.
- Goldthorpe, J.H. (with Catriona Llewellyn and Clive Payne) (1980, 2nd ed. 1987): *Social Mobility and Class Structure in Modern Britain*. Oxford: Clarendon Press.
- Goldthorpe, J.H. (1985): 'On Economic Development and Social Mobility'. *British Journal of Sociology* 36.
- Goldthorpe, J.H. and K. Hope (1974): *The Social Grading of Occupations: A New Approach and Scale*. Oxford: Clarendon Press.
- Goodman, L.A. (1972): 'A General Model for the Analysis of Surveys'. *American Journal of Sociology* 77.
- Grusky, D.B. and R.M. Hauser (1984): 'Comparative Social Mobility Revisited: Models of Convergence and Divergence in 16 Countries'. *American Sociological Review* 49.
- Hannan, D.F. (1979): *Displacement and Development: Class, Kinship and Social Change in Irish Rural Communities*. Dublin: The Economic and Social Research Institute.
- Hauser, R.M., P.J. Dickinson, H.P. Travis and J.M. Koffel (1975): 'Temporal Change in Occupational Mobility: Evidence for Men in the United States'. *American Sociological Review* 40.
- Hoem, J.M. and P. Linneman (1987): 'The Tails in Moving Average Graduation'. Stockholm: Research Reports in Demography 37, University of Stockholm.
- Hout, M. (1989): *Following in Father's Footsteps: Social Mobility in Ireland*. Cambridge, Mass.: Harvard University Press.

- Kaelble, H. (1981): *Historical Research on Social Mobility*. London: Croom Helm.
- Kaelble, H. (1984): 'Eras of Social Mobility in 19th and 20th Century Europe'. *Journal of Social History* 17.
- Kemp, T. (1978): *Historical Patterns of Industrialization*. London: Longman.
- Kerr, C. (1969): *Marshall, Marx and Modern Times*. Cambridge: Cambridge University Press.
- Kerr, C. (1983): *The Future of Industrial Societies*. Cambridge, Mass.: Harvard University Press.
- Kerr, C., J.T. Dunlop, F.H. Harbison and C.A. Myers (1960, 2nd ed. 1973): *Industrialism and Industrial Man*. Cambridge, Mass.: Harvard University Press.
- Kindleberger, C.P. (1967): *Europe's Postwar Growth: The Role of Labor Supply*. Cambridge, Mass.: Harvard University Press.
- Kocka, J. (1980): 'The Study of Social Mobility and the Formation of the Working Class in the 19th Century'. *Le mouvement social* 111.
- Kulcsár, K. (1984): *Contemporary Hungarian Society*. Budapest: Corvina.
- Kuznets, S. (1966): *Modern Economic Growth*. New Haven: Yale University Press.
- Landes, D.S. (1957): 'Observations on France: Economy, Society and Politics'. *World Politics*, April.
- Landes, D.S. (1965, 2nd ed. 1972): *The Unbound Prometheus: Technological Change and Industrial Development in Western Europe from 1750 to the Present*. Cambridge: Cambridge University Press.
- Lewis, P. (1973): 'The Peasantry' in D. Lane and G. Kolankiewicz, eds., *Social Groups in Polish Society*. London: Macmillan.
- Liebersohn, S. (1985, 2nd ed. 1987): *Making It Count*. Berkeley: University of California Press.
- de Lillo, A. and A. Schizzerotto (1985): *La valutazione sociale delle occupazioni*. Bologna: II Mulino.
- Lipset, S.M. and H.L. Zetterberg (1956): 'A Theory of Social Mobility'. *Transactions of the Third World Congress of Sociology*, vol. 3. London: International Sociological Association.
- Lipset, S.M. and H.L. Zetterberg (1959): 'Social Mobility in Industrial Societies' in S.M. Lipset and R. Bendix, *Social Mobility in Industrial Society*. Berkeley: University of California Press.
- Lutz, B. (1984): *Der kurze Traum immerwährender Prosperität*. Frankfurt: Campus.
- Mayer, A.J. (1981): *The Persistence of the Old Regime: Europe to the Great War*. New York: Pantheon.
- Mendels, F.F. (1976): 'Social Mobility and Phases of Industrialisation'. *Journal of Interdisciplinary History* 7.
- Naoi, A. (1979): 'Shokugyoteki Chiishakudo no Kosei' (The Construction of the Occupational Status Scale) in K. Tominaga, ed., *Nihon no Kaiso Kozo* (The Stratification Structure in Japan). Tokyo: Todai Shuppan Kai.
- O'Brien, P. and C. Keyder (1978): *Economic Growth in Britain and France, 1780–1914*. London: Allen and Unwin.
- O'Muircheartaigh, C.A. and R.D. Wiggins (1977): 'Sample Design and Evaluation for an Occupational Mobility Study'. *Economic and Social Review* 8.
- Parsons, T. (1960): *Structure and Process in Modern Societies*. Glencoe: Free Press.
- Parsons, T. (1967): *Sociological Theory and Modern Society*. New York: Free Press.
- Parsons, T. (1971): *The System of Modern Societies*. Englewood Cliffs: Prentice-Hall.
- Payne, G. (1987): *Employment and Opportunity*. London: Macmillan.
- Pohl, R., C. Thélot and M-F. Jousset (1974): *L'Enquête Formation-Qualification Professionnelle de 1970*. Paris: INSEE.
- Priebe, H. (1976): 'The Changing Role of Agriculture, 1920–1970' in C.M. Cipolla, ed., *Fonana Economic History of Europe*, vol. 5 (ii). London: Fontana.
- Rostow, W.W. (1960): *The Stages of Economic Growth: A Non-Communist Manifesto*. Cambridge: Cambridge University Press.
- Schwartz, J.E. (1985): 'Goodman's Coefficient of Multiple Determination: Why It Is Not Analogous to R^2 '. Stockholm: Swedish Institute for Social Research.
- Simkus, A.A., R. Andorka, J. Jackson, K-B. Yip and D.J. Treiman (1990): 'Changes in Social Mobility in Two Societies in the Crux of Transition: A Hungarian-Irish Comparison, 1943–1973'. *Research in Social Stratification and Mobility* 9.
- Singelmann, J. (1978): *From Agriculture to Services: The Transformation of Industrial Employment*. Beverly Hills: Sage.
- Singelmann, J. and M. Tienda (1985): 'The Process of Occupational Change in a Service Society: The Case of the United States, 1960–1980' in B. Roberts, R. Finnegan and D. Gallie, eds., *New Approaches to Economic Life*. Manchester: Manchester University Press.
- Sorokin, P.A. (1927, 2nd ed. 1959): *Social and Cultural Mobility*. Glencoe: Free Press.
- Treiman, D.J. (1970): 'Industrialisation and Social Stratification' in E.O. Laumann, ed., *Social Stratification: Research and Theory for the 1970s*. Indianapolis: Bobbs Merrill.
- Treiman, D.J. (1977): *Occupational Prestige in Comparative Perspective*. New York: Academic Press.

- Treiman, D.J. and K-B. Yip (1989): 'Educational and Occupational Attainment in 21 Countries' in M.L. Kohn, ed., *Cross-National Research in Sociology*. Newbury Park: Sage.
- Wegener, B. (1988): *Kritik des Prestiges*. Opladen: Westdeutscher Verlag.
- Wright, E.O. and J. Singelmann (1982): 'Proletarianization in the Changing American Class Structure'. *American Journal of Sociology* 88.
- Wright, E.O. and B. Martin (1987): 'The Transformation of the American Class Structure, 1960–1980'. *American Journal of Sociology* 93.
- Wrigley, E.A. (1972): 'The Process of Modernization and the Industrial Revolution in England'. *Journal of Interdisciplinary History* 3.
- Zagórski, K. (1977–8): 'Transformations of Social Structure and Social Mobility in Poland'. *International Journal of Sociology* 7.

48 ■ Richard Breen

Social Mobility in Europe

I discuss here the findings of a project that looks at social mobility in 11 European countries over a period of almost 30 years (see Breen 2004 for details). I want to summarize some of the main results and discuss how they relate to previous research in this area; and I also want to make some comments about the consequences of this research and of these findings for the study of social mobility. But I will begin by briefly explaining what the project is all about.

Data

The data used in the project comprise 117 mobility surveys covering the period 1970 to 2000. The sources of the data are shown in Table 1. The number of tables per country ranges from two in Israel and Italy to 35 in the Netherlands.

In general the age range of the respondents in our mobility tables is 25 to 64, and we coded social class according to the seven-class CASMIN scheme (see Table 2). This identifies classes I + II (the service class); III (routine non-manual class); IVab (petty bourgeoisie with and without employees); IVc

(farmers); V + VI (skilled manual workers, technicians and supervisors of manual workers); VIIa (unskilled manual workers not in agriculture); and VIIb (farm workers). One consequence of choosing this categorization is that it allows our results to be compared with those of *The Constant Flux* (Erikson and Goldthorpe 1993), where the same categories were used.

Methodological Issues

As Table 1 showed, the 11 countries contribute rather different numbers of mobility tables to our cross-national analyses. Sweden, for example, has a table for every year from 1976 to 1999, whereas Poland and Ireland have only three tables each, covering the years between the early 1970s and 1994. The amount of information we possess regarding change over time, and the reliability of the conclusions based on this information, will vary between countries. If we have a small number of observations, any one of them may be very influential in determining whether or not the data display a trend (as we shall see) and this will inevitably lead to

Table 1. Sources of Data

| <i>Country</i> | <i>No. of Tables</i> | <i>Sources of Data</i> | <i>Years for Which Data Are Included</i> |
|----------------|----------------------|--|---|
| Germany | 22 | Zumabus Allbus Politik in der BRD Wohlfahrtssurvey German socio-economic panel | 1976–77 1979(2) 1980 1982 1980 1982 1984 1986 1988 1990–92 1994 1996 1998 1978 1980 1978 1986 1999 |
| France | 4 | Formation-qualification professionnelle Insee surveys | 1970 1977 1985 1993 |
| Italy | 2 | National survey on social mobility Italian household longitudinal survey | 1985 1997 |
| Ireland | 3 | Survey of the determinants of occupational status and mobility Survey of income distribution and poverty Living in Ireland survey | 1973 1987 1994 |
| Great Britain | 15 | General household survey | 1973 1975–76 1979–1984 1987–1992 |
| Sweden | 24 | Annual surveys of living conditions (ULF) | 1976–1999 |
| Norway | 3 | Colbjørnsen et al. 1987 Moen et al. 1996 Level of Living Survey | 1982 1994 1995 |
| Poland | 3 | Zagorski 1976 Słomczynski 1989 Treiman/ Szelényi | 1972 1988 1994 |
| Hungary | 4 | Social mobility and life history survey Way of Life and Time Use Survey (Hungarian Central Statistical Office) | 1973 1983 1992 2000 |
| Israel | 2 | Matras and Weintraub, 1977 Kraus and Toren 1992 | 1974 1991 |
| Netherlands | 35 | Parliamentary Election Study Political Action Survey I Justice of Income Survey CBS Life Situation Survey National Labour Market Survey National Prestige and Mobility Survey Strategic Labour Market Survey Cultural Changes [ISSP] Justice of Income Survey Primary and Social Relationships Social and Cultural Trends Justice of Income Survey [ISJP] Family Survey I, 1992–93 Households in the Netherlands pilot Households in the Netherlands Social Inequality in the Netherlands National Crime Study Social and Economic Attitudes Netherlands Family Survey II Use of Information Technology | 1970 1971 1977 1981 1982 1986 1994 1998 1974 1979 1976 1977 1986 1982 1982 1985 1988 1990 1992 1994 1996 1998 1987 1987 1987 1990 1991 1992 1994 1995 1996 1996 1998 1998 1999 |

Table 2. The Goldthorpe Class Schema

| <i>Goldthorpe Class</i> | <i>Description</i> | <i>Employment Relations</i> | <i>CASMIN Version</i> |
|-------------------------|---|----------------------------------|---|
| I | Higher-grade professionals, administrators and officials; managers in large industrial establishments; large proprietors | Employer or service relationship | I + II Service class |
| II | Lower-grade professionals, administrators and officials; higher-grade technicians; managers in small industrial establishments; supervisors of non-manual employees | Service relationship | |
| IIIa | Routine non-manual employees, higher-grade (administration and commerce) | Intermediate | III Routine non-manual |
| IIIb | Routine non-manual employees, lower-grade (sales and services) | Labour contract | |
| IVa | Small proprietors, artisans, etc., with employees | Employer | IVab Non-farm petty-bourgeoisie |
| IVb | Small proprietors, artisans, etc., without employees | Self-employed | |
| IVc | Farmers and smallholders; other self-employed workers in primary production | Employer or self-employed | IVc Farmers and other self-employed workers in primary production |
| V | Lower-grade technicians; supervisors of manual workers | Intermediate | V + VI Technicians, supervisors and skilled manual workers |
| VI | Skilled manual workers | Labour contract | |
| VIIa | Semi- and unskilled manual workers (not in agriculture, etc.) | Labour contract | VIIa Semi- and unskilled manual workers (not in agriculture) |
| VIIb | Semi- and unskilled manual workers in agriculture | Labour contract | VIIb Semi- and unskilled manual workers in agriculture |

uncertainty in the conclusions we draw. All else equal we must, as a consequence, attach more credence to results about temporal trends drawn from countries with a larger number of observations (Sweden, the Netherlands, Great Britain and Germany).

Furthermore, the data that we use are never free of error, and differences in data quality may easily be mistaken for substantive differences. We have used the best-quality data

available from each of our 11 countries, but we still need to be aware of the potential for differential reliability and validity to induce spurious cross-national variation and temporal change. As far as the differences between countries are concerned, the fieldwork for the surveys we use was in all cases carried out according to internationally accepted procedures, and the subsequent coding of the variables—notably class origins and destinations—

Table 3a. Aggregate Class Structures (All Men) in the 11 Countries by Decade

| <i>Class</i> | <i>1970s</i> | <i>1980s</i> | <i>1990s</i> |
|--------------|--------------|--------------|--------------|
| I + II | 23.1 | 28.6 | 30.8 |
| III | 8.7 | 9.0 | 10.1 |
| IVa + b | 7.9 | 8.6 | 10.4 |
| IVc | 8.6 | 5.7 | 4.0 |
| V + VI | 27.7 | 27.6 | 27.1 |
| VIIa | 20.6 | 18.3 | 15.7 |
| VIIb | 3.5 | 2.3 | 2.0 |

Note: Each country is weighted equally in computation of aggregate class structure.

Table 3b. Aggregate Class Structures (Women in the Labour Force) in the 10 Countries by Decade

| <i>Class</i> | <i>1970s</i> | <i>1980s</i> | <i>1990s</i> |
|--------------|--------------|--------------|--------------|
| I + II | 22.1 | 30.5 | 34.6 |
| III | 32.8 | 32.3 | 35.1 |
| IVa + b | 6.5 | 6.0 | 6.1 |
| IVc | 8.6 | 4.4 | 2.3 |
| V + VI | 6.1 | 6.3 | 7.1 |
| VIIa | 21.1 | 18.6 | 13.7 |
| VIIb | 2.8 | 1.7 | 1.2 |

Note: Each country is weighted equally in computation of aggregate class structure.

followed a common and widely implemented procedure. Nevertheless, while adherence to such norms is some reassurance that the data attain high standards of quality, the surveys in the various countries were carried out independently of each other, and so we should be cautious about what we infer from them concerning cross-national differences. As far as change within countries is concerned, we can have more faith in our findings when the various surveys have been administered in a consistent fashion. In three cases the data always come from the same survey series: these are France (the FQP—Formation-Qualification Professionnelle—surveys), Britain (the General Household Survey) and Sweden (the ULF series). In a further five countries the data sets come from highly comparable sources: these are Ireland (where the three surveys were all carried out by the same field-work organization), Hungary (where the four surveys were all fielded by the Hungarian Central Statistical Office), Italy (where a

number of the same academics were involved in the design and execution of the two surveys), Germany and Israel. But in the remaining three cases—Norway, Poland and the Netherlands—the data come from various sources within each country and thus the possibility that variations in data quality might be mistaken for temporal change is greatest here. We believe that more reliance can be placed on estimates of trends *within* countries than measures of differences *between* them: thus our discussion, later in this chapter, of which countries are more or less open in their mobility regime, should be interpreted with some caution. Finally, while the data that we have are probably adequate for presenting a picture of broad trends and differences, we would have less confidence in the extent to which they allow the specifics of the pattern of social fluidity to be compared across either time or countries. This consideration has then dictated our choice of models. Rather than seeking to develop detailed models of

the fluidity regime we prefer instead to fit rather general models and to assess their adequacy using several measures (including the conventional chi-squared goodness of fit test and the index of dissimilarity).

Absolute Mobility and Class Structure

In contemporary studies of social mobility a key distinction is drawn between observed patterns of social mobility, sometimes referred to as ‘absolute mobility’, and social fluidity (or ‘relative mobility’). Absolute mobility is concerned with patterns and rates of mobility, where mobility is understood simply as movement between class origins (the social class in which someone was brought up) and class destinations (the class they occupy at the time of the survey). Social fluidity concerns the relationship between class origins and current class position: specifically it is based on the comparison, between people of different class origins, of their chances of being found in one destination class rather than another. If these chances were the same regardless of origins then the mobility table would display perfect mobility: the class in which a respondent was found would not depend on (would be independent of) the class in which she or he was brought up. Social fluidity is often interpreted as an index of equality in the chances of access to more or less advantageous social positions between people coming from different social origins (in other words, as an index of societal openness),¹ and contemporary research on social mobility accordingly pays much more attention to social fluidity than to absolute mobility. I will follow that precedent here and so will make only a few comments about absolute mobility.

In fact, our results concerning absolute mobility can be summarised quite easily. There has been a marked convergence in the class structures of European countries and in their patterns of absolute mobility,

and these things are true for both men and women.

- A. Convergence in class structures has been driven by some internationally consistent trends, such as the growth in the service class, I + II, and the decline in manual work, particularly of the unskilled kind. Among women, increased rates of labour force participation have been associated with a reduction in international variation as more and more of them enter occupations in the white collar classes, I + II and III. But the single biggest cause of this convergence has been the declining significance of the farm classes, IVc and VIIb, in those countries (such as Poland, Ireland, and Hungary) where a large farm sector persisted until the last quarter of the 20th century.
- B. This trend towards convergence in class structures has occurred together with decreasing variation between countries in their rates of overall mobility, of vertical, of upward and of downward mobility—and, again, this is evident among both sexes, as Tables 4 and 5 show. But, further, the distribution of people in the mobility tables of the different countries has also grown more similar. If we calculate the index of dissimilarity (Δ) from comparisons, between all pairs of countries, of their entire mobility tables, we find that the average Δ (the average difference between countries) falls from 43 per cent in the 1970s, to 33 in the 1980s and 30 in the 1990s, among women, with the comparable figures for men being 39, 30 and 30 per cent. And the variance around these means has also declined: from 163.2 to 62.6 to 41.6 among women and from 137.5 to 62.9 to 56.1 among men. Although European countries continue to show differences in their absolute mobility flows, these have become smaller.

Table 4. Percentage Mobile by Country in Each Decade (All Men)

| Total Mobility | Germany | France | Italy | Ireland | Great Britain | Sweden | Norway | Poland | Hungary | Israel | N'lands | Mean | Variance |
|--------------------------|---------|--------|-------|---------|---------------|--------|--------|--------|---------|--------|---------|------|----------|
| 1970s | 61.6 | 66.6 | — | 56.7 | 63.0 | 70.8 | — | 59.4 | 77.5 | 74.4 | 66.3 | 66.3 | 48.0 |
| 1980s | 62.1 | 67.5 | 69.5 | 61.3 | 61.8 | 71.4 | 71.9 | 61.0 | 74.9 | — | 67.7 | 66.9 | 25.8 |
| 1990s | 60.3 | 67.0 | 72.1 | 66.1 | 60.8 | 71.0 | 68.1 | 67.4 | 71.6 | 74.3 | 65.7 | 67.7 | 19.9 |
| Vertical Mobility | | | | | | | | | | | | | |
| 1970s | 44.1 | 43.8 | — | 39.9 | 50.7 | 54.0 | — | 40.9 | 53.0 | 43.7 | 50.6 | 46.7 | 28.5 |
| 1980s | 45.8 | 45.9 | 40.8 | 42.6 | 50.8 | 54.7 | 55.2 | 42.9 | 55.8 | — | 54.1 | 48.9 | 34.6 |
| 1990s | 46.3 | 46.3 | 46.3 | 45.5 | 50.7 | 55.2 | 52.1 | 45.9 | 53.7 | 50.4 | 54.0 | 49.7 | 13.9 |
| Upward Mobility | | | | | | | | | | | | | |
| 1970s | 31.7 | 25.9 | — | 21.6 | 32.8 | 35.1 | — | 22.1 | 26.9 | 20.1 | 36.1 | 28.0 | 37.1 |
| 1980s | 33.6 | 29.1 | 29.0 | 27.9 | 33.1 | 35.3 | 39.3 | 24.8 | 34.7 | — | 38.9 | 32.6 | 22.9 |
| 1990s | 33.3 | 29.9 | 35.9 | 31.4 | 31.7 | 36.6 | 34.2 | 26.3 | 35.9 | 35.0 | 37.7 | 33.4 | 11.4 |
| Downward Mobility | | | | | | | | | | | | | |
| 1970s | 12.4 | 17.9 | — | 18.4 | 17.9 | 19.0 | — | 18.8 | 26.2 | 23.5 | 14.5 | 18.7 | 17.3 |
| 1980s | 12.2 | 16.8 | 11.8 | 14.7 | 17.7 | 19.4 | 15.9 | 18.0 | 21.1 | — | 15.2 | 16.3 | 8.7 |
| 1990s | 13.0 | 16.4 | 10.4 | 14.1 | 19.0 | 18.6 | 17.9 | 19.6 | 17.8 | 15.4 | 16.3 | 16.2 | 8.0 |

Table 5. Percentage Mobile by Country in Each Decade (Women in the Labour Force)

| Total Mobility | | Germany | France | Italy | Great Britain | Sweden | Norway | Poland | Hungary | Israel | N'lands | Mean | Variance |
|-------------------|------|---------|--------|-------|---------------|--------|--------|--------|---------|--------|---------|------|----------|
| 1970s | 74.0 | 71.4 | — | — | 78.8 | 73.1 | — | 50.8 | 81.0 | 76.5 | 74.0 | 72.5 | 86.4 |
| 1980s | 75.6 | 77.6 | 74.3 | 76.3 | 73.6 | 73.6 | 76.2 | 66.3 | 79.5 | — | 73.9 | 74.8 | 13.7 |
| 1990s | 72.6 | 77.2 | 75.0 | 73.9 | 73.2 | 77.4 | 76.2 | 76.2 | 76.5 | 82.2 | 72.3 | 75.7 | 8.8 |
| Vertical Mobility | | | | | | | | | | | | | |
| 1970s | 48.6 | 41.7 | — | 52.1 | 55.4 | — | 34.0 | 34.0 | 54.1 | 44.9 | 51.4 | 47.8 | 52.2 |
| 1980s | 48.8 | 45.7 | 51.0 | 52.6 | 56.4 | 54.1 | 48.5 | 48.5 | 58.2 | — | 51.4 | 51.9 | 15.7 |
| 1990s | 47.3 | 46.0 | 47.9 | 53.2 | 57.9 | 53.0 | 50.3 | 50.3 | 55.7 | 53.5 | 53.6 | 51.8 | 14.8 |
| Upward Mobility | | | | | | | | | | | | | |
| 1970s | 25.8 | 27.8 | — | 27.5 | 23.9 | — | 19.5 | 19.5 | 23.3 | 26.0 | 30.9 | 25.6 | 11.7 |
| 1980s | 29.6 | 32.9 | 38.5 | 29.0 | 27.5 | 34.4 | 31.7 | 31.7 | 38.8 | — | 33.6 | 32.9 | 15.7 |
| 1990s | 32.2 | 33.2 | 36.7 | 30.6 | 33.5 | 37.1 | 34.1 | 34.1 | 42.0 | 39.0 | 34.8 | 35.3 | 11.6 |
| Downward Mobility | | | | | | | | | | | | | |
| 1970s | 22.8 | 13.9 | — | 24.6 | 31.5 | — | 14.4 | 14.4 | 30.8 | 19.0 | 20.5 | 22.2 | 44.2 |
| 1980s | 19.2 | 12.8 | 12.5 | 23.7 | 28.9 | 19.8 | 16.8 | 16.8 | 19.4 | — | 17.8 | 19.0 | 26.0 |
| 1990s | 15.2 | 12.8 | 11.3 | 22.5 | 24.4 | 15.9 | 16.2 | 16.2 | 13.7 | 14.5 | 18.8 | 16.5 | 17.6 |

Absolute mobility concerns the observed rates and patterns of flows between origin and destination classes and, in mobility analysis, is treated as the consequence of social fluidity (the relative chances of people from each origin being found in each destination class) operating within fixed origin and destination distributions. A model in which origins and destinations are independent, given the observed distributions of these two in each country and at each point in time, correctly classifies over 80 per cent of cases, while a model which also assumes a common level and pattern of social fluidity correctly classifies around 95 per cent of cases. It is evident, therefore, that changes over time, and differences between countries, in absolute mobility are driven by variation in the origin and destination distributions rather than in social fluidity.

Can such variation be said to follow a pattern? We believe that the answer, in very broad terms, is yes. We might imagine societies following a developmental path that incorporates two major transitions: from an agricultural to an industrial society, and from an industrial to a post-industrial society. The consequences, for the class structure, of the former transition are a decline in the proportions in classes IVc and VIIb and a growth in the remaining classes, especially (among men) the manual working classes V + VI and VIIa. The transition to a post-industrial society sees the decline of V + VI and VIIa and the growth of I + II and III. Everywhere the decline in agriculture is either more or less complete (Britain, Germany, Sweden, Israel, the Netherlands) or well underway while, in eight of our 11 countries (Ireland, Poland and Hungary being the exceptions), between the 1970s and 1990s, the class structure saw a steady fall in the proportion of men in classes V + VI and VIIa and a consistent increase in the proportion in I + II and III. Among women the pattern was exactly the same. These differences mean that some countries display a post-industrial class structure with a heavy concentra-

tion of people in classes I + II and III: this is particularly true of the male class structure in Britain and the Netherlands and it is true of the female class structure in several countries. But the important thing, from the point of view of the study of absolute mobility, is the recent rapidity of the transition out of agriculture. Similarly, we saw in our comparative analysis that the shift towards a concentration of women in the white-collar classes has been more rapid in countries such as Hungary and Poland where the class distribution in the 1970s differed most from this. The result has been the growing similarity in destination distributions that we have already remarked upon. But because countries embarked on this developmental path long before the first of our surveys was fielded, there is also decreasing variation in class origins. The mean value of the Δ between class origins for each pair of countries fell from 33 per cent in the 1970s to 23 in the 1980s and 24 in the 1990s.² Absolute mobility flows converged because their main determinants did.

This convergence chiefly occurred between the 1970s and 1980s and whether the trend will persist, or even strengthen, is, of course, difficult to say. Clearly, if the working classes continue to decline in those countries where the decline has begun, and if this extends from VIIa to V + VI, then further convergence will be inevitable as men, like women, come to be heavily concentrated in classes I + II and III. Recent historical experience of the location of industrial production would suggest that we can expect further convergence; in any event, it seems unlikely that any of these countries will display a growth in classes V + VI and VIIa, while some at least will experience a decline. As for the countries in which these classes have not yet begun to decline (Ireland, Poland and Hungary), the outlook seems less certain. In Ireland the growth of classes I + II and III has outstripped that of V + VI and VIIa over this period, but this is not true of the male class structure in Poland and Hungary. On the other hand, among women in

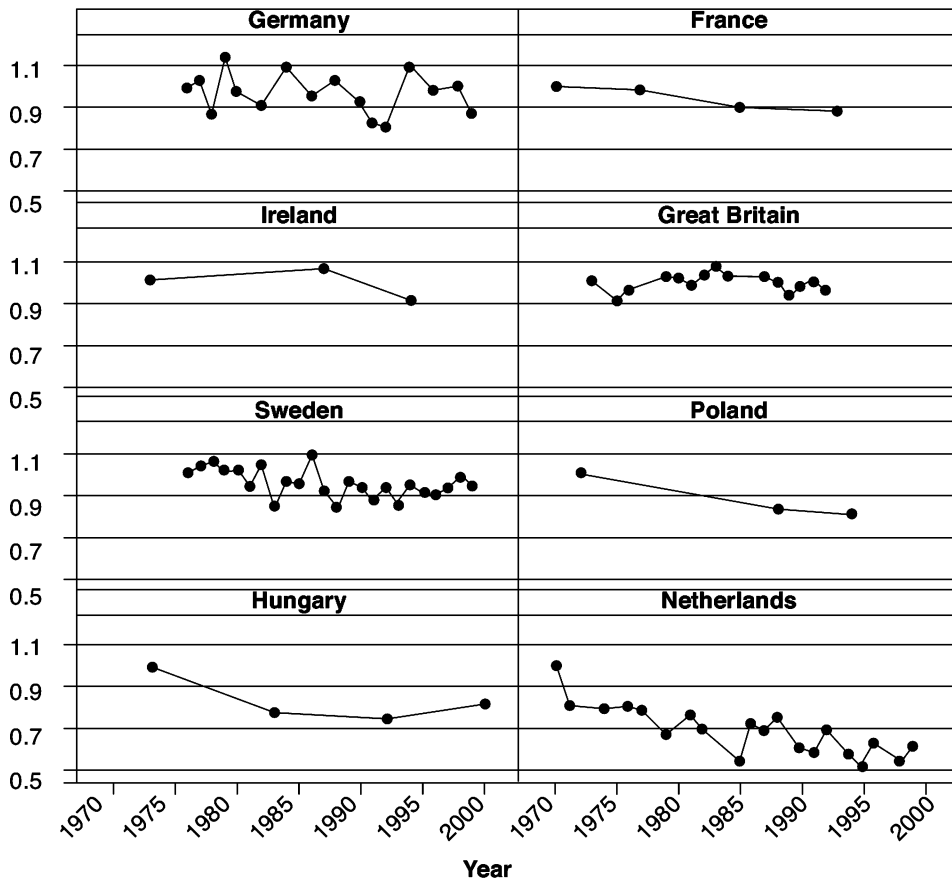


Figure 1. Annual LmSF (or Unidiff) Coefficients by Country (All Men)

Poland and Hungary there has been a steady growth in classes I + II and III and an increase, then a decline, in V + VI and VIIa, suggesting that the second transition may be under way. Much here depends on the nature of economic development. Foreign direct investment in manufacturing, as in the Irish case, is one mechanism by which the size of the working class may be sustained and the rate of convergence consequently slowed.

Social Fluidity

In our comparative analysis we found that trends in social fluidity are very similar among men and women, showing a widespread tendency towards greater fluidity. Britain is the sole clear exception to this: here there has been

little or no change. In other cases—notably Germany—there is no statistically significant change, though the trend, at least for men, is towards a weaker association between origins and destinations. Elsewhere—in France, Ireland, Sweden, Poland, Hungary and the Netherlands—there is a statistically significant increase in fluidity, though the small number of observations for Ireland (three), Poland (three), and Hungary (four), and the lack of a consistent pattern of change in these countries, must leave some room for uncertainty. But in contrast to absolute mobility we see no evidence of convergence among countries in their social fluidity. Figures 1 (for men) and 2 (for women) show these within-country trends in the form of the annual β coefficients from an LmSF model³ with common

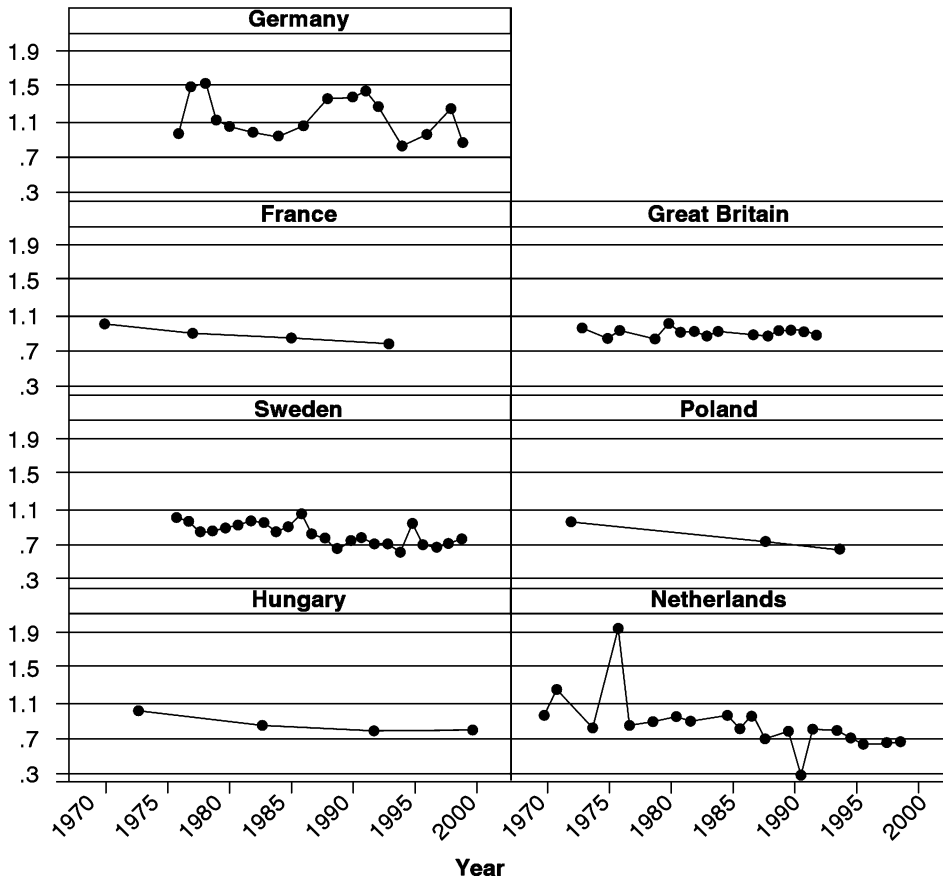


Figure 2. Annual LmSF (or Unidiff) Coefficients by Country (Women in the Workforce)

local association among all the yearly tables of a given country.

Figure 3 (for men) and Figure 4 (for women) show the LmSF β parameters from a model, applied to decade data from each country, which assumes common local association in each table, varying only by β . The value for Britain in the 1980s is set to unity. Among men, Figure 3 shows that, in the 1970s, levels of social fluidity were lowest in Germany, France, Italy, Ireland, Hungary and the Netherlands and highest in Britain, Sweden, Norway, Poland and Israel. Fluidity increased in France, Sweden and the Netherlands and possibly in Ireland, Hungary and Poland too. The increases in the Netherlands and Hungary were particularly marked. These different trends have left several countries—Sweden, Norway, Poland, Hungary

and the Netherlands—with, as far as we can tell, rather similar rates of fluidity, followed by Britain (where the absence of change has led to a shift in its relative position), Ireland, France, Italy and Germany, which remains the country with the strongest association between class origins and class destinations. At the other extreme, Israel is consistently more open than any other country. Overall, however, we can find no convincing evidence of convergence in fluidity regimes: for example, the within decade variance of the β s shown in Figure 3 is largest for the 1980s and the Δ for CnSF across countries in each decade is larger for the 1990s than for the 1980s.

The picture among women (4) is very similar. Once again, the points for the 1970s are above those for the 1980s, which are above those for the 1990s, indicating a general ten-

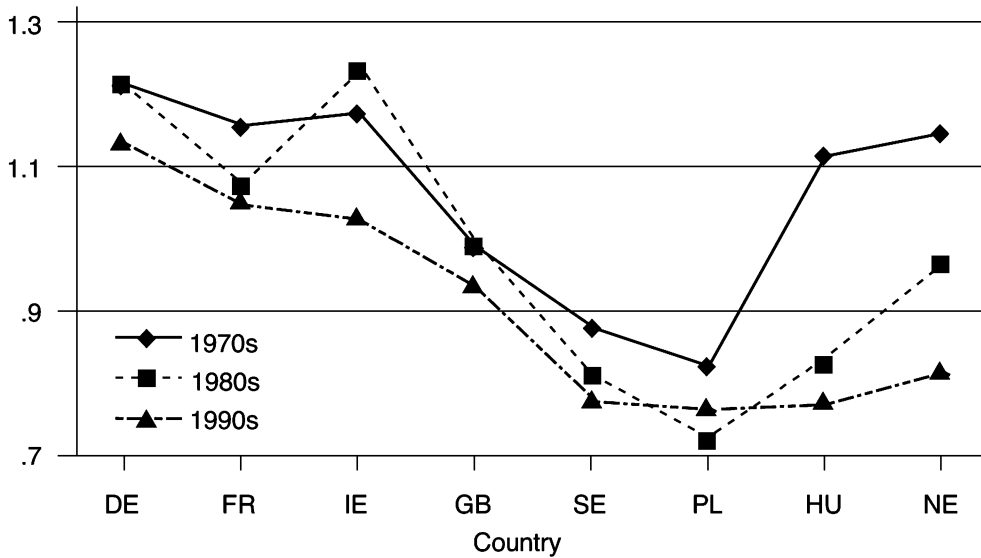


Figure 3. LmSF (or Unidiff) Coefficients per Decade, per Country (All Men)

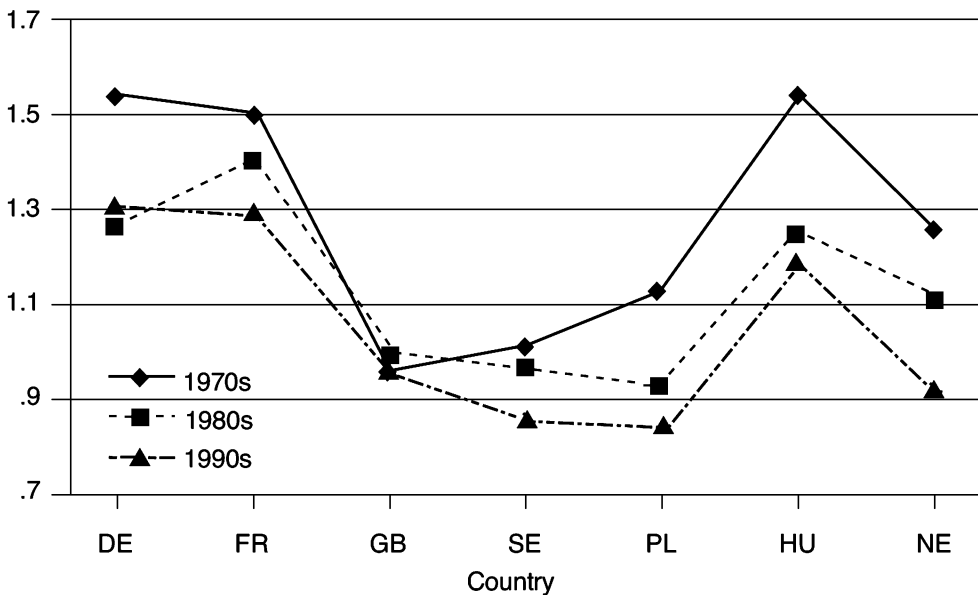


Figure 4. LmSF (or Unidiff) Coefficients per Decade, per Country (Women in the Labor Force)

dency for fluidity to increase, with Britain being an exception. The average β falls from 1.28 in the 1970s to 1.14 in the 1980s and 1.05 in the 1990s. France and Germany are the least fluid societies; Britain, Sweden, Poland and, by the 1990s, the Netherlands are the most fluid. Hungary presents a differ-

ent picture for women than men, the former showing much lower fluidity, compared with other countries, than the latter. In Israel the values are 0.84 in the 1970s and 0.71 in the 1990s. Taken together with the results for men this is evidence of the exceptionally fluid nature of Israeli society.

Overall, the results from our 11 countries point to a clear conclusion: there is a widespread tendency for social fluidity to increase, even though this might not be a statistically significant trend in every case. Among men, fluidity is greater at the end of the period than at the start in every country except Britain and Israel (where the values remain the same). Furthermore, of the 20 decade-to-decade changes in fluidity reported in Breen (2004), we find that in 16 of them fluidity increased and it declined in three—in Ireland and Britain between the 1970s and 1980s and in Norway between the 1980s and 1990s. There is just one further notable case in which fluidity fell (but which is obscured by the use of aggregated decade data) and that is Hungary, where fluidity declined significantly between the 1992 and 2000 observations. Although there are some cases (such as Sweden) where we cannot be unequivocal about an increase in fluidity, we can say with confidence that nowhere (with the possible exception of post-Communist Hungary) is there any evidence of a trend in the opposite direction. For women the picture is very similar. Of 18 decade-to-decade changes, two of them show a decline in fluidity (Germany between the 1980s and 1990s, Britain between the 1970s and 1980s) while 14 show an increase.

Previous Research

These results differ quite substantially from those of the previous major comparative study of social mobility, namely *The Constant Flux*, in which Erikson and Goldthorpe (1993) argue strongly in favour of the so-called FJH hypothesis of a basic similarity in social fluidity in all industrial societies ‘with a market economy and a nuclear family system’ (Featherman, Jones and Hauser 1975: 340) and they also claim (1992: 367) that ‘relative rates possess a high degree of temporal stability’. Indeed, our results are somewhat closer to those of the slightly earlier comparative study by Ganzeboom, Luijkx and Treiman (1989).

They used 149 mobility tables for men drawn from 35 countries spanning the period 1947–86. They claim that they show that ‘although . . . there is a basic similarity in mobility patterns . . . at the same time there are substantial cross-national and cross-temporal differences in the *extent* of mobility.’ Furthermore, ‘a smaller but significant part of the variance in mobility regimes can be explained by the trend towards increasing openness over time’ (Ganzeboom, Luijkx and Treiman 1989: 47). However, although our results and those of Ganzeboom et al. are rather similar, it should be noted that the latter have been heavily criticised, notably by Wong (1994: 138), who reanalyses their data and finds that ‘the model of temporal invariance cannot be rejected for a majority of countries . . . Hungary and Sweden are the only countries giving irrefutable evidence of temporal variation’.

Why do we find evidence for change and cross-national difference when Erikson and Goldthorpe did not? As far as temporal change is concerned, because Erikson and Goldthorpe have only one mobility table per country, their claim that there is little systematic temporal variation in patterns of social fluidity within countries is based on analyses in which age groups are taken to represent different birth cohorts. But as Breen and Jonsson (2003) have pointed out, this approach confounds lifecycle and cohort effects and makes no allowance for either selective mortality or recall errors. In an assessment of the reliability of measures of class origin and class destination in mobility tables, Breen and Jonsson (1997) found that reliability was lower for origin information from older age groups, implying that the common practice of using age groups to draw conclusions about cohort change over time in mobility regimes may be unsound. More simply, one cannot properly make inferences about change from cross-sectional data.

An important reason why we observe a trend towards increasing social fluidity and Erikson and Goldthorpe do not is that our data, by and large, refer to later-born cohorts.

Although our analyses, like those of the overwhelming majority of mobility studies, are based on period data, there is, I believe, good reason to suppose that social fluidity is driven by cohort, rather than period, effects.

Erikson and Goldthorpe's data represent cohorts born between, approximately, 1900 and 1945, whereas our data extend this to cohorts born around 1970. This means that our data contain a larger share of people who benefited from what has been called the Golden Age of Welfare Capitalism—in other words, the long post-war economic boom and the generally more egalitarian educational and social welfare policies that followed in its wake. Work that I have done with Janne Jonsson on Sweden clearly shows that here the major reduction in class inequalities in educational attainment took place among cohorts of people born between 1910 and 1950 and that since then there has been no further equalization. But this means that, as we compare period data for the 1970s, 1980s and 1990s, we gradually lose the older, less fluid cohorts, who are replaced by younger, more fluid cohorts. In the case of Sweden such a process will continue until the second decade of the 21st century, after which we expect to see no further increase in period social fluidity. But it seems likely that the specific cohorts which benefited from the Golden Age will have differed between countries, so leading to variation between them in the fluidity they display in particular periods.

Whereas we argue that countries display considerable variation in social fluidity, Erikson and Goldthorpe argue the contrary. But in their own analyses (Erikson and Goldthorpe 1992: 381) the logarithm of each odds ratio in their mobility table from Czechoslovakia (which is the country in which they find the greatest fluidity) is two-thirds as strong as that from Scotland (the country with the least fluidity). So, for example, given an odds ratio of 3 in Czechoslovakia, the comparable odds ratio would be 5.3 in Scotland. On this basis it is difficult to argue for commonality in the

strength of social fluidity, and, indeed, these results are quite similar to ours: in our data the logarithm of the odds ratios in Israel is around half those in Germany. So, an odds ratio of 3 in Israel would be an odds ratio of 7 in Germany. Given this, one might ask why Erikson and Goldthorpe, like many others, conclude that there is little variation in fluidity. The reason is the following. All comparative mobility studies find that the difference in the goodness of fit between a model in which social fluidity is common across all countries (the model of common social fluidity or CnSF) and a model in which fluidity can differ between them is small, relative to the difference between the CnSF model and one in which origins and destinations are independent. It is certainly true that the model of variation in fluidity is a significantly better fit to the data than CnSF, but this incremental improvement to any goodness of fit measure is small. For example, when we analyzed our data according to decade, a very small index of dissimilarity was returned by a model that allowed for no temporal or cross-national variation in social fluidity (3.95 per cent for men and 3.81 for women) and allowing for such variation only improved Δ by, at most, two percentage points. This compares with a Δ of around 15 per cent in models in which origins and destinations are independent. Much the same picture emerged when we used annual data, and arguments like this usually lead to the conclusion that most social fluidity is common and invariant over time. Sometimes the same point is made using the deviance, rather than Δ , as the yardstick, and here the result is even more extreme. For example, 90 per cent of the deviance returned by the model of independence disappears when we add common (among countries and over decades) social fluidity, and a model allowing fluidity to change over time and differ between countries improves it only by a further seven per cent. Taken together, the decompositions of the deviance and of Δ would seem to indicate that more than 85 per cent of social fluidity is common over nations and time.

This then seems to conflict with our finding of large variation between countries in their odds ratios. The problem with using arguments based on Δ or the deviance to support the argument that social fluidity is largely invariant is that variation in fluidity is assessed using measures of the consequences of fluidity for the whole mobility regime (i.e. for absolute mobility), and, as we have seen, these consequences are unanimously agreed to be quite minor. An analogy may help to make the point. In a linear regression, $Y = a + bX$; X (analogous to social fluidity) may display a lot of variation, but it will have little impact on Y (analogous to overall mobility) if the coefficient, b , is close to zero. But simply because b is small we could not then claim that there was little variation in X . Measures such as the change in Δ or in G^2 might be said to capture the strength of effect of fluidity on overall mobility, but they do not measure the variation in fluidity itself, and it is therefore mistaken to conclude, on this basis, that social fluidity is common and invariant. It is empirically the case that, between countries or over time, large variations in social fluidity can be found which nevertheless have little impact on the overall mobility regime. To illustrate this: if we take the fluidity pattern from the 1997 Italian men's table in our data and insert it into the 1991 Israeli men's table, while preserving the Israeli marginal distributions, the Δ between the real and the constructed Israeli tables is 6 per cent.⁴ When we consider that the Israeli and Italian mobility regimes are close to the extremes of the range of fluidity found in our data, this suggests that 6 per cent represents the maximum impact of differences in fluidity on the distribution of individuals in the mobility table. The conclusion to be drawn from these apparently contradictory measures of the variation in fluidity is not that it is common or invariant, but, rather, that even quite substantial differences in fluidity have little impact on the distribution of cases over the mobility table—i.e. on observed, absolute mobility flows.

Explanation

Thus far I have said nothing about how we might explain patterns of social mobility or social fluidity, aside from some rather vague remarks about the impact of policies pursued during the Golden Age of Welfare Capitalism. In broad terms, explanations of variations in social fluidity are usually pitched at a macro-sociological level: that is, fluidity is assumed to be related to other, societal level measures. The very influential 'liberal theory of industrialism', for example, argues that fluidity will increase with economic development. In our case, the stage of economic development of our countries varies rather little, but, even so, we could find no evident link between their ranking in fluidity terms and their GDP per capita. Nor did we find any support for Erikson and Goldthorpe's argument that fluidity is related to income inequality: there is no significant association between a country's fluidity and its Gini coefficient. Overall we could discern no tendencies towards either convergence or divergence in fluidity, and thus the hypothesis that, as nations have come to follow different policy trajectories—particularly in economic policy—so we might see growing differences between them in fluidity, receives no support. There is some indication, however, that fluidity is greater in state socialist (Poland and Hungary) and social democratic (Norway and Sweden) countries, and the argument for such a political explanation receives additional support from the finding of declining fluidity in Hungary during the 1990s. But, on the other hand, we observe very high fluidity in Israel and data from the General Social Survey (made available to us by Mike Hout) show that fluidity is high in the United States. This leads to the conclusion that direct political intervention of the kinds associated with state socialist and social democratic societies may be one means by which a society can reach relatively high rates of fluidity, but it is not the only one.

One major difficulty in devising theories to explain variation in mobility or fluidity is that mobility tables, especially period-based tables, reflect a large number of underlying processes—artefactual, contingent and substantive. For one thing, this aggregation of processes renders it difficult to explain variations in fluidity; for another, it may also be the case that some of the commonality that has often been observed in comparisons of social fluidity derives from the mixing together in the mobility tables of processes that, when investigated separately, might show greater and more systematic societal and temporal differences.

While it is reasonable to suppose that fluidity in a nation is shaped by government policy, the education system, the workings of the labour market and the like, it is also the case that what we observe in a mobility table may also reflect purely artefactual sources of variation arising from differences in the way that the data themselves represent the underlying phenomenon of interest. Furthermore, what we might call contingent factors, which are usually omitted from any theoretical discussion of social fluidity, may play an important role in shaping what we observe. Walter Müller and Reinhard Pollak (see Breen 2004) attribute the high fluidity they find among Germans born in the 1920s to the massive migration from the eastern part of Germany that occurred following the Second World War. The measured class origins of this cohort are thus their pre-migration origins, which had very little relevance in shaping their subsequent mobility patterns: the physical detaching of a large share of the cohort from their true origins led to higher measured social fluidity. The same argument may explain the high level of fluidity in Israel, a country in which a very large share of the population is comprised of immigrants.

Conclusions

The experience of this project should lead us to question the balance that mobility re-

search has struck between social fluidity and absolute mobility. The emphasis has lain heavily on the former but, insofar as we are concerned with the mobility regime, this now seems inappropriate. This is by no means to deny that social fluidity tells us important things about the prevailing degree of inequality in the chances of attaining one class position rather than another, and may be indicative of other characteristics of society. Nevertheless, although one would not want to say that fluidity can never make a difference (since we can easily construct examples in which extreme patterns of fluidity will be highly consequential for the distribution of cases in a mobility table), within the advanced industrial and post-industrial societies, the range of fluidity that we observe is relatively inconsequential in determining variation in mobility flows and in the life chances of individuals and families as these are captured in measures of class position. Many previous authors (such as Grusky and Hauser 1984; Goldthorpe 1985) have called for more attention to be paid to structural change, but, as Erikson and Goldthorpe (1992: 104, 189) suggest, it is not clear how such change should be explained nor, indeed, whether it might not be better approached as a matter of historical description rather than sociological explanation. But while this might be a valid concern if we conceive of class structures as macro-sociological phenomena, it may be less so, and may leave open the possibility of sociological explanation, if we were to turn our attention to the detailed evolution of businesses and firms and of the jobs that constitute classes.

Furthermore, it now seems to me that a period perspective on social mobility may not be the most appropriate if, as I have suggested, social fluidity is driven by factors related to cohorts, rather than periods. The cohort perspective on social mobility has not been entirely neglected, but it has certainly played a very secondary role to the usual period view. Under the cohort approach, change in the

association between origins and destinations is believed to occur in specific cohorts, while, from a period point of view, this association is considered to change among all cohorts at similar historical points in time. The importance, for social change, of the replacement of older by younger cohorts has been stressed by several sociologists. In the specific field of mobility research, Erikson and Goldthorpe (1992), for example, argue that equality of condition is essential for equality of opportunity, and that changes in the latter will then predominantly take effect during a person's upbringing and schooling. As a result, the overall level of inequality of opportunity in society will change mainly through cohort replacement, with younger cohorts experiencing a different degree of inequality of educational attainment than older ones. It has long been education that is the major channel of social mobility: if this is so, then social fluidity should respond to changes in the level of class inequality in educational attainment, which, because educational inequality is itself a characteristic of cohorts, implies a cohort explanation of change in fluidity.

Though rarely spelled out, there is a tension between the cohort interpretation of change in social fluidity and the belief that period effects drive variation in the origin-destination association over the lifecycle through labour market changes that simultaneously affect different birth cohorts. We believe that there are strong arguments for adopting a cohort rather than, or in addition to, a period perspective. It is widely agreed that changes in an individual's social class position are relatively rare after the age of about 35 (Goldthorpe 1980: 51–2, 69–71; Erikson and Goldthorpe 1992: 72). Such stability implies that, except in unusual circumstances in which this stability is disrupted, period effects will be a less important source of change in fluidity when compared with cohort replacement and within-cohort change that occurs during the early years in the labour force.

NOTES

1. The terms association and social fluidity can be used interchangeably: greater social fluidity implies lower origin–destination association. In the log-linear modeling context in which this paper is situated, this association is captured by odds ratios.

2. These figures are for men. For women the figures are 36, 24 and 24 per cent. The slight differences arise because our samples of women include only those in the labour force and we have no data for women in Ireland.

3. LmSF means 'log-multiplicative social fluidity' (by analogy with CnSF—'common social fluidity'). We use LmSF to refer to the unidiff (Erikson and Goldthorpe 1992) or log-multiplicative layer effect model (Xie 1992) when the local association is modeled in a completely unspecified way (as it is in CnSF).

4. We use the observed Italian fluidity pattern, and thus the magnitude of the difference that we report does not depend on the adequacy of any particular model of fluidity.

REFERENCES

- Breen, Richard (ed.). 2004. *Social Mobility in Europe*. Oxford: Oxford University Press.
- Erikson, Robert and John H. Goldthorpe. 1992. *The Constant Flux: A Study of Class Mobility in Industrial Societies*. Oxford: Oxford University Press.
- Featherman, David L., F. L. Jones, and Robert M. Hauser. 1975. "Assumptions of Mobility Research in the United States: The Case of Occupational Status." *Social Science Research* 4: 329–60.
- GANZEBOOM, Harry B. G., Ruud Luijkx, and Donald J. Treiman. 1989. "Intergenerational Class Mobility in Comparative Perspective." *Research in Social Stratification and Mobility* 8: 3–84.
- Goldthorpe, John H. 1980. *Social Mobility and Class Structure in Modern Britain*. Oxford: Clarendon Press.
- Goldthorpe, John H. 1985. "On Economic Development and Social Mobility." *British Journal of Sociology* 36: 549–573.
- Grusky, David B. and Robert M. Hauser. 1984. "Comparative Social Mobility Revisited: Models of Convergence and Divergence in 16 Countries." *American Sociological Review* 49: 19–38.
- Wong, Raymond S. 1994. "Postwar Mobility Trends in Advanced Industrial Societies." *Research in Social Stratification and Mobility* 13: 121–44.

■ INCOME MOBILITY ■

49 ■ Gary Solon

Intergenerational Income Mobility

Intergenerational income mobility is the degree to which position in the income distribution persists or changes from one generation to the next.

For example, a society in which individuals' adult income is altogether independent of their parents' income is a highly mobile society. A society in which one's percentile in the income distribution is surely identical to one's parents' percentile is completely immobile. Neither extreme is ideal, and neither corresponds to the intergenerational patterns typically observed in actual societies. Which intergenerational association between the extremes is desirable depends on the processes generating it. In any case, reasonable and well-informed observers are likely to disagree about the optimal level of intergenerational mobility because of differences in their values, such as feelings about equity/efficiency tradeoffs. See Jencks and Tach (forthcoming) for a discussion of some of the normative issues.

Because understanding intergenerational mobility is important for understanding the nature of income inequality, intergenerational mobility has received a great deal of attention from both theoretical and empirical researchers. Over the last two decades, empirical research describing the extent of intergenerational mobility has made considerable progress. Much work remains to improve our understanding of the causal processes under-

lying the observed extent of intergenerational mobility.

Theory

The classic theoretical analysis by Becker and Tomes (1979) encompasses a multitude of reasons that relative income status is correlated across generations. In a recent variant of that analysis, Solon (2004) adopts the functional form assumptions consistent with the log-linear intergenerational mobility regression commonly estimated by empirical researchers. In this model, an individual parent divides her income between her own consumption and investment in an individual child's human capital so as to maximize a Cobb-Douglas utility function in which the two goods are the parent's consumption and the child's adult income. The mapping from the parent's investment in her child's human capital to the child's subsequent income as an adult operates through two functions.

First, a semi-logarithmic human capital production function relates the child's level of human capital to the logarithm of the sum of the parent's investment and public investment (e.g., publicly supported education and health care for children) plus a variable representing the human capital endowments children receive regardless of the investment choices of their families and the government.

These more mechanically determined endowments to children are, in Becker and Tomes's words, "determined by the reputation and 'connections' of their families, the contribution to the ability, race, and other characteristics of children from the genetic constitutions of their families, and the learning, skills, goals, and other 'family commodities' acquired through belonging to a particular family culture." The transmission of these endowments is assumed to follow a first-order autoregressive process across generations. Thus, intergenerational transmission occurs both because higher-income parents have greater wherewithal to invest in the human capital of their children and because of the genetic and cultural heritability of human capital. Second, the mapping to the child's income is completed by a semi-logarithmic earnings function that relates the child's log earnings to her level of human capital.

This simple model leaves out some important aspects of intergenerational transmission. For example, it assumes that the parent cannot borrow against the child's prospective earnings and does not bequeath financial assets to the child. See Becker and Tomes (1986) and Mulligan (1997) for analyses that relax this assumption. Also, the model's single-parent/single-child structure ignores the role of assortative mating, which is discussed by Lam and Schoeni (1994) and Chadwick and Solon (2002). Nevertheless, the model is rich enough to illustrate some key aspects of the intergenerational transmission process. These are embodied in the following result concerning the intergenerational income elasticity β , which is the coefficient in the regression of the child's log income on the parent's log income:

$$\beta \cong \frac{(1-\gamma)\theta + \lambda}{1 + (1-\gamma)\theta\lambda} \quad (1)$$

where θ is the elasticity of earnings with respect to human capital investment, λ is the

autoregressive parameter representing the genetic and cultural heritability of income-generating traits, and λ is an index of the progressivity of public investment in children's human capital.

This result implies that the intergenerational income elasticity increases with the responsiveness of earnings to human capital investment and with the heritability of income-generating traits, and it decreases with the progressivity of public investment in children's human capital. Cross-country differences in intergenerational mobility could arise from differences in any of these factors. So could changes over time in a particular country's intergenerational mobility. Finally, it can be shown that the same factors that increase the intergenerational income elasticity also increase the equilibrium level of cross-sectional income inequality within a generation. Thus, we should not be surprised if societies with particularly high income inequality also exhibit high intergenerational persistence of income status.

Empirical Evidence

If lifetime income data were available for both the parents' and children's generations in a nationally representative sample, estimation of the intergenerational income elasticity β could be performed simply by applying least squares to the regression of the children's log lifetime income on the parents' log lifetime income. In most countries, however, the ideal data are not available. As of the 1980s, data constraints had forced most of the then-small empirical literature to rely on short-term income measures, such as annual income in only one year, for peculiarly homogeneous samples. As summarized in Becker and Tomes (1986), the resulting estimates suggested that "a 10% increase in father's earnings (or income) raises son's earnings by less than 2%." As discussed in detail in Solon (1989, 1992), however, these esti-

mates were biased substantially downward. The “right-sided” measurement error from using short-term parental income measures to proxy for parents’ lifetime income can serve as a good classroom example for the econometrics textbook analysis of the attenuation bias resulting from “noisy” measurement of an explanatory variable. And when the estimates were based on relatively homogeneous parent samples, this bias was aggravated by the diminished “signal variance” in the explanatory variable.

By the 1990s, empirical researchers in the United States had the benefit of better data. By that time, two longitudinal surveys initiated in the late 1960s, the Panel Study of Income Dynamics (PSID) and the National Longitudinal Surveys (NLS) of labor market experience, had generated new data with an intergenerational span. Because these surveys used national probability samples, they were less subject to the problems from homogeneous samples. And because the longitudinal surveys repeatedly collected income information at each reinterview, they enabled exploration of the impact of using longer-term income measures. Many of the new studies, surveyed in Solon (1999), treated the errors-in-variables issue by averaging the parental income measured over several years. A typical finding was that, in a regression of son’s log earnings on a multi-year measure of father’s log earnings, the estimated slope coefficient was about 0.4—that is, double the 0.2 value that previously had been described as an upper bound for the intergenerational elasticity. A few studies treated the errors-in-variables problem by performing instrumental variables estimation with parental characteristics like education or occupation used to instrument for measured parental income. That approach usually produced somewhat higher intergenerational elasticity estimates, but, as explained in Solon (1992), the consistency of such instrumental variables estimation depends on the “excludability” of

the instruments from the model for children’s income.

Even the new estimates based on multi-year parental income data probably were too low. As emphasized in Solon (1992) and Mazumder (2005), averaging parental income over several years reduces but does not eliminate attenuation bias.

Nonrandom attrition from the longitudinal surveys probably generated a weaker version of the sample homogeneity that had plagued earlier data sets. And, as discussed in Reville (1995) and Haider and Solon (2005), many of the newer estimates have been biased by “left-sided” measurement error. At the time researchers began to use intergenerational data from the PSID and NLS, the offspring were only about thirty years old or even younger. For workers in their twenties, the log of current income as a proxy for log lifetime income is subject to “mean-reverting” measurement error, instead of the classical measurement error typically analyzed in econometrics textbooks. The mean-reversion occurs because the workers who eventually will have high lifetime earnings typically experience steeper earnings growth. As a result, the early-career gap in current earnings between workers with high and low lifetime earnings tends to understate their lifetime gap. This sort of mean-reverting measurement error in a dependent variable is still another source of attenuation bias. Once all these downward biases in the estimation of the intergenerational elasticity are considered, it becomes plausible that the intergenerational elasticity in the United States may well be as large as 0.5 or 0.6.

In recent years, researchers have estimated intergenerational elasticities for many other countries, sometimes with much larger samples than are available from the U.S. surveys. As summarized in Solon (2002), the elasticity estimates for the United States and United Kingdom are towards the high end

among developed countries, with considerably smaller estimates appearing for Canada, Sweden, Finland, and Norway. Some new estimates for developing countries in Latin America (Dunn, 2004; Ferreira and Veloso, 2004; Grawe, 2004) are even higher than the U.S. and U.K. estimates. By and large, these cross-country comparisons accord with the theoretical prediction of greater intergenerational income persistence in countries with greater income inequality, higher returns to human capital, and less progressive public investment in children's human capital. A related question is whether the changes in income inequality experienced by many countries over the last quarter-century have been accompanied by changing intergenerational elasticities. In most of the time-trends research conducted so far, the time spans and sample sizes have been too limited to permit strong conclusions.

Cross-country comparisons have only begun to illuminate *why* intergenerational income associations are as large (and as small) as they are. To what extent does intergenerational transmission occur because higher-income parents invest more in their children's human capital? What are the roles of genetic and cultural heritability? One intriguing line of research seeks clues from comparisons of relatives with varying degrees of genetic and environmental relatedness. Siblings studies of this type (Taubman, 1976; Bjorklund, Jantti, and Solon, 2005) have compared correlations in socioeconomic status among monozygotic twins, dizygotic twins, non-twin full siblings, half-siblings, and biologically unrelated adoptive siblings and also have compared biological siblings reared together and apart.¹ Intergenerational studies (Bjorklund, Lindahl, and Plug, 2004; Plug, 2004; Sacerdote, 2004) have compared parent-child outcome associations in biological and adoptive families. Some empirical patterns consistent with an important role for ge-

netic transmission are: outcome correlations are particularly high among monozygotic twins; correlations for dizygotic twins and non-twin full siblings exceed those for half-siblings and adoptive siblings; correlations for biological siblings are positive even when the siblings are reared apart; intergenerational associations are higher for biologically related parents and children; and adoptive children's outcomes are positively associated with those of their biological parents (even after controlling for the adoptive parents' outcomes). Empirical patterns consistent with important environmental factors are: outcome correlations are positive among biologically unrelated adoptive siblings; correlations among biological siblings tend to be higher when the siblings are reared together; and adoptive children's outcomes are positively associated with those of their adoptive parents (even after controlling for the biological parents' outcomes).

NOTES

1. A related literature (Solon, Page, and Duncan, 2000; Page and Solon, 2003a, 2003b; Oreopoulos, 2003; Raaum, Salvanes, and Sorensen, forthcoming) has compared sibling correlations and correlations among unrelated children who grew up in the same neighborhood. The typical finding that the sibling correlations are considerably larger than the neighbor correlations suggests that family influences loom larger than neighborhood influences in accounting for the effects of origins on socioeconomic outcomes.

REFERENCES

- Becker, G.S. and Tomes, N. 1979. An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of Political Economy* 87, 1153–1189.
- Becker, G.S. and Tomes, N. 1986. Human capital and the rise and fall of families. *Journal of Labor Economics* 4, S1-S39.
- Bjorklund, A., Jantti, M., and Solon, G. 2005. Influences of nature and nurture on earnings variation: a report on a study of various sibling types in Sweden. In S. Bowles, H. Gintis, and M. Osborne

- Groves, *Unequal Chances: Family Background and Economic Success*, 145–164. New York: Russell Sage Foundation.
- Bjorklund, A., Lindahl, M., and Plug, E. 2004. Intergenerational effects in Sweden: what can we learn from adoption data? Discussion Paper No. 1194, IZA.
- Chadwick, L. and Solon, G. 2002. Intergenerational income mobility among daughters. *American Economic Review* 92, 335–344.
- Dunn, C.E. 2004. The intergenerational transmission of earnings: evidence from Brazil. Doctoral dissertation, University of Michigan.
- Ferreira, S.G. and Veloso, F.A. 2004. Intergenerational mobility of wages in Brazil. Unpublished manuscript.
- Grawe, N.D. 2004. Intergenerational mobility for whom? The experience of high- and low-earning sons in international perspective. In M. Corak (ed.), *Generational Income Mobility in North America and Europe*, 58–89. Cambridge: Cambridge University Press.
- Haider, S. and Solon, G. 2005. Life-cycle variation in the association between current and lifetime earnings. Unpublished manuscript.
- Jencks, C. and Tach, L. Forthcoming. Would equal opportunity mean more mobility? In S.L. Morgan, D.B. Grusky, and G.S. Fields (eds.), *Mobility and Inequality: Frontiers of Research from Sociology and Economics*. Stanford: Stanford University Press.
- Lam, D. and Schoeni, R.F. 1994. Family ties and labor markets in the United States and Brazil. *Journal of Human Resources* 29, 1235–1258.
- Mazumder, B. 2005. Fortunate sons: new estimates of intergenerational mobility in the U.S. using Social Security earnings data. *Review of Economics and Statistics* 87, 235–255.
- Mulligan, C.B. 1997. *Parental Priorities and Economic Inequality*. Chicago: University of Chicago Press.
- Oreopoulos, P. 2003. The long-run consequences of living in a poor neighborhood. *Quarterly Journal of Economics* 118, 1533–1575.
- Page, M.E. and Solon, G. 2003a. Correlations between brothers and neighboring boys in their adult earnings: the importance of being urban. *Journal of Labor Economics* 21, 831–855.
- Page, M.E. and Solon, G. 2003b. Correlations between sisters and neighbouring girls in their subsequent income as adults. *Journal of Applied Econometrics* 18, 545–562.
- Plug, E. 2004. Estimating the effect of mother's schooling on children's schooling using a sample of adoptees. *American Economic Review* 94, 358–368.
- Raam, O., Salvanes, K., and Sorensen, E.O. Forthcoming. The neighbourhood is not what it used to be. *Economic Journal*.
- Reville, R.T. 1995. Intertemporal and life cycle variation in measured intergenerational earnings mobility. Unpublished manuscript.
- Sacerdote, B. 2004. What happens when we randomly assign children to families? Working Paper No. 10894, National Bureau of Economic Research.
- Solon, G. 1989. Biases in the estimation of intergenerational earnings correlations. *Review of Economics and Statistics* 71, 172–174.
- _____. 1992. Intergenerational income mobility in the United States. *American Economic Review* 82, 393–408.
- _____. 1999. Intergenerational mobility in the labor market. In O. C. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, vol. 3A, 1761–2000. Amsterdam: North-Holland.
- _____. 2002. Cross-country differences in intergenerational earnings mobility. *Journal of Economic Perspectives* 16, 59–66.
- _____. 2004. A model of intergenerational mobility over time and place. In M. Corak (ed.), *Generational Income Mobility in North America and Europe*, 38–47. Cambridge: Cambridge University Press.
- Solon, G., Page, M.E., and Duncan, G.J. 2000. Correlations between neighboring children in their subsequent educational attainment. *Review of Economics and Statistics* 82, 383–392.
- Taubman, P. 1976. The determinants of earnings: genetics, family, and other environments; a study of white male twins. *American Economic Review* 66, 858–870.

Status and Income Attainment

■ BASIC MODELS ■

50 ■ Peter M. Blau and Otis Dudley Duncan,
with the collaboration of Andrea Tyree

The Process of Stratification

Stratification systems may be characterized in various ways. Surely one of the most important has to do with the processes by which individuals become located, or locate themselves, in positions in the hierarchy comprising the system. At one extreme we can imagine that the circumstances of a person's birth—including the person's sex and the perfectly predictable sequence of age levels through which he is destined to pass—suffice to assign him unequivocally to a ranked status in a hierarchical system. At the opposite extreme his prospective adult status would be wholly problematic and contingent at the time of birth. Such status would become entirely determinate only as adulthood was reached, and solely as a consequence of his own actions taken freely—that is, in the absence of any constraint deriving from the circumstances of his birth or rearing. Such a pure achievement system is, of course, hypothetical, in much the same way that motion without friction is a purely hypothetical possibility in the physical world. Whenever the stratification system of any moderately large and complex society is described, it is seen to

involve both ascriptive and achievement principles.

In a liberal democratic society we think of the more basic principle as being that of achievement. Some ascriptive features of the system may be regarded as vestiges of an earlier epoch, to be extirpated as rapidly as possible. Public policy may emphasize measures designed to enhance or to equalize opportunity—hopefully, to overcome ascriptive obstacles to the full exercise of the achievement principle.

The question of how far a society may realistically aspire to go in this direction is hotly debated, not only in the ideological arena but in the academic forum as well. Our contribution, if any, to the debate will consist largely in submitting measurements and estimates of the strength of ascriptive forces and of the scope of opportunities in a large contemporary society. The problem of the relative importance of the two principles in a given system is ultimately a quantitative one. We have pushed our ingenuity to its limit in seeking to contrive relevant quantifications.

The governing conceptual scheme in the analysis is quite a commonplace one. We think of the individual's life cycle as a sequence in time that can be described, however partially and crudely, by a set of classificatory or quantitative measurements taken at successive stages. Ideally we should like to have under observation a cohort of births, following the individuals who make up the cohort as they pass through life. As a practical matter we resorted to retrospective questions put to a representative sample of several adjacent cohorts so as to ascertain those facts about their life histories that we assumed were both relevant to our problem and accessible by this means of observation.

Given this scheme, the questions we are continually raising in one form or another are: how and to what degree do the circumstances of birth condition subsequent status? and, how does status attained (whether by ascription or achievement) at one stage of the life cycle affect the prospects for a subsequent stage? The questions are neither idle nor idiosyncratic ones. Current policy discussion and action come to a focus in a vaguely explicated notion of the "inheritance of poverty." Thus a spokesman for the Social Security Administration writes:

It would be one thing if poverty hit at random and no one group were singled out. It is another thing to realize that some seem destined to poverty almost from birth—by their color or by the economic status or occupation of their parents.¹

Another officially sanctioned concept is that of the "dropout," the person who fails to graduate from high school. Here the emphasis is not so much on circumstances operative at birth but on the presumed effect of early achievement on subsequent opportunities. Thus the "dropout" is seen as facing "a lifetime of uncertain employment,"² probable assignment to jobs of inferior status, reduced

earning power, and vulnerability to various forms of social pathology.

In this study we do not have measurements on all the factors implicit in a full-blown conception of the "cycle of poverty" nor all those variables conceivably responding unfavorably to the achievement of "dropout" status. . . . This limitation, however, is not merely an analytical convenience. We think of the selected quantitative variables as being sufficient to describe the major outlines of status changes in the life cycle of a cohort. Thus a study of the relationships among these variables leads to a formulation of a basic model of the process of stratification.

A Basic Model

To begin with, we examine only five variables. For expository convenience, when it is necessary to resort to symbols, we shall designate them by arbitrary letters but try to remind the reader from time to time of what the letters stand for. These variables are:

- V*: Father's educational attainment
- X*: Father's occupational status
- U*: Respondent's educational attainment
- W*: Status of respondent's first job
- Y*: Status of respondent's occupation in 1962

Each of the three occupational statuses is scaled by the [socioeconomic] index described [elsewhere],³ ranging from 0 to 96. The two education variables are scored on the following arbitrary scale of values ("rungs" on the "educational ladder") corresponding to specified numbers of years of formal schooling completed:

- 0: No school
- 1: Elementary, one to four years
- 2: Elementary, five to seven years
- 3: Elementary, eight years
- 4: High school, one to three years

- 5: High school, four years
- 6: College, one to three years
- 7: College, four years
- 8: College, five years or more (i.e., one or more years of postgraduate study)

Actually, this scoring system hardly differs from a simple linear transformation, or “coding,” of the exact number of years of school completed. In retrospect, for reasons given [elsewhere],⁴ we feel that the score implies too great a distance between intervals at the lower end of the scale; but the resultant distortion is minor in view of the very small proportions scored 0 or 1.

A basic assumption in our interpretation of regression statistics—though not in their calculation as such—has to do with the causal or temporal ordering of these variables. In terms of the father’s career we should naturally assume precedence of V (education) with respect to X (occupation when his son was 16 years old). We are not concerned with the father’s career, however, but only with his statuses that comprised a configuration of background circumstances or origin conditions for the cohorts of sons who were respondents in the Occupational Changes in a Generation (OCG) study. Hence we generally make no assumption as to the priority of V with respect to X ; in effect, we assume the measurements on these variables to be contemporaneous from the son’s viewpoint. The respondent’s education, U , is supposed to follow in time—and thus to be susceptible to causal influence from—the two measures of father’s status. Because we ascertained X as of respondent’s age 16, it is true that some respondents may have completed school before the age to which X pertains. Such cases were doubtlessly a small minority and in only a minor proportion of them could the father (or other family head) have changed status radically in the two or three years before the respondent reached 16.

The next step in the sequence is more problematic. We assume that W (first job

status) follows U (education). The assumption conforms to the wording of the questionnaire, which stipulated “the first full-time job you had after you left school.” In the years since the OCG study was designed we have been made aware of a fact that should have been considered more carefully in the design. Many students leave school more or less definitively, only to return, perhaps to a different school, some years later, whereupon they often finish a degree program.⁵ The OCG questionnaire contained information relevant to this problem, namely the item on age at first job. Through an oversight no tabulations of this item were made for the present study. Tables prepared for another study⁶ using the OCG data, however, suggest that approximately one-eighth of the respondents report a combination of age at first job and education that would be very improbable unless (a) they violated instructions by reporting a part-time or school-vacation job as the first job, or (b) they did, in fact, interrupt their schooling to enter regular employment. (These “inconsistent” responses include men giving 19 as their age at first job and college graduation or more as their education; 17 or 18 with some college or more; 14, 15, or 16 with high-school graduation or more; and under 14 with some high school or more.) When the two variables are studied in combination with occupation of first job, a very clear effect is evident. Men with a given amount of education beginning their first jobs early held lower occupational statuses than those beginning at a normal or advanced age for the specified amount of education.

Despite the strong probability that the U - W sequence is reversed for an appreciable minority of respondents, we have hardly any alternative to the assumption made here. If the bulk of the men who interrupted schooling to take their first jobs were among those ultimately securing relatively advanced education, then our variable W is downwardly biased, no doubt, as a measure of their

occupational status immediately after they finally left school for good. In this sense, the correlations between U and W and between W and Y are probably attenuated. Thus, if we had really measured "job after completing education" instead of "first job," the former would in all likelihood have loomed somewhat larger as a variable intervening between education and 1962 occupational status. We do not wish to argue that our respondents erred in their reports on first job. We are inclined to conclude that their reports were realistic enough, and that it was our assumption about the meaning of the responses that proved to be fallible.

The fundamental difficulty here is conceptual. If we insist on *any* uniform sequence of the events involved in accomplishing the transition to independent adult status, we do violence to reality. Completion of schooling, departure from the parental home, entry into the labor market, and contracting of a first marriage are crucial steps in this transition, which all normally occur within a few short years. Yet they occur at no fixed ages nor in any fixed order. As soon as we aggregate individual data for analytical purposes we are forced into the use of simplifying assumptions. Our assumption here is, in effect, that "first job" has a uniform significance for all men in terms of its temporal relationship to educational preparation and subsequent work experience. If this assumption is not strictly correct, we doubt that it could be improved by substituting any other *single* measure of initial occupational status. (In designing the OCG questionnaire, the alternative of "job at the time of first marriage" was entertained briefly but dropped for the reason, among others, that unmarried men would be excluded thereby.)

One other problem with the U - W transition should be mentioned. Among the younger men in the study, 20 to 24 years old, are many who have yet to finish their schooling or to take up their first jobs or both—not to mention the men in this age group missed by the survey on account of their military service.⁷

Unfortunately, an early decision on tabulation plans resulted in the inclusion of the 20 to 24 group with the older men in aggregate tables for men 20 to 64 years old. We have ascertained that this results in only minor distortions by comparing a variety of data for men 20 to 64 and for those 25 to 64 years of age. Once over the U - W hurdle, we see no serious objection to our assumption that both U and W precede Y , except in regard to some fraction of the very young men just mentioned.

In summary, then, we take the somewhat idealized assumption of temporal order to represent an order of priority in a causal or processual sequence, which may be stated diagrammatically as follows:

$$(V, X) - (U) - (W) - (Y).$$

In proposing this sequence we do not overlook the possibility of what Carlsson calls "delayed effects,"⁸ meaning that an early variable may affect a later one not only via intervening variables but also directly (or perhaps through variables not measured in the study).

In translating this conceptual framework into quantitative estimates the first task is to establish the pattern of associations between the variables in the sequence. This is accomplished with the correlation coefficient. Table 1 supplies the correlation matrix on which much of the subsequent analysis is based. In discussing causal interpretations of these correlations, we shall have to be clear about the distinction between two points of view. On the one hand, the simple correlation—given our assumption as to direction of causation—measures the gross magnitude of the effect of the antecedent upon the consequent variable. Thus, if $r_{YW} = .541$, we can say that an increment of one standard deviation in first job status produces (whether directly or indirectly) an increment of just over half of one standard deviation in 1962 occupational status. From another point of view we are more concerned with net effects. If both first

Table 1. Simple Correlations for Five Status Variables

| Variable | Variable | | | | |
|-------------------------|----------|------|------|------|------|
| | Y | W | U | X | V |
| Y: 1962 occ. status | — | .541 | .596 | .406 | .322 |
| W: First-job status | | — | .538 | .417 | .332 |
| U: Education | | | — | .438 | .453 |
| X: Father's occ. status | | | | — | .516 |
| V: Father's education | | | | | — |

job and 1962 status have a common antecedent cause—say, father's occupation—we may want to state what part of the effect of *W* on *Y* consists in a transmission of the prior influence of *X*. Or, thinking of *X* as the initial cause, we may focus on the extent to which its influence on *Y* is transmitted by way of its prior influence on *W*.

We may, then, devote a few remarks to the pattern of gross effects before presenting the apparatus that yields estimates of net direct and indirect effects. Since we do not require a causal ordering of father's education with respect to his occupation, we may be content simply to note that $r_{XV} = .516$ is somewhat lower than the corresponding correlation, $r_{YU} = .596$, observed for the respondents themselves. The difference suggests a heightening of the effect of education on occupational status between the fathers' and the sons' generations. Before stressing this interpretation, however, we must remember that the measurements of *V* and *X* do not pertain to some actual cohort of men, here designated "fathers." Each "father" is represented in the data in proportion to the number of his sons who were 20 to 64 years old in March 1962.

The first recorded status of the son himself is education (*U*). We note that r_{UV} is just slightly greater than r_{UX} . Apparently both measures on the father represent factors that may influence the son's education.

In terms of gross effects there is a clear ordering of influences on first job. Thus $r_{WU} > r_{WX} > r_{WV}$. Education is most strongly correlated with first job, fol-

lowed by father's occupation, and then by father's education.

Occupational status in 1962 (*Y*) apparently is influenced more strongly by education than by first job; but our earlier discussion of the first-job measure suggests we should not overemphasize the difference between r_{YW} and r_{YU} . Each, however, is substantially greater than r_{YX} , which in turn is rather more impressive than r_{YV} .

Figure 1 is a graphic representation of the system of relationships among the five variables that we propose as our basic model. The numbers entered on the diagram, with the exception of r_{XV} , are path coefficients, the estimation of which will be explained shortly. First we must become familiar with the conventions followed in constructing this kind of diagram. The link between *V* and *X* is shown as a curved line with an arrowhead at both ends. This is to distinguish it from the other lines, which are taken to be paths of influence. In the case of *V* and *X* we may suspect an influence running

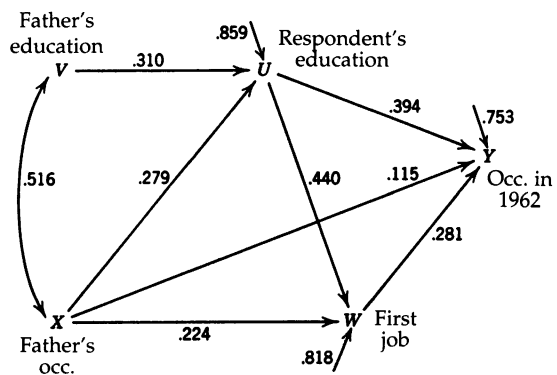


Figure 1. Path Coefficients in Basic Model of the Process of Stratification

from the former to the latter. But if the diagram is logical for the respondent's generation, we should have to assume that for the fathers, likewise, education and occupation are correlated not only because one affects the other but also because common causes lie behind both, which we have not measured. The bidirectional arrow merely serves to sum up all sources of correlation between V and X and to indicate that the explanation thereof is not part of the problem at hand.

The straight lines running from one measured variable to another represent *direct* (or net) influences. The symbol for the path coefficient, such as P_{YV} , carries a double subscript. The first subscript is the variable at the head of the path, or the effect; the second is the causal variable. (This resembles the convention for regression coefficients, where the first subscript refers to the "dependent" variable, the second to the "independent" variable.)

Finally, we see lines with no source indicated carrying arrows to each of the effect variables. These represent the residual paths, standing for all other influences on the variable in question, including causes not recognized or measured, errors of measurement, and departures of the true relationships from additivity and linearity, properties that are assumed throughout the analysis.

An important feature of this kind of causal scheme is that variables recognized as effects of certain antecedent factors may, in turn, serve as causes for subsequent variables. For example, U is caused by V and X , but it in turn influences W and Y . The algebraic representation of the scheme is a system of equations, rather than the single equation more often employed in multiple regression analysis. This feature permits a flexible conceptualization of the *modus operandi* of the causal network. Note that Y is shown here as being influenced directly by W , U , and X , but not by V (an assumption that will be justified shortly). But this does not imply that V has no influence on Y . V affects U , which does

affect Y both directly and indirectly (via W). Moreover, V is correlated with X , and thus shares in the gross effect of X on Y , which is partly direct and partly indirect. Hence the gross effect of V on Y , previously described in terms of the correlation r_{YV} , is here interpreted as being entirely indirect, in consequence of V 's effect on intervening variables and its correlation with another cause of Y .

Path Coefficients

Whether a path diagram, or the causal scheme it represents, is adequate depends on both theoretical and empirical considerations. At a minimum, before constructing the diagram we must know, or be willing to assume, a causal ordering of the observed variables (hence the lengthy discussion of this matter earlier in this chapter). This information is external or *a priori* with respect to the data, which merely describe associations or correlations. Moreover, the causal scheme must be complete, in the sense that all causes are accounted for. Here, as in most problems involving analysis of observational data, we achieve a formal completeness of the scheme by representing unmeasured causes as a residual factor, presumed to be uncorrelated with the remaining factors lying behind the variable in question. If any factor is known or presumed to operate in some other way it must be represented in the diagram in accordance with its causal role, even though it is not measured. Sometimes it is possible to deduce interesting implications from the inclusion of such a variable and to secure useful estimates of certain paths in the absence of measurements on it, but this is not always so. A partial exception to the rule that all causes must be explicitly represented in the diagram is the unmeasured variable that can be assumed to operate strictly as an intervening variable. Its inclusion would enrich our understanding of a causal system without invalidating the causal scheme that omits it. Sociologists have only recently begun to

appreciate how stringent are the logical requirements that must be met if discussion of causal processes is to go beyond mere impressionism and vague verbal formulations.⁹ We are a long way from being able to make causal inferences with confidence, and schemes of the kind presented here had best be regarded as crude first approximations to adequate causal models.

On the empirical side, a minimum test of the adequacy of a causal diagram is whether it satisfactorily accounts for the observed correlations among the measured variables. In making such a test we employ the fundamental theorem in path analysis, which shows how to obtain the correlation between any two variables in the system, given the path coefficients and correlations entered on the diagram.¹⁰ Without stating this theorem in general form, we may illustrate its application here. For example,

$$r_{YX} = p_{YX} + p_{YU}r_{UX} + p_{YW}r_{WX}$$

and

$$r_{WX} = p_{WX} + p_{WU}r_{UX}.$$

We make use of each path leading to a given variable (such as Y in the first example) and the correlations of each of its causes with all other variables in the system. The latter correlations, in turn, may be analyzed; for example, r_{WX} , which appeared as such in the first equation, is broken down into two parts in the second. A complete expansion along these lines is required to trace out all the indirect connections between variables; thus,

$$r_{YX} = p_{YX} + p_{YU}p_{UX} + p_{YU}p_{UV}r_{VX} + p_{YW}p_{WX} \\ + p_{YW}p_{WU}p_{UX} + p_{YW}p_{WU}p_{UV}r_{VX}.$$

Now, if the path coefficients are properly estimated, and if there is no inconsistency in

the diagram, the correlations calculated by a formula like the foregoing must equal the observed correlations. Let us compare the values computed from such a formula with the corresponding observed correlations:

$$r_{WV} = p_{WX}r_{XV} + p_{WU}r_{UV} \\ = (.224)(.516) + (.440)(.453) \\ = .116 + .199 = .315,$$

which compares with the observed value of .332; and

$$r_{YV} = p_{YU}r_{UV} + p_{YX}r_{XV} + p_{YW}r_{WV} \\ = (.394)(.453) + (.115)(.516) + \\ (.281)(.315) \\ = .326$$

(using here the calculated rather than the observed value of r_{WV}), which resembles the actual value, .322. Other such comparisons—for r_{YX} for example—reveal, at most, trivial discrepancies (no larger than .001).

We arrive, by this roundabout journey, at the problem of getting numerical values for the path coefficients in the first place. This involves using equations of the foregoing type inversely. We have illustrated how to obtain correlations if the path coefficients are known, but in the typical empirical problem, we know the correlations (or at least some of them) and have to estimate the paths. For a diagram of the type of Figure 1 the solution involves equations of the same form as those of linear multiple regression, except that we work with a recursive system of regression equations¹¹ rather than a single regression equation.

Table 2 records the results of the regression calculations. It can be seen that some alternative combinations of independent variables were studied. It turned out that the net re-

Table 2. Partial Regression Coefficients in Standard Form (Beta Coefficients) and Coefficients of Determination, for Specified Combinations of Variables

| Dependent Variables ^a | Independent Variables ^a | | | | Coefficient of Determination (R^2) |
|----------------------------------|------------------------------------|----------|----------|----------|--|
| | <i>W</i> | <i>U</i> | <i>X</i> | <i>V</i> | |
| <i>U</i> ^b | — | — | .279 | .310 | .26 |
| <i>W</i> | — | .433 | .214 | .026 | .33 |
| <i>W</i> ^b | — | .440 | .224 | — | .33 |
| <i>Y</i> | .282 | .397 | .120 | -.014 | .43 |
| <i>Y</i> ^b | .281 | .394 | .115 | — | .43 |
| <i>Y</i> | .311 | .428 | — | — | .42 |

^a *V*: father's education, *X*: father's occ. status, *U*: respondent's education, *W*: first-job status, *Y*: 1962 occ. status.

^b Beta coefficients in these sets were taken as estimates of path coefficients for Figure 1.

gressions of both *W* and *Y* on *V* were so small as to be negligible. Hence *V* could be disregarded as a direct influence on these variables without loss of information. The net regression of *Y* on *X* was likewise small but, as it appears, not entirely negligible. Curiously, this net regression is of the same order of magnitude as the proportion of occupational inheritance in this population—about 10 percent, as discussed [elsewhere].¹² We might speculate that the direct effect of the father's occupation on the occupational status of a mature man consists of this modest amount of strict occupational inheritance. The remainder of the effect of *X* on *Y* is indirect, inasmuch as *X* has previously influenced *U* and *W*, the son's education and the occupational level at which he got his start. For reasons noted [elsewhere]¹³ we do not assume that the full impact of the tendency to take up the father's occupation is registered in the choice of first job.

With the formal properties of the model in mind we may turn to some general problems confronting this kind of interpretation of our results. One of the first impressions gained from Figure 1 is that the largest path coefficients in the diagram are those for residual factors, that is, variables not measured. The residual path is merely a convenient representation of the extent to which measured causes in the system fail to ac-

count for the variation in the effect variables. (The residual is obtained from the coefficient of determination; if $R^2_{Y(WUX)}$ is the squared multiple correlation of *Y* on the three independent variables, then the residual for *Y* is $\sqrt{1 - R^2_{Y(WUX)}}$.) Sociologists are often disappointed in the size of the residual, assuming that this is a measure of their success in "explaining" the phenomenon under study. They seldom reflect on what it would mean to live in a society where nearly perfect explanation of the dependent variable could be secured by studying causal variables like father's occupation or respondent's education. In such a society it would indeed be true that some are "destined to poverty almost from birth . . . by the economic status or occupation of their parents" (in the words of the reference cited in endnote 1). Others, of course, would be "destined" to affluence or to modest circumstances. By no effort of their own could they materially alter the course of destiny, nor could any stroke of fortune, good or ill, lead to an outcome not already in the cards.

Thinking of the residual as an index of the adequacy of an explanation gives rise to a serious misconception. It is thought that a high multiple correlation is presumptive evidence that an explanation is correct or nearly so, whereas a low percentage of determination means that a causal interpretation is almost

certainly wrong. The fact is that the size of the residual (or, if one prefers, the proportion of variation “explained”) is *no* guide whatever to the validity of a causal interpretation. The best-known cases of “spurious correlation”—a correlation leading to an egregiously wrong interpretation—are those in which the coefficient of determination is quite high.

The relevant question about the residual is not really its size at all, but whether the unobserved factors it stands for are properly represented as being uncorrelated with the measured antecedent variables. We shall entertain [elsewhere]¹⁴ some conjectures about unmeasured variables that clearly are not uncorrelated with the causes depicted in Figure 1. It turns out that these require us to acknowledge certain possible modifications of the diagram, whereas other features of it remain more or less intact. A delicate question in this regard is that of the burden of proof. It is all too easy to make a formidable list of unmeasured variables that someone has alleged to be crucial to the process under study. But the mere existence of such variables is already acknowledged by the very presence of the residual. It would seem to be part of the task of the critic to *show*, if only hypothetically, but *specifically*, how the modification of the causal scheme to include a new variable would disrupt or alter the relationships in the original diagram. His argument to this effect could then be examined for plausibility and his evidence, if any, studied in terms of the empirical possibilities it suggests.

Our supposition is that the scheme in Figure 1 is most easily subject to modification by introducing additional measures of the same kind as those used here. If indexes relating to socioeconomic background other than V and X are inserted we will almost certainly estimate differently the direct effects of these particular variables. If occupational statuses of the respondent intervening between W and Y were known we should have to modify more or less radically the right-hand

portion of the diagram. Yet we should argue that such modifications may amount to an enrichment or extension of the basic model rather than an invalidation of it. The same may be said of other variables that function as intervening causes. In theory, it should be possible to specify these in some detail, and a major part of the research worker’s task is properly defined as an attempt at such specification. In the course of such work, to be sure, there is always the possibility of a discovery that would require a fundamental reformulation, making the present model obsolete. Discarding the model would be a cost gladly paid for the prize of such a discovery.

Postponing the confrontation with an altered model, the one at hand is not lacking in interest. An instructive exercise is to compare the magnitudes of gross and net relationships. Here we make use of the fact that the correlation coefficient and the path coefficient have the same dimensionality. The correlation $r_{YX} = .405$ (Table 1) means that a unit change (one standard deviation) in X produces a change of 0.4 unit in Y , in gross terms. The path coefficient, $P_{YX} = .115$ (Figure 1), tells us that about one-fourth of this gross effect is a result of the direct influence of X on Y . (We speculated above on the role of occupational inheritance in this connection.) The remainder ($.405 - .115 = .29$) is indirect, via U and W . The sum of all indirect effects, therefore, is given by the difference between the simple correlation and the path coefficient connecting two variables. We note that the indirect effects on Y are generally substantial, relative to the direct. Even the variable temporally closest (we assume) to Y has “indirect effects”—actually, common antecedent causes—nearly as large as the direct ones. Thus $r_{YW} = .541$ and $p_{YW} = .281$, so that the aggregate of “indirect effects,” which in this case are common determinants of Y and W that spuriously inflate the correlation between them, is .26.

To ascertain the indirect effects along a given chain of causation we must multiply

the path coefficients along the chain. The procedure is to locate on the diagram the dependent variable of interest, and then trace back along the paths linking it to its immediate and remote causes. In such a tracing we may reverse direction once but only once, following the rule “first back, then forward.” Any bidirectional correlation may be traced in either direction. If the diagram contains more than one such correlation, however, only one may be used in a given compound path. In tracing the indirect connections no variable may be intersected more than once in one compound path. Having traced all such possible compound paths, we obtain the entirety of indirect effects as their sum.

Let us consider the example of effects of education on first job, *U* on *W*. The gross or total effect is $r_{WU} = .538$. The direct path is $p_{WU} = .440$. There are two indirect connections or compound paths: from *W* back to *X*, then forward to *U*; and from *W* back to *X*, then back to *V*, and then forward to *U*. Hence we have:

$$r_{WU} = p_{WU} + \underbrace{p_{WX}p_{UX} + p_{WX}r_{XV}p_{UV}}_{(indirect)}$$

(gross) (direct)

or, numerically,

$$\begin{aligned} .538 &= .440 + (.224)(.279) + \\ &\quad (.224)(.516)(.310) \\ &= .440 + .062 + .036 \\ &= .440 + .098. \end{aligned}$$

In this case all the indirect effect of *U* on *W* derives from the fact that both *U* and *W* have *X* (plus *V*) as a common cause. In other instances, when more than one common cause is involved and these causes are themselves interrelated, the complexity is too great to permit a succinct verbal summary.

A final stipulation about the scheme had best be stated, though it is implicit in all the

previous discussion. The form of the model itself, but most particularly the numerical estimates accompanying it, are submitted as valid only for the population under study. No claim is made that an equally cogent account of the process of stratification in another society could be rendered in terms of this scheme. For other populations, or even for subpopulations within the United States, the magnitudes would almost certainly be different, although we have some basis for supposing them to have been fairly constant over the last few decades in this country. The technique of path analysis is not a method for discovering causal laws but a procedure for giving a quantitative interpretation to the manifestations of a known or assumed causal system as it operates in a particular population. When the same interpretive structure is appropriate for two or more populations there is something to be learned by comparing their respective path coefficients and correlation patterns. We have not yet reached the stage at which such comparative study of stratification systems is feasible. . . .

The Concept of a Vicious Circle

Although the concept of a “cycle of poverty” has a quasi-official sanction in U.S. public policy discussion, it is difficult to locate a systematic explication of the concept. As clear a formulation as any that may be found in academic writing is perhaps the following:¹⁵

Occupational and social status are to an important extent self-perpetuating. They are associated with many factors which make it difficult for individuals to modify their status. Position in the social structure is usually associated with a certain level of income, education, family structure, community reputation, and so forth. These become part of a vicious circle in which each factor acts on the other in such a way as to preserve the social structure in its present form, as well as the individual family’s position in that structure.

. . . The cumulation of disadvantages (or of advantages) affects the individual's entry into the labor market as well as his later opportunities for social mobility.

The suspicion arises that the authors in preparing this summary statement were partly captured by their own rhetoric. Only a few pages earlier they had observed that the "widespread variation of educational attainment within classes suggests that one's family background plays an enabling and motivating rather than a determining role."¹⁶ But is an "enabling and motivating role" logically adequate to the function of maintaining a "vicious circle"? In focusing closely on the precise wording of the earlier quotation we are not interested in splitting hairs or in generating a polemic. It merely serves as a convenient point of departure for raising the questions of what is specifically meant by "vicious circle," what are the operational criteria for this concept, and what are the limits of its usefulness.

To begin with, there is the question of fact—or, rather, of how the quantitative facts are to be evaluated. How "difficult" is it, in actuality, "for individuals to modify their status" (presumably reference is to the status of the family of orientation)? We have found that the father-son correlation for occupational status is of the order of .4. (Assuming attenuation by errors of measurement, this should perhaps be revised slightly upward.) Approaching the measurement problem in an entirely different way, we find that the amount of intergenerational mobility between census major occupation groups is no less than seven-eighths as much as would occur if there were no statistical association between the two statuses whatsoever, or five-sixths as much as the difference between the "minimum" mobility involved in the intergenerational shift in occupation distributions and the amount required for "perfect" mobility.¹⁷ Evidently a very considerable amount of "status modification" or occupational mo-

bility does occur. (There is nothing in the data exhibited by Lipset and Bendix to indicate the contrary.) If the existing amount of modification of status is insufficient in terms of some functional or normative criterion implicitly employed, the precise criterion should be made explicit: *How much mobility must occur to contradict the diagnosis of a "vicious circle"?*

Next, take the postulate that occupational status (of origin) is "associated with many factors" and that "each factor acts on the other" so as "to preserve . . . the individual family's position." Here the exposition virtually cries out for an explicit *quantitative* causal model; if not one of the type set forth in the first section of this chapter, then some other model that also takes into account the way in which several variables combine their effects. Taking our own earlier model, for want of a better alternative, as representative of the situation, what do we learn about the "associated factors"? Family "position" is, indeed, "associated with . . . education," and education in turn makes a sizable difference in early and subsequent occupational achievement. Yet of the total or gross effect of education (U) on Y , occupational status in 1962 ($r_{YU} = .596$), only a minor part consists in a transmission of the prior influence of "family position," at least as this is indicated by measured variables V (father's education) and X (father's occupation). . . . A relevant calculation concerns the compound paths through V and X linking Y to U . Using data for men 20 to 64 years old with non-farm background, we find:

$$p_{YX}p_{UX} = .025$$

$$p_{YX}r_{XV}p_{UV} = .014$$

$$p_{YX}p_{WX}p_{UX} = .014$$

$$p_{YW}p_{WX}r_{XV}p_{UV} = .008$$

$$\text{Sum} = .061$$

This is the *entire* part of the effect of education that has to do with “perpetuating” the “family’s position.” By contrast, the direct effect is $p_{YU} = .407$ and the effect via W (exclusive of prior influence of father’s education and occupation on respondent’s first job) is $p_{YW}p_{WU} = .128$, for a total of .535. Far from serving in the main as a factor perpetuating initial status, education operates *primarily* to induce variation in occupational status that is independent of initial status. The simple reason is that the large residual factor for U is an indirect cause of Y . But by definition it is quite uncorrelated with X and V . This is not to gainsay the equally cogent point that the degree of “perpetuation” (as measured by r_{YX}) that does occur is mediated in large part by education. . . .

Our model also indicates where the “vicious circle” interpretation is vulnerable. In the passage on the vicious circle quoted, there seems to be an assumption that because of the substantial intercorrelations between a number of background factors, each of which has a significant relationship to subsequent achievement, the total effect of origin on achievement is materially enhanced. Here, in other words, the concept of “cumulation” appears to refer to the intercorrelations of a collection of independent variables. But the effect of such intercorrelations is quite opposite to what the writers appear to suppose. They are not alone in arguing from a fallacious assumption that was caustically analyzed by Karl Pearson half a century ago.¹⁸ The crucial point is that if the several determinants are indeed substantially intercorrelated with each other, then their combined effect will consist largely in redundancy, not in “cumulation.” This circumstance does not relieve us from the necessity of trying to understand better *how* the effects come about (a point also illustrated in a less fortunate way in Pearson’s work). It does imply that a refined esti-

mate of how much effect results from a combination of “associated factors” will not differ greatly from a fairly crude estimate based on the two or three most important ones.

NOTES

1. Mollie Orshansky, “Children of the Poor,” *Social Security Bulletin*, 26(July 1963).

2. Forrest A. Bogan, “Employment of High School Graduates and Dropouts in 1964,” *Special Labor Force Report*, No. 54 (U.S. Bureau of Labor Statistics, June 1965), p. 643.

3. Peter M. Blau and Otis Dudley Duncan, *The American Occupational Structure*, New York: The Free Press, 1967, ch. 4.

4. *Ibid.*

5. Bruce K. Eckland, “College Dropouts Who Came Back,” *Harvard Educational Review*, 34(1964), 402–420.

6. Beverly Duncan, *Family Factors and School Dropout: 1920–1960*, U.S. Office of Education, Cooperative Research Project No. 2258, Ann Arbor: Univ. of Michigan, 1965.

7. Blau and Duncan, *op. cit.*, Appendix C.

8. Gösta Carlsson, *Social Mobility and Class Structure*, Lund: CWK Gleerup, 1958, p. 124.

9. H. M. Blalock, Jr., *Causal Inferences in Nonexperimental Research*, Chapel Hill: Univ. of North Carolina Press, 1964.

10. Sewall Wright, “Path Coefficients and Path Regressions,” *Biometrics*, 16(1960), 189–202; Otis Dudley Duncan, “Path Analysis,” *American Journal of Sociology*, 72(1966), 1–16.

11. Blalock, *op. cit.*, pp. 54 ff.

12. Blau and Duncan, *op. cit.*, ch. 4.

13. *Ibid.*, ch. 3.

14. *Ibid.*, ch. 5.

15. Seymour M. Lipset and Reinhard Bendix, *Social Mobility in Industrial Society*, Berkeley: Univ. of California Press, 1959, pp. 198–199.

16. *Ibid.*, p. 190.

17. U.S. Bureau of the Census, “Lifetime Occupational Mobility of Adult Males: March 1962,” *Current Population Reports*, Series P–23, No. 11 (May 12, 1964), Table B.

18. Karl Pearson, “On Certain Errors with Regard to Multiple Correlation Occasionally Made by Those Who Have Not Adequately Studied This Subject,” *Biometrika*, 10(1914), 181–187.

51 ■ Christopher Jencks, Marshall Smith, Henry Acland, Mary Jo Bane, David Cohen, Herbert Gintis, Barbara Heyns, and Stephan Michelson

Inequality

A Reassessment of the Effect of Family and Schooling in America

Most Americans say they believe in equality. But when pressed to explain what they mean by this, their definitions are usually full of contradictions. Many will say, like the Founding Fathers, that “all men are created equal.” Many will also say that all men are equal “before God,” and that they are, or at least ought to be, equal in the eyes of the law. But most Americans also believe that some people are more competent than others, and that this will always be so, no matter how much we reform society. Many also believe that competence should be rewarded by success, while incompetence should be punished by failure. They have no commitment to ensuring that everyone’s job is equally desirable, that everyone exercises the same amount of political power, or that everyone receives the same income.

But while most Americans accept inequality in virtually every sphere of day-to-day life, they still believe in what they often call “equal opportunity.” By this they mean that the rules determining who succeeds and who fails should be fair. People are, of course, likely to disagree about precisely what is “fair” and what is “unfair.” Still, the general principle of fair competition is almost universally endorsed.

During the 1960s, many reformers devoted enormous effort to equalizing opportunity. More specifically, they tried to eliminate inequalities based on skin color, and to a lesser extent on economic background. They also wanted to eliminate absolute deprivation: “poverty,” “ignorance,” “powerlessness,” and so forth. But only a handful of radicals talked about eliminating inequality *per se*. Almost none of the national legislation passed during the 1960s tried to reduce disparities in adult status, power, or income in any direct way. There was no significant effort, for example, to make taxation more progressive, and very little effort to reduce wage disparities between highly paid and poorly paid workers. Instead, attention focused on helping workers in poorly paid jobs to move into better paid jobs. Nor was there much effort to reduce the social or psychological distance between high- and low-status occupations. Instead, the idea was to help people in low-status occupations leave these occupations for more prestigious ones. Even in the political arena, “maximum feasible participation” implied mainly that more “leaders” should be black and poor, not that power should be equally distributed between leaders and followers.

Because the reforms of the 1960s did not tackle the problem of adult inequality directly, they accomplished only a few of their goals. Equalizing opportunity is almost impossible without greatly reducing the absolute level of inequality, and the same is true of eliminating deprivation.

Consider the case of equal opportunity. One can equalize the opportunities available to blacks and whites without equalizing anything else, and considerable progress was made in this direction during the late 1960s. But equalizing the opportunities available to different children of the same race is far more difficult. If a society is competitive and rewards adults unequally, some parents are bound to succeed while others fail. Successful parents will then try to pass along their advantages to their children. Unsuccessful parents will inevitably pass along some of their disadvantages. Unless a society completely eliminates ties between parents and children, inequality among parents guarantees some degree of inequality in the opportunities available to children. The only real question is how serious these inequalities must be.

Or consider the problem of deprivation. When the war on poverty began in late 1963, it was conceived as an effort to raise the living standards of the poor. The rhetoric of the time described the persistence of poverty in the midst of affluence as a “paradox,” largely attributable to “neglect.” Official publications all assumed that poverty was an absolute rather than a relative condition. Having assumed this, they all showed steady progress toward the elimination of poverty, since fewer and fewer people had incomes below the official “poverty line.”

Yet despite all the official announcements of progress, the feeling that lots of Americans were poor persisted. The reason was that most Americans define poverty in relative rather than absolute terms. Public opinion surveys show, for example, that when people are asked how much money an American

family needs to “get by,” they typically name a figure about half what the average American family actually receives.¹ This has been true for the last three decades, despite the fact that real incomes (i.e. incomes adjusted for inflation) have doubled in the interval.

Political definitions of poverty have reflected these popular attitudes. During the Depression, the average American family was living on about \$30 a week. A third of all families were living on less than half this amount, i.e. less than \$15 a week. This made it natural for Franklin Roosevelt to speak of “one third of a nation” as ill-housed, ill-clothed, and ill-fed. One third of the nation was below what most people then regarded as the poverty line.

By 1964, when Lyndon Johnson declared war on poverty, incomes had risen more than fivefold. Even allowing for inflation, living standards had doubled. Only about 10 percent of all families had real incomes as low as the bottom third had had during the Depression. But popular conceptions of what it took to “get by” had also risen since the Depression. Mean family income was about \$160 a week, and popular opinion now held that it took \$80 a week for a family of four to make ends meet. About a quarter of all families were still poor by this definition. As a matter of political convenience, the Administration set the official poverty line at \$60 a week for a family of four rather than \$80, ensuring that even conservatives would admit that those below the line were poor. But by 1970 inflation had raised mean family income to about \$200 a week, and the National Welfare Rights Organization was rallying liberal support for a guaranteed income of \$100 a week for a family of four.

These political changes in the definition of poverty were not just a matter of “rising expectations” or of people’s needing to “keep up with the Joneses.” The goods and services that made it possible to live on \$15 a week during the Depression were no longer available to a family with the same “real” income

(i.e. \$40 a week) in 1964. Eating habits had changed, and many cheap foods had disappeared from the stores. Most people had enough money to buy an automobile, so public transportation had atrophied, and families without automobiles were much worse off than during the Depression. The labor market had also changed, and a person without a telephone could not get or keep many jobs. A home without a telephone was more cut off socially than when few people had telephones and more people “dropped by.” Housing arrangements had changed, too. During the Depression, many people could not afford indoor plumbing and “got by” with a privy. By the 1960s, privies were illegal in most places. Those who could not afford an indoor toilet ended up in buildings which had broken toilets. For this they paid more than their parents had paid for privies.

Examples of this kind suggest that the “cost of living” is not the cost of buying some fixed set of goods and services. It is the cost of participating in a social system. The cost of participation depends in large part on how much other people habitually spend to participate. Those who fall far below the norm, whatever it may be, are excluded. It follows that raising the incomes of the poor will not eliminate poverty if the incomes of other Americans rise even faster. If people with incomes less than half the national average cannot afford what “everyone” regards as “necessities,” the only way to eliminate poverty is to make sure everyone has an income at least half the average.

This line of reasoning applies to wealth as well as poverty. The rich are not rich because they eat filet mignon or own yachts. Millions of people can now afford these luxuries, but they are still not “rich” in the colloquial sense. The rich are rich because they can afford to buy other people’s time. They can hire other people to make their beds, tend their gardens, and drive their cars. These are not privileges that become more widely avail-

able as people become more affluent. If all workers’ wages rise at the same rate, the highly paid professional will have to spend a constant percentage of his income to get a maid, a gardener, or a taxi. The number of people who are “rich,” in the sense of controlling more than their share of other people’s time and effort, will therefore remain the same, even though consumption of yachts and filet mignon is rising.

If the distribution of income becomes more equal, as it did in the 1930s and 1940s, the number of people who are “rich” in this sense of the term will decline, even though absolute incomes are rising. If, for example, the wages of domestic servants rise faster than the incomes of their prospective employers, fewer families will feel they can afford full-time servants. This will lower the living standards of the elite to some extent, regardless of what happens to consumption of yachts and filet mignon.

This same logic applies not only to income but to the cognitive skills taught in school. Young people’s performance on standardized tests rose dramatically between World War I and World War II, for example. But the level of competence required for many adult roles rose too. When America was a polyglot nation of immigrants, all sorts of jobs were open to those who could not read English. Such people could, for example, join the army, drive a truck, or get a job in the construction industry. Today, when almost everyone can read English, the range of choices open to nonreaders has narrowed. The military no longer takes an appreciable number of illiterates, a driver’s license requires a written examination, and apprenticeships in the construction trades are restricted to those who can pass tests. Those who cannot read English are at a disadvantage, simply because they are atypical. America is not organized with their problems in mind. The same thing applies to politics. If the average citizen’s vocabulary expands, the vocabulary used by politicians

and newspapers will expand too. Those with very limited vocabularies relative to their neighbors will still have trouble following events, even though their vocabulary is larger than, say, their parents' vocabulary was.

Arguments of this kind suggest that it makes more sense to think of poverty and ignorance as relative than as absolute conditions. They also suggest that eliminating poverty and ignorance, at least as these are usually defined in America, depends on eliminating, or at least greatly reducing, inequality. This is no simple matter. Since a competitive system means that some people "succeed" while others "fail," it also means that people will end up unequal. If we want to reduce inequality, we therefore have two options. The first possibility is to make the system less competitive by reducing the benefits that derive from success and the costs paid for failure. The second possibility is to make sure that everyone enters the competition with equal advantages and disadvantages.

The basic strategy of the war on poverty during the 1960s was to try to give everyone entering the job market or any other competitive arena comparable skills. This meant placing great emphasis on education. Many people imagined that if schools could equalize people's cognitive skills this would equalize their bargaining power as adults. In such a system nobody would end up very poor—or, presumably, very rich.

This strategy rested on a series of assumptions which went roughly as follows:

1. Eliminating poverty is largely a matter of helping children born into poverty to rise out of it. Once families escape from poverty, they do not fall back into it. Middle-class children rarely end up poor.
2. The primary reason poor children do not escape from poverty is that they do not acquire basic cognitive skills. They cannot read, write, calculate, or articu-

late. Lacking these skills, they cannot get or keep a well-paid job.

3. The best mechanism for breaking this vicious circle is educational reform. Since children born into poor homes do not acquire the skills they need from their parents, they must be taught these skills in school. This can be done by making sure that they attend the same schools as middle-class children, by giving them extra compensatory programs in school, by giving their parents a voice in running their schools, or by some combination of all three approaches.

So far as we can discover, each of these assumptions is erroneous.

1. Poverty is not primarily hereditary. While children born into poverty have a higher-than-average chance of ending up poor, there is still an enormous amount of economic mobility from one generation to the next. Indeed, there is nearly as much economic inequality among brothers raised in the same homes as in the general population. This means that inequality is recreated anew in each generation, even among people who start life in essentially identical circumstances.
2. The primary reason some people end up richer than others is not that they have more adequate cognitive skills. While children who read well, get the right answers to arithmetic problems, and articulate their thoughts clearly are somewhat more likely than others to get ahead, there are many other equally important factors involved. Thus there is almost as much economic inequality among those who score high on standardized tests as in the general population. Equalizing everyone's reading scores would not appreciably reduce the number of economic "failures."

3. There is no evidence that school reform can substantially reduce the extent of cognitive inequality, as measured by tests of verbal fluency, reading comprehension, or mathematical skill. Neither school resources nor segregation has an appreciable effect on either test scores or educational attainment.

Our work suggests, then, that many popular explanations of economic inequality are largely wrong. We cannot blame economic inequality primarily on genetic differences in men's capacity for abstract reasoning, since there is nearly as much economic inequality among men with equal test scores as among men in general. We cannot blame economic inequality primarily on the fact that parents pass along their disadvantages to their children, since there is nearly as much inequality among men whose parents had the same economic status as among men in general. We cannot blame economic inequality on differences between schools, since differences between schools seem to have very little effect on any measurable attribute of those who attend them.

Economic success seems to depend on varieties of luck and on-the-job competence that are only moderately related to family background, schooling, or scores on standardized tests. The definition of competence varies greatly from one job to another, but it seems in most cases to depend more on personality than on technical skills. This makes it hard to imagine a strategy for equalizing competence. A strategy for equalizing luck is even harder to conceive.

The fact that we cannot equalize luck or competence does *not* mean that economic inequality is inevitable. Still less does it imply that we cannot eliminate what has traditionally been defined as poverty. It only implies that we must tackle these problems in a different way. Instead of trying to reduce people's capacity to gain a competitive advantage on one another, we would have to

change the rules of the game so as to reduce the rewards of competitive success and the costs of failure. Instead of trying to make everyone equally lucky or equally good at his job, we would have to devise "insurance" systems which neutralize the effects of luck, and income-sharing systems which break the link between vocational success and living standards.

This could be done in a variety of ways. Employers could be constrained to reduce wage disparities between their best- and worst-paid workers.² The state could make taxes more progressive, and could provide income supplements to those who cannot earn an adequate living from wages alone. The state could also provide free public services for those who cannot afford to buy adequate services in the private sector. Pursued with vigor, such a strategy would make "poverty" (i.e. having a living standard less than half the national average) virtually impossible. It would also make economic "success," in the sense of having, say, a living standard more than twice the national average, far less common than it now is. The net effect would be to make those with the most competence and luck subsidize those with the least competence and luck to a far greater extent than they do today.

This strategy was rejected during the 1960s for the simple reason that it commanded relatively little popular support. The required legislation could not have passed Congress. Nor could it pass today. But that does not mean it was the wrong strategy. It simply means that until we change the political and moral premises on which most Americans now operate, poverty and inequality of opportunity will persist at pretty much their present level.

At this point the reader may wonder whether trying to change these premises is worthwhile. Why, after all, should we be so concerned about economic equality? Is it not enough to ensure equal opportunity? And does not the evidence we have described sug-

gest that opportunities are already quite equal in America? If economic opportunities are relatively equal, and if the lucky and the competent then do better for themselves than the unlucky and incompetent, why should we feel guilty about this? Such questions cannot be answered in any definitive way, but a brief explanation of our position may help avoid misunderstanding.

We begin with the premise that every individual's happiness is of equal value. From this it is a short step to Bentham's dictum that society should be organized so as to provide the greatest good for the greatest number. In addition, we assume that the law of diminishing returns applies to most of the good things in life. In economic terms this means that people with low incomes value extra income more than people with high incomes.³ It follows that if we want to maximize the satisfaction of the population, the best way to divide any given amount of money is to make everyone's income the same. Income disparities (except those based on variations in "need") will always reduce overall satisfaction, because individuals with low incomes will lose more than individuals with high incomes gain.

The principal argument against equalizing incomes is that some people contribute more to the general welfare than others, and that they are therefore entitled to greater rewards. The most common version of this argument is that unless those who contribute more than their share are rewarded (and those who contribute less than their share punished) productivity will fall and everyone will be worse off. A more sophisticated version is that people will only share their incomes on an equal basis if all decisions that affect these incomes are made collectively. If people are left free to make decisions on an individual basis, their neighbors cannot be expected to pay the entire cost of their mistakes.

We accept the validity of both these arguments. We believe that men need incentives to contribute to the common good, and we

prefer monetary incentives to social or moral incentives, which tend to be inflexible and very coercive. We believe, in other words, that virtue should be rewarded, and we assume that there will be considerable variation in virtue from one individual to another. This does not, however, mean that incomes must remain as unequal as they are now. Even if we assume, for example, that the most productive fifth of all workers accounts for half the Gross National Product, it does not follow that they need receive half the income. A third or a quarter might well suffice to keep both them and others productive.

Most people accept this logic to some extent. They believe that the rich should pay more taxes than the poor, although they often disagree about how much more. Conversely, they believe that the poor should not starve, even if they contribute nothing to the general welfare. They believe, in other words, that people should not be rewarded solely for their contribution to the general welfare, but that other considerations, such as need, should also be taken into account. Our egalitarianism is simply another way of saying that we think need should play a larger role than it now does in determining what people get back from society. We do not think it can or should be the sole consideration.

When we turn from the distribution of income to the distribution of other things, our commitment to equality is even more equivocal. We assume, for example, that occupational prestige resembles income in that those who have low-prestige occupations usually value additional prestige more than those who have high-prestige occupations. Insofar as prestige is an end in itself, then, the optimal distribution is again egalitarian. But occupational prestige derives from a variety of factors, most of which are more difficult to redistribute than income. We cannot imagine a social system in which all occupations have equal prestige, except in a society where all workers are equally competent. Since we do not see any likelihood of equalizing

competence, we regard the equalization of occupational prestige as a desirable but probably elusive goal.

When we turn from occupational prestige to educational attainment and cognitive skills, the arguments for and against equality are reversed. If schooling and knowledge are thought of strictly as ends in themselves, it is impossible to make a case for distributing them equally. We can see no reason to suppose, for example, that people with relatively little schooling value additional schooling more than people who have already had a lot of schooling. Experience suggests that the reverse is the case. Insofar as schooling is an end in itself, then, Benthamite principles imply that those who want a lot should get a lot, and those who want very little should get very little. The same is true of knowledge and cognitive skills. People who know a lot generally value additional knowledge and skills more than those who know very little. This means that insofar as knowledge or skill is valued for its own sake, an unequal distribution is likely to give more satisfaction to more people than an equal distribution.

The case for equalizing the distribution of schooling and cognitive skill derives not from the idea that we should maximize consumer satisfaction, but from the assumption that equalizing schooling and cognitive skill is necessary to equalize status and income. This puts egalitarians in the awkward position of trying to impose equality on people, even though the natural demand for both cognitive skill and schooling is very unequal. Since we have found rather modest relationships between cognitive skill and schooling on the one hand and status and income on the other, we are much less concerned than most egalitarians with making sure that people end up alike in these areas.

Our commitment to equality is, then, neither all-embracing nor absolute. We do not believe that everyone can or should be made equal to everyone else in every respect. We assume that some differences in cognitive skill and vocational competence are inevitable, and that efforts to eliminate such differences can never be 100 percent successful. But we also believe that the distribution of income can be made far more equal than it is, even if the distribution of cognitive skill and vocational competence remains as unequal as it is now. We also think society should get on with the task of equalizing income rather than waiting for the day when everyone's earning power is equal.

NOTES

1. This material has been collected and analyzed by Lee Rainwater at Harvard University, as part of a forthcoming study of the social meaning of low income.

2. Lester C. Thurow and Robert E.B. Lucas, in "The American Distribution of Income" [Washington, D.C.: U.S. Government Printing Office, March 17, 1972], discuss the possibility of such constraints in some detail. The principal virtue of this approach is that it reduces the incomes of the rich *before* they are defined as "income" rather than afterwards. This means that the recipient is less conscious of what he is giving up and less likely to feel he is being cheated of his due.

3. If everyone had equal earning power we could assume that people "chose" their incomes voluntarily and that those with low incomes were those who were maximizing something else (e.g. leisure, autonomy, etc.). But as we note [elsewhere], people's concern with income as against other objectives has no apparent effect on their actual income, at least while they are young [see Christopher Jencks, Marshall Smith, Henry Acland, Mary Jo Bane, David Cohen, Herbert Gintis, Barbara Heyns, and Stephan Michelson, *Inequality: A Reassessment of the Effect of Family and Schooling in America*, New York: Harper and Row, 1972, ch. 7, note 64]. Thus we infer that income differences derive largely from differences in earning power and luck.

52 ■ David J. Harding, Christopher Jencks,
Leonard M. Lopoo, and Susan E. Mayer

Family Background and Income in Adulthood, 1961–1999

Most Americans endorse the ideal of equal opportunity, and many interpret this ideal as requiring that children from different backgrounds have an equal chance of achieving economic success. Most Americans also recognize that children whose parents have “all the advantages” are more likely to prosper than children whose parents lack these advantages. Understanding this relationship has therefore become a prominent goal of social research, especially since the 1960s. This chapter investigates how the relationship between American family income during adulthood and family background during childhood changed from 1961 to 1999.

Our approach to measuring trends in intergenerational inheritance differs from earlier studies. Almost all earlier studies have focused on the determinants of individuals’ occupational rank or earnings. We focus instead on an individual’s total family income. This change allows us to assess the impact of family background on the economic status of individuals who are not working. It also allows us to take account of the fact that individuals’ economic status often depends more on how they fare in the “marriage market” than on how they fare in the labor market.

The changing importance of marriage for one’s family income is clear when we examine changes in the correlation between a mature adult’s own earnings in a given year and his or her total family income in the same year. Among men the correlation¹ between earnings and total family income fell from 0.87 in 1968 to 0.75 in 1996.² Among women, the correlation rose from 0.18 in 1968 to 0.39 in 1996. Earnings are still a much better proxy for living standards among men than among women. But since family background influences an individual’s chance of having a well-paid spouse, ignoring this fact is likely to bias trend estimates of how background affects economic status among both men and women.

How Family Background Affects Adult Family Income

Figure 1 shows some of the pathways by which family background can influence children’s eventual income in adulthood.³ It contains several concepts that require brief discussion.

Genetic Advantages

A growing body of evidence suggests that genetic differences can influence an individual’s

David J. Harding, Christopher Jencks, Leonard M. Lopoo, and Susan E. Mayer. This chapter draws heavily on material in a previous publication (David J. Harding, Christopher Jencks, Leonard M. Lopoo, and Susan E. Mayer. “The Changing Effect of Family Background on the Incomes of American Adults,” in *Unequal Chances: Family Background and Economic Success*, Samuel Bowles, Herbert Gintis, and Melissa Osborne, eds). Copyright © 2005 by Russell Sage Foundation and Princeton University Press.

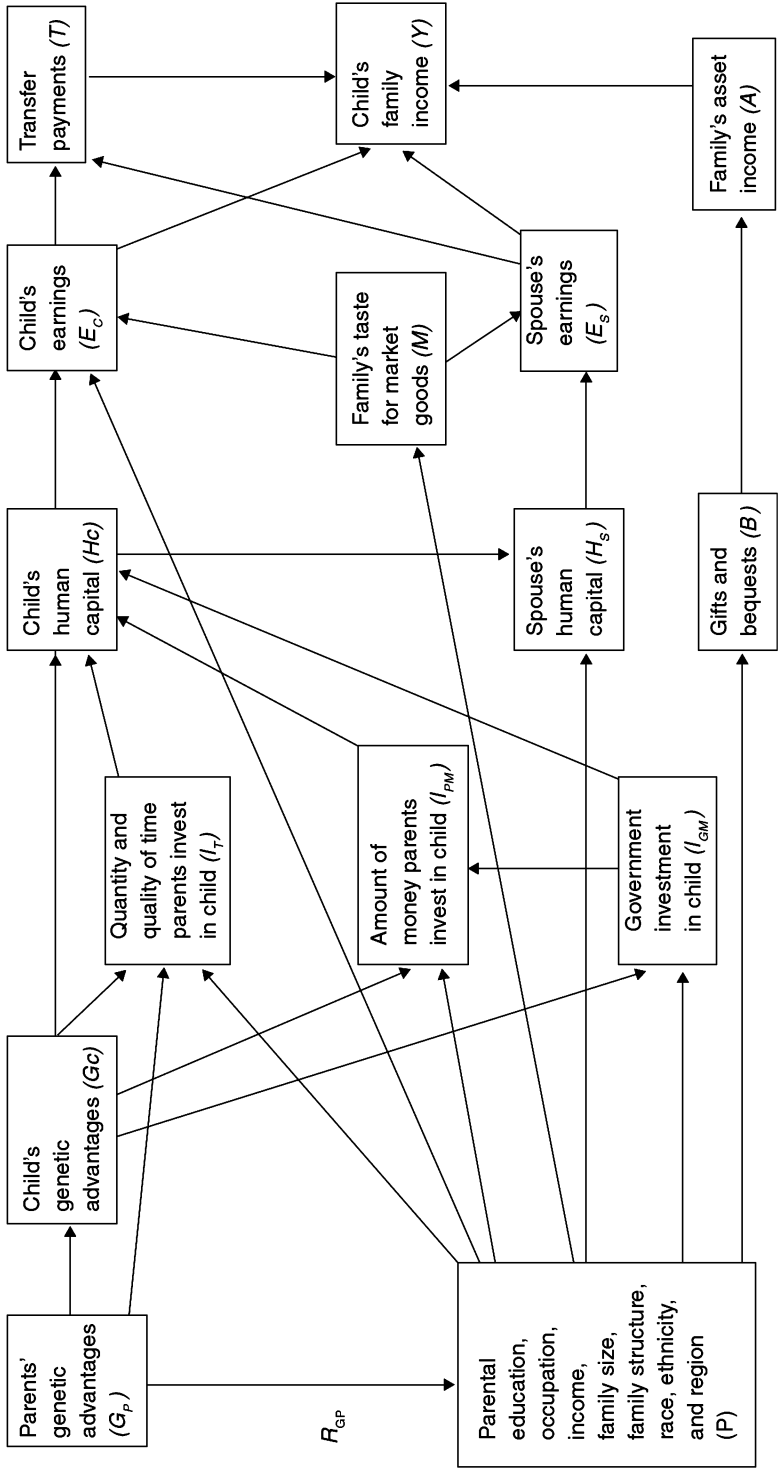


Figure 1. Some Possible Links Between Respondents' Adult Incomes and Their Family Background

earnings. Studies based on comparisons of monozygotic (MZ) twins and dizygotic (DZ) twins (Ashenfelter and Rouse 1998; Taubman 1976) suggest that genes are likely to play a significant role in explaining the intergenerational inheritance of economic advantages.⁴ However, because estimates based on twin studies require a number of questionable assumptions (Björklund, Jäntti, and Solon 2005; Goldberger 1978), the exact magnitude of the importance of genes is difficult to measure. Björklund and Chadwick's (2002) study of the earnings of fathers and sons in Sweden also suggests that genes play a role.

However, the fact that genetic resemblance helps explain economic resemblance between parents and children does not mean that this source of economic resemblance operates independently of the environment. Suppose, for example, that myopia has a genetic component. Most affluent societies ensure that myopic children get glasses. As a result, myopic children can usually see almost as well as other children. In poor societies where glasses are not available to everyone, the genes that contribute to myopia have a larger impact. People's genes also influence the environments they choose for themselves. Consider two sisters, one of whom finds reading easy and one of whom finds it difficult. The sister who finds reading easier is likely to enjoy reading more and do more of it. As a result, she is likely to score higher on vocabulary and reading tests, get higher grades, stay in school longer, and earn more when she enters the labor force.

Human Capital

Economists use the term "human capital" to describe the skills, knowledge, and character traits that influence a worker's potential earnings. If employers value the same attributes for several generations, parents with above-average human capital are likely to have children with above-average human capital. To

begin with, parents with traits that employers value tend to have above-average earnings, so they can invest more in their children's nutrition, health care, and education. Such parents also tend to live in political jurisdictions where the government invests a lot in children.

Although economists use the term "human capital" to describe characteristics that pay off in the labor market, one can easily broaden the concept to include characteristics that pay off in the marriage market. As we shall see, family background explains slightly more of the variation in daughters' family incomes than in sons' family incomes. This difference suggests that family background may exert even more influence on a daughter's success in the marriage market than on a son's success in the labor market.

Taste for Market Goods

Families vary dramatically in the number of hours that their adult members spend doing paid labor. Family members also vary in the extent to which they maximize their hourly wages. Some always take the best-paid job they can find, and others settle for less money to get shorter hours, more congenial colleagues, more interesting work, or other nonmonetary advantages. Family background appears to influence how much people work (Altonji and Dunn 2000), and it probably also influences the weight that individuals assign to earnings relative to goals like short hours and interesting work. If these associations change over time, the association between adult income and family background is also likely to change.

Gifts and Bequests

Dividends, interest, and rent accounted for 7 percent of all income received by individuals between the ages of 25 and 64 in 1999.⁵ Estimates suggest that income from assets accounts for between 1 and 5 percent

of all money income among thirty- to fifty-nine-year-olds (Gale and Scholz 1994; Wolff 2002). We do not know how this figure has changed over time, but the age at which adults inherit their parents' assets is rising, so the relative importance of inherited assets could be falling for those under the age of sixty.

Two Sources of Change over Time

Figure 1 suggests that the effects of family background on incomes can change for two quite different reasons. First, the effects of various family background characteristics on children's characteristics, such as years of schooling, cognitive ability, or taste for market goods, can change. Second, the effects of children's characteristics on their eventual adult income can change. Children's education illustrates both of these possibilities.

The association between parents' socioeconomic advantages and the number of years of school that their children complete declined between World War I and the early 1970s (Hauser and Featherman 1976). As a result, educational disparities among adults from different kinds of families kept shrinking throughout the twentieth century. This trend should have reduced the impact of family background on children's eventual family income. However, the economic value of an extra year of schooling also changed. The effect of an extra year of school on men's annual earnings fell between 1940 and 1975. This change accentuated the declining impact of family background on children's educational attainment, making it doubly difficult for economically advantaged parents to pass along their advantages by keeping their children in school. But after 1975 the labor market returns to schooling rose.

The payoff to schooling also rose in the marriage market. If we compare the most educated third of all mothers to the least edu-

cated third, about 93 percent of the most educated and 87 percent of the least educated were living with a husband in both 1940 and 1960. By 1990 the proportions were 84 percent for the most educated and 70 percent for the least educated.⁶ The marriage gap between the most and least educated mothers had thus risen from 6 to 14 percentage points. The net result was that although family background had less impact on how much schooling young people got in the last third of the century than earlier in the century, the value of each year of school rose during the last decades of the century.

Mature adults seldom go back to school, so the association between educational attainment and family background when a birth cohort reaches sixty is about the same as when it was thirty. The association does change as more recent birth cohorts replace earlier ones, but this takes a long time. In principle, we can hold changes of this kind constant by tracking the same birth cohort over time. When we do this, almost all the remaining changes in the association between adult family income and family background are attributable to changes in the way employment, marriage, and earnings are distributed among individuals with different background characteristics.

Whereas cohort replacement is slow, the economic value of respondents' characteristics can change relatively rapidly. When job opportunities improved for black workers in the wake of the civil rights movement, for example, young blacks benefited the most, but older blacks also gained. Likewise, when the wage gap between high school and college graduates widened in the 1980s, the change was largest among younger workers but it also affected mature workers. Changes in the marriage market can also change the value of personal characteristics quite rapidly. When divorce rates rise, all women are more likely to experience sudden reductions in family income, but the change is

likely to be largest among those with disadvantaged parents, because their divorce rates are likely to rise the most.

Data

We concentrate on trends among men and women between the ages of thirty and fifty-nine, whom we refer to as “mature adults.” We use two sources of data, the Occupational Changes in a Generation (OCG) surveys and the General Social Survey (GSS). The OCG surveys were conducted in 1962 and 1973. They both asked men (but not women) about their family background. Our first sample of mature OCG men was born between 1902 and 1932. Our second sample was born between 1913 and 1943. The GSS is a smaller survey that has been conducted either annually or biennially since 1972. The GSS asks both men and women about their family income and family background. We analyze the surveys conducted between 1972 and 2000, so our samples of mature adults were born between 1912 and 1970.

Our primary dependent variable is family income in the year preceding the survey. Although there is some variation across data sources in how our variables are measured, one can think of adult family income as the sum of all sources of income for the respondent and his or her spouse (if married).⁷ Variation in income due to survey year, gender, and age has been removed, and we analyze the effect of parental characteristics on the natural logarithm of this adjusted income. We measure family background using seven characteristics of the family in which the respondent grew up, including the family head’s occupational ranking (measured by an updated version of Duncan’s socioeconomic index, which is based on occupations’ educational requirements and economic rewards; see Hauser and Warren 1997), the family head’s education in years, the number of sib-

lings, and whether the respondent is black, Hispanic, grew up in the South, and/or grew up in a family with both parents.

Changes in Inequality of Family Background

Suppose that a 10 percent income difference between parents was always associated with a 4 percent income difference between their children. The income gap between children raised by rich and poor parents would then depend on the income gap separating rich and poor parents. If inequality between rich and poor parents falls, as it did during the first two-thirds of the twentieth century, and if all else remains equal, the subsequent income gap between their children will also fall. If inequality between rich and poor parents rises, as it did in the last third of the twentieth century, and if all else remains equal, the income gap between their children will also widen. The same logic holds for other parental advantages, like educational attainment and occupational position.

The degree to which the distribution of background characteristics changed over time varies according to which family background characteristics we examine. Inequality in parental occupation rose by about 15 percent between 1962 and 1999. Inequality in parental years of schooling rose slightly between 1962 and 1973 and declined slightly thereafter, so educational inequality among parents was about the same in 1999 as in 1962. Disparities in the size of children’s families declined by about 10 percent between the early 1960s and the 1990s.

Of the four remaining measures of background characteristics—intact family, black, Hispanic, and Southern origins, the only one that changes appreciably is Hispanic. The GSS shows substantial growth in the English-speaking Hispanic population between the 1970s and 1990s. Because Hispanics have

Table 1. Multiple Correlations of Family Income with Seven Family Background Measures in OCG and GSS, By Age, Sex, and Decade

| <i>Respondents' Age</i> | <i>OCG Men</i> | | <i>GSS Men</i> | | | <i>GSS Women</i> | | |
|---|----------------|-------------------|----------------|--------|--------|------------------|--------|-------------------|
| | 1961 | 1972 ^a | 1970s | 1980s | 1990s | 1970s | 1980s | 1990s |
| 30 to 39 | .399 ↘ | .309 | .330 ↘ | .289 ↘ | .306 | .399 ↘ | .361 ↘ | .374 |
| 40 to 49 | .397 ↘ | .305 | .355 ↘ | .362 ↘ | .352 | .437 ↘ | .388 ↘ | .361 |
| 50 to 59 | .414 ↘ | .323 | .327 ↘ | .410 ↘ | .360 | .440 ↘ | .429 ↘ | .350 |
| Weighted mean for all 30- to 59-year-olds | .403 | .312 | .337 | .342 | .335 | .423 | .385 | .364 ^b |
| Approximate sampling error of multiple R^c | (.008) | (.006) | (.014) | (.017) | (.015) | (.012) | (.015) | (.014) |
| Sample size | 12,829 | 23,824 | 2,382 | 3,040 | 3,817 | 2,761 | 3,703 | 4,657 |

Note: The seven measures of family background are race, ethnicity, Southern origins, living with both parents at age 16, number of siblings, parental education, and parental occupation (see text for details).

^a The decline between 1961 and 1972 is significant at the .01 level for all age groups.

^b The average decline between the 1970s and the 1990s is significant at the .01 level.

^c Approximated using $[(1-R^2)/(N-8)]^5$.

relatively low family incomes, the fact that more adults had Hispanic parents implies an increase in the overall level of inequality among parents.

Changes in the Intergenerational Inheritance of Advantage

We turn now to changes in the relationship between mature adults' income and their family background (as measured by the seven background characteristics). We measure this relationship with the multiple correlations between income and family background characteristics, which we denote as R_{pc} .⁸ When R_{pc} falls, we can infer that intergenerational changes in rank (what sociologists call "exchange mobility") have become larger or more frequent. When R_{pc} rises, intergenerational changes in rank have become either smaller or less frequent.

Table 1 shows multiple correlations between respondents' adult family income and their background characteristics. Among OCG men we see a dramatic decline in R_{pc} between 1961 and 1972. This decline recurs for all age groups. The diagonal arrows

track birth cohorts as they age. Here again we see declines between 1961 and 1972. The GSS tells a very different story about changes in R_{pc} after 1972. The multiple correlations for men in different age groups show no consistent trend between the 1970s and the 1990s. When we track the same cohort of men over time, the multiple correlations almost always rise as a cohort ages. This is in sharp contrast to what we saw between 1961 and 1972.

The GSS also provides data on post-1972 trends among women. Table 1 shows that although R_{pc} was essentially stable among GSS men, it declined by about one-seventh among GSS women. Whereas R_{pc} was substantially higher among women than men in the 1970s, this gender difference had largely disappeared by the 1990s. This pattern of differences is consistent with the hypothesis that family background influences success in the marriage market somewhat more than it influences success in the labor market, although other explanations are also possible.

For those who find it hard to decide whether the correlations in Table 1 are "large"

Table 2. Percentage of 30- to 59-Year-Old Sons and Daughters in Each Quartile of Family Income By Quartile of Family Background and Year in OCG and GSS

| <i>Parents' Quartile and Survey</i> | <i>Family Income Quartile</i> | | | | Total | Sample Size |
|-------------------------------------|-------------------------------|--------|-------|------|-------|-------------|
| | Bottom | Second | Third | Top | | |
| Sons from bottom quartile | | | | | | |
| OCG: 1961 | 46.2 | 27.7 | 16.7 | 9.5 | 100 | 2,537 |
| OCG: 1972 | 40.7 | 29.8 | 21.2 | 8.4 | 100 | 6,454 |
| GSS: 1970s | 40.3 | 29.6 | 19.8 | 10.2 | 100 | 595 |
| GSS: 1980s | 39.4 | 30.0 | 21.2 | 9.4 | 100 | 840 |
| GSS: 1990s | 40.6 | 25.8 | 21.1 | 12.6 | 100 | 954 |
| Trend: 1961 to 1990s ^a | -5.3 | -1.7 | 5.7 | 1.4 | | |
| Sons from top quartile | | | | | | |
| OCG-I | 9.5 | 20.9 | 32.1 | 37.6 | 100 | 2,678 |
| OCG-II | 11.8 | 24.0 | 29.2 | 35.0 | 100 | 4,899 |
| GSS: 1970s | 11.6 | 20.5 | 30.2 | 37.8 | 100 | 596 |
| GSS: 1980s | 12.3 | 19.7 | 28.8 | 39.2 | 100 | 728 |
| GSS: 1990s | 11.6 | 21.4 | 29.0 | 38.0 | 100 | 954 |
| Trend: 1961 to 1990s ^a | 2.4 | 3.9 | -4.1 | -2.3 | | |
| Daughters from bottom quartile | | | | | | |
| GSS: 1970s | 44.4 | 26.2 | 19.6 | 9.9 | 100 | 690 |
| GSS: 1980s | 44.5 | 25.5 | 18.8 | 11.2 | 100 | 1,070 |
| GSS: 1990s | 43.7 | 25.9 | 18.8 | 11.6 | 100 | 1,164 |
| Trend: 1970s to 1990s | -0.6 | -0.4 | -0.8 | 1.7 | | |
| Daughters from top quartile | | | | | | |
| GSS: 1970s | 10.1 | 19.1 | 27.1 | 42.8 | 100 | 691 |
| GSS: 1980s | 12.1 | 19.3 | 29.7 | 38.9 | 100 | 878 |
| GSS: 1990s | 11.7 | 23.5 | 27.1 | 37.7 | 100 | 1,164 |
| Trend: 1970s to 1990s | 1.5 | 4.4 | 0.0 | -5.2 | | |

Parents are assigned to quartiles of parental advantages using the same seven measures as in Table 1, with each advantage weighted by its coefficient in the equation predicting log family income. Family incomes are assigned to quartiles after eliminating variation due to age, gender, and survey year.

^a Sum of within-survey changes: (OCG II – OCG I) + (GSS 1990s – GSS 1970s).

or “small,” Table 2 displays the association between family background and income in adulthood in a more intuitively understandable form. Table 2 shows the proportion of respondents in the top and bottom quartiles of family background who ended up in each of the four income quartiles as adults. We focus on the probability that those born into the top or bottom quartile were in the top or bottom income quartile as mature adults.

Consistent with the results in Table 1, the proportion of men who remained in the bottom or top quartiles declined be-

tween 1961 and 1972 and stayed the same afterward.

Among women the story is more complicated. Among GSS women, R_{pc} fell between the 1970s and the 1990s, so we expected a parallel decline in the proportions of GSS women who stayed in both the top and bottom quartiles. Table 2 confirms this prediction for women born into the top quartile: their chances of having slipped into a lower quartile were clearly higher in the 1990s than in the 1970s. However, there was little change in the percentage of women born into the bottom quartile who had risen into a higher quartile,

although those women who did move up moved further in the 1990s than in the 1970s.

Specific Background Characteristics

Previously we examined the relationship between adult income and family background by combining the seven background measures. Now we separate these characteristics to see which ones became more and less important. Column 1 of Table 3 shows the bivariate relationship of logged family income in 1961 to each of the seven family background characteristics. Column 2 shows the change in this coefficient between 1961 and 1972. Almost all the coefficients decline significantly between 1961 and 1972. However, when we examine each coefficient controlling for the other six, only the changes in the effects of parental occupation, black race, and Southern origins remain statistically significant (results not shown).

The decline in the effect of parental occupation is partly attributable to the steady decline in the proportion of respondents whose fathers were farmers. Among respondents not raised on farms, the main reason for decline was that respondents from the top tenth of the occupational distribution enjoyed less of an economic advantage in 1972 than in 1961 (results not shown). The declining effect of Southern origins probably reflects the narrowing of the income gap between the South and the North. The declining effect of race reflects a decline in the educational gap between blacks and whites, movement of blacks from the South to the North, and some reduction in discrimination.

The remainder of Table 3 shows the bivariate results from GSS. Column 3 shows the estimated coefficients of family background measures for men in 1971, and column 4 shows the estimated change in the bivariate coefficients for men between 1971 and 1999. The bivariate relationships of

men's family income to their parents' education, number of siblings, and whether they grew up in an intact family become significantly stronger over time. The bivariate effect of parental occupation also grows slightly stronger, but the change is not significant. The bivariate effect of being Hispanic grows weaker over time.

Columns 5 and 7 show that in 1971 almost every family background measure had more impact on women than on men. Columns 6 and 8 show that most of the coefficients for men and women converged between 1971 and 1999. The effects of parental education, race, and Southern origins all weakened significantly for women and were very close to those for men by the late 1990s. The adverse effect of coming from a large family was significantly larger for women than men in 1971, but this difference had also disappeared by the late 1990s. The multivariate GSS results (not shown) are broadly similar to the bivariate results except that the coefficients are smaller.

Summary

Intergenerational Inheritance

The relationship between a mature man's family income in adulthood and his family background fell during the 1960s and then remained stable from the 1970s through the 1990s. But men's incomes began growing more unequal in the 1970s, so the income gap between men raised by advantaged rather than disadvantaged parents widened between the 1970s and the 1990s.

We have no data on trends among women during the 1960s. In the early 1970s intergenerational correlations were higher among daughters than among sons, and daughters' family incomes were also more unequal. The economic cost of having grown up in a disadvantaged family was therefore considerably larger for daughters than for sons. But

Table 3. Bivariate Regression Coefficients of Seven Family Background Measures When Predicting Log Family Income at Ages 30 to 59 Among OCG Men and GSS Men and Women

| <i>Background Measure</i> | <i>OCG Men</i> | | | <i>GSS Men</i> | | | <i>GSS Women</i> | | | <i>GSS Women Minus Men</i> | | |
|---------------------------|----------------------|--------------------------|----------------------|--------------------------|----------------------|--------------------------|----------------------|--------------------------|----------------------|----------------------------|----------------------|--------------------------|
| | Coefficient for 1961 | Change from 1961 to 1972 | Coefficient for 1971 | Change from 1971 to 1999 | Coefficient for 1971 | Change from 1971 to 1999 | Coefficient for 1971 | Change from 1971 to 1999 | Coefficient for 1971 | Change from 1971 to 1999 | Coefficient for 1971 | Change from 1971 to 1999 |
| Father's occupation | | | | | | | | | | | | |
| <i>B</i> | .257 | -0.089 | .131 | .037 | .207 | -.010 | .076 | -.047 | .076 | -.047 | .076 | -.047 |
| (S.E.) | (.007) | (.009) | (.017) | (.026) | (.017) | (.028) | (.024) | (.038) | (.024) | (.038) | (.024) | (.038) |
| Parent's education | | | | | | | | | | | | |
| <i>B</i> | .053 | -0.012 | .037 | .015 | .057 | -.010 | .020 | -.025 | .020 | -.025 | .020 | -.025 |
| (S.E.) | (.002) | (.002) | (.004) | (.006) | (.004) | (.007) | (.006) | (.009) | (.006) | (.009) | (.006) | (.009) |
| Number of siblings | | | | | | | | | | | | |
| <i>B</i> | -0.058 | -0.011 | -0.043 | -0.029 | -0.067 | .007 | -0.024 | .036 | -0.024 | .036 | -0.024 | .036 |
| (S.E.) | (.002) | (.003) | (.006) | (.010) | (.006) | (.010) | (.008) | (.014) | (.008) | (.014) | (.008) | (.014) |
| Intact family | | | | | | | | | | | | |
| <i>B</i> | .164 | .033 | .090 | .205 | .262 | .081 | .172 | -.124 | .172 | -.124 | .172 | -.124 |
| (S.E.) | (.018) | (.024) | (.039) | (.062) | (.039) | (.062) | (.055) | (.088) | (.055) | (.088) | (.055) | (.088) |
| Black | | | | | | | | | | | | |
| <i>B</i> | -0.824 | .359 | -0.415 | -0.065 | -0.770 | .179 | -0.355 | .244 | -0.355 | .244 | -0.355 | .244 |
| (S.E.) | (.023) | (.031) | (.051) | (.081) | (.051) | (.079) | (.072) | (.133) | (.072) | (.133) | (.072) | (.133) |
| Hispanic | | | | | | | | | | | | |
| <i>B</i> | -0.405 | .082 | -0.573 | .443 | -0.482 | .168 | .091 | -.275 | .091 | -.275 | .091 | -.275 |
| (S.E.) | (.047) | (.055) | (.095) | (.147) | (.095) | (.144) | (.134) | (.206) | (.134) | (.206) | (.134) | (.206) |
| Southern regions | | | | | | | | | | | | |
| <i>B</i> | -0.424 | .191 | -0.260 | .093 | -0.367 | .226 | -0.107 | .133 | -0.107 | .133 | -0.107 | .133 |
| (S.E.) | (.015) | (.019) | (.035) | (.057) | (.036) | (.060) | (.050) | (.083) | (.050) | (.083) | (.050) | (.083) |
| Unweighted sample | | | | | | | | | | | | |
| | | 32,242 | | 9,268 | | 11,107 | | | | | | |

*Coefficients in boldface are significant at the .05 level.

whereas the association between a mature man's family income and his parents' socioeconomic position was almost constant between the 1970s and the 1990s, this association fell among women. As a result, the economic cost of growing up in a disadvantaged family was roughly constant for women, although it remained larger for women than for men.

Race, Ethnicity, and Region

Disparities in adult family income between blacks and whites, between Hispanics and Anglos, and between those raised in the South and North narrowed between 1961 and 1999.

Parental Education

The effect of parental education on men's family income fell during the 1960s but rose again over the next three decades. Parental education mattered more for mature women's family income than for men's in the 1970s, but this difference had disappeared by the late 1990s.

Parental Occupation

When we rank occupations in terms of their educational requirements and economic rewards, the effect of differences between fathers' occupations on their mature sons' family incomes declined during the 1960s. The effect of a father's occupation shows no clear trend after that for either men or women, but it is consistently larger for women than for men, partly because women's family incomes are more unequal than men's.

Taken together these findings suggest that there were significant changes in equality of opportunity in the United States between 1961 and 1999. For men, equality of opportunity increased during the 1960s but changed little thereafter. Among women there was less equality of opportunity in the

early 1970s than among men, but equality of opportunity among women increased during the 1970s. By the late 1990s, the importance of family background for a woman's economic prospects (relative to other women) was similar to its importance for a man's economic prospects (relative to other men). In addition, the importance of race, ethnicity, and region declined between 1961 and 1999. Nevertheless, as we begin the twenty-first century, there are still sizable inequalities of opportunity in the United States. Gender and family background, including race, ethnicity, and region, remain strong predictors of economic well-being in adulthood.

NOTES

1. In this chapter we often measure the relationship between two attributes with a "correlation," which ranges from zero (no relationship) to one (perfect relationship) and which can be thought of as assessing the degree to which two characteristics are associated in the data. A "multiple correlation" can be thought of as a correlation between one attribute and a set of attributes considered together.

2. Men and women without earnings are included and assigned earnings of zero.

3. Figure 1 is not a complete causal model. It omits many links that are well established in the research literature, such as the reciprocal effects of spouses' earnings on one another, to highlight the links that we think are most likely to have changed.

4. Monozygotic (MZ) twins have all their genes in common, and dizygotic (DZ) twins share roughly half their genes.

5. Estimate based on data from U.S. Bureau of the Census (2000, pp. 55 and D-4). Our estimates of family income do not include the implicit return on home equity.

6. These estimates are from tabulations by Andrew Clarkwest using the Integrated Public Use Microsamples from the decennial census.

7. See Harding et al. (2005) for further details on the data and methodology.

8. The square of the multiple correlation is the percentage of the income variance explained by parental characteristics.

REFERENCES

- Altonji, Joseph, and Thomas Dunn. 2000. "An Intergenerational Model of Wages, Hours, and Earnings." *Journal of Human Resources* 35(2): 221–258.
- Ashenfelter, Orley, and Cecelia Rouse. 1998. "Income, Schooling, and Ability: Evidence from a New Sample of Identical Twins." *Quarterly Journal of Economics* 113: 317–323.
- Björklund, Anders, and Laura Chadwick. 2002. "Intergenerational Income Mobility in Permanent and Separated Families." Swedish Institute for Social Research, Stockholm University.
- Björklund, Anders, Markus Jäntti, and Gary Solon. 2005. "Influences of Nature and Nurture of Earnings: A Report on a Study of Various Sibling Types in Sweden," in Samuel Bowles, Herbert Gintis, and Melissa Osborne, eds., *Unequal Chances: Family Background and Economic Success*. Princeton, NJ: Princeton University Press and Russell Sage.
- Gale, William G., and John Karl Scholz. 1994. "Intergenerational Transfers and the Accumulation of Wealth." *Journal of Economic Perspectives* 8(4): 145–160.
- Goldberger, Arthur. 1978. "The Genetic Determination of Income: Comment." *American Economic Review* 68(5): 960–969.
- Harding, David J., Christopher Jencks, Leonard M. Lopoo, and Susan E. Mayer. 2005. "The Changing Effect of Family Background on the Incomes of American Adults." In Samuel Bowles, Herbert Gintis, and Melissa Osborne, eds., *Unequal Chances: Family Background and Economic Success*. Princeton, NJ: Princeton University Press and Russell Sage.
- Hauser, Robert, and David Featherman. 1976. "Equality of Schooling: Trends and Prospects." *Sociology of Education* 49(2): 99–120.
- Hauser, Robert, and John Robert Warren. 1997. "Socioeconomic Indexes for Occupations: A Review, Update, and Critique." *Sociological Methodology* 27: 177–298.
- Taubman, Paul. 1976. "The Determinants of Earnings: Genetics, Family, and Other Environments; A Study of Male Twins." *American Economic Review* 66(5): 858–870.
- U.S. Bureau of the Census. 2000. "Money Income in the United States, 1999." *Current Population Reports*. Washington, DC: Government Printing Office.
- Wolff, Edward. 2002. "Inheritances and Wealth Inequality, 1989–1998." *American Economic Review* 92(2): 260–264.

■ SOCIAL PSYCHOLOGICAL MODELS ■

53 ■ William H. Sewell, Archibald O. Haller, and Alejandro Portes

The Educational and Early Occupational Attainment Process

Blau and Duncan (1967:165–172) have recently presented a path model of the occupational attainment process of the American adult male population. This basic model begins with two variables describing the early stratification position of each person; these are his father's educational and occupational attainment statuses. It then moves to two behavioral variables; these are the educational level the individual has completed and the prestige level of his first job. The dependent variable is the person's occupational prestige position in 1962. That the model is not without power is attested by the fact that it accounts for about 26 percent of the variance in educational attainment, 33 percent of the variance in first job, and 42 percent of the variance in 1962 level of occupational attainment. Various additions to the basic model are presented in the volume, but none is clearly shown to make much of an improvement in it. These include nativity, migration, farm origin, subgroup position, marriage, and assortative mating. Without detracting from the excellence of the Blau and Duncan analysis, we may make several observations.

1) Because the dependent behaviors are occupational prestige attainments—attainment levels in a stratification system—it is appropri-

ate to single out variables indicating father's stratification position as the most relevant social structural inputs. It is unfortunate that practical considerations prevented the inclusion of psychological inputs in their model, especially considering the repeated references to one such—mental ability—in the literature on differential occupational attainment (Lipset and Bendix, 1959:203–226; Sewell and Armer, 1966). More recently, this gap has been partially filled (Duncan, 1968).

2) Also omitted are social psychological factors which mediate the influence of the input variables on attainment. This, too, is unfortunate in view not only of the speculative theory but also the concrete research in social psychology, which suggests the importance of such intervening variables as reference groups (Merton, 1957:281–386), significant others (Gerth and Mills, 1953:84–91), self-concept (Super, 1957:80–100), behavior expectations (Gross et al., 1958), levels of educational and occupational aspiration (Haller and Miller, 1963; Kuvlesky and Ohlendorf, 1967; Ohlendorf et al., 1967), and experiences of success or failure in school (Parsons, 1959; Brookover et al., 1965).

It remains to be seen whether the addition of such psychological and social psychological

variables is worthwhile, although there are reasons for believing that at least some of them may be. First, an explanation of a behavior system requires a plausible causal argument, not just a set of path coefficients among temporally ordered variables. As indicated in Duncan's (1969) recent work, the introduction of social psychological mediating variables offers this possibility, but it does not guarantee it. As it stands, the Blau-Duncan model fails to indicate why any connection at all would be expected between the input variables, father's education and occupation, and the three subsequent factors: respondent's education, respondent's first job, and respondent's 1962 occupation. Granting differences among social psychological positions, they all agree that one's cognitions and motivations (including, among others, knowledge, self-concept, and aspirations) are developed in structured situations (including the expectations of others), and that one's actions (attainments in this case) are a result of the cognitive and motivational orientations one brings to the action situation, as well as the factors in the new situation itself. Second, if valid, a social psychological model will suggest new points at which the causal system may be entered in order to change the attainment behaviors of persons, an issue not addressed by the Blau and Duncan volume. Variables such as the expectations of significant others offer other possibilities for manipulating the outcomes, including educational attainments. Third, in addition to the above advantages, a social psychological model of educational and occupational attainment might add to the explanation of variance in the dependent variables.

The Problem

The present report extends the attempts of the writers (Sewell and Armer, 1966; Sewell and Orenstein, 1965; Sewell and Shah, 1967; Sewell, 1964; Haller and Sewell, 1967; Portes et al., 1968; Haller, 1966; Haller and Miller, 1963; Miller and Haller, 1964; Sewell et al., 1957) to apply social psychological

concepts to the explanation of variation in levels of educational and occupational attainment. We assume (1) that certain social structural and psychological factors—initial stratification position and mental ability, specifically—affect both the sets of significant others' influences bearing on the youth, and the youth's own observations of his ability; (2) that the influence of significant others, and possibly his estimates of his ability, affect the youth's levels of educational and occupational aspiration; (3) that the levels of aspiration affect subsequent levels of educational attainment; (4) that education in turn affects levels of occupational attainment. In the present analysis we assume that all effects are linear; also, that the social psychological variables perform only mediating functions.

More specifically, we present theory and data regarding what we believe to be a logically consistent social psychological model. This provides a plausible causal argument to link stratification and mental ability inputs through a set of social psychological and behavioral mechanisms to educational and occupational attainments. One compelling feature of the model is that some of the inputs may be manipulated through experimental or other purposive interventions. This means that parts of it can be experimentally tested in future research and that practical policy agents can reasonably hope to use it in order to change educational and occupational attainments.

A Social Psychological Model

The model treats causal relationships among eight variables. X_1 is the occupational prestige level attained by the adult person, or *occupational attainment* (OccAtt); X_2 is the educational level he had previously attained, or *educational attainment* (EdAtt); X_3 is the occupational prestige level to which he aspired as a youth, or *level of occupational aspiration* (LOA); X_4 is his *level of educational aspiration* as a youth (LEA); X_5 is the influence for educational achievement exerted upon him by

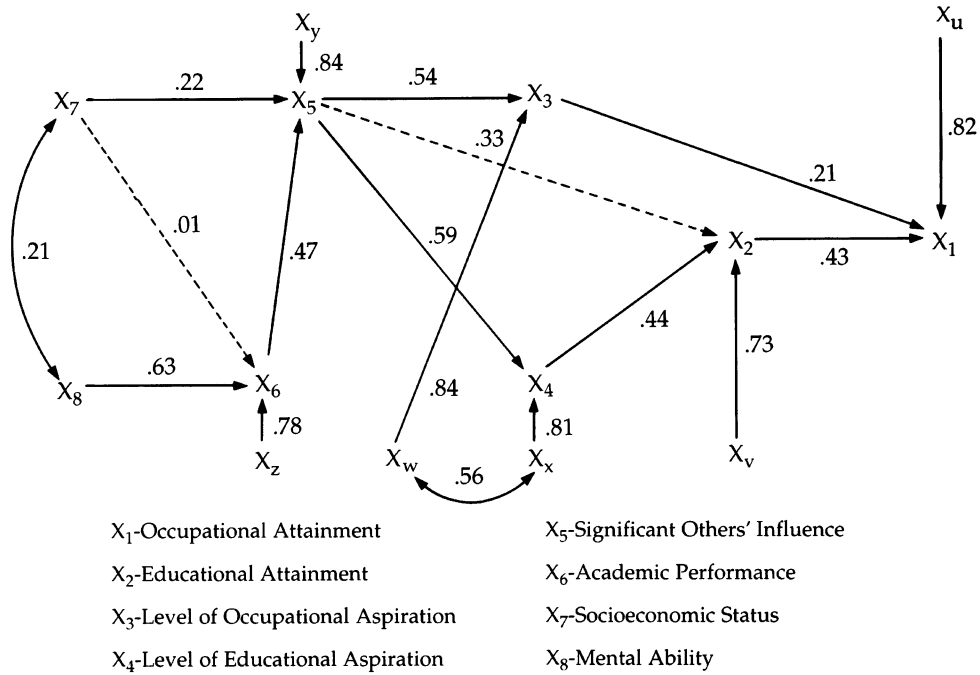


Figure 1. Path Coefficients of Antecedents of Educational and Occupational Attainment Levels

significant others while still in high school, or *significant others' influence* (SOI); X_6 is the quality of his *academic performance* in high school (AP); X_7 is the level of his family in the stratification system, or *socioeconomic status* (SES); and X_8 is his *mental ability* as measured while he was in high school (MA). Path models (Blau and Duncan, 1967:165–172; Wright, 1934; Wright, 1960; Heise, 1969) require a knowledge of the causal order among the variables. Beyond the causal arguments presented below, additional credibility is suggested by the existence of a plausible temporal order among variables. X_7 (SES) and X_8 (MA) precede everything else. X_5 (SOI) and X_6 (AP) precede both aspirations and attainments, and it can be assumed that for the most part X_6 precedes X_5 . Youthful aspirations obviously precede later educational and occupational attainments. Pre-adult educational attainments precede adult occupational attainments.

By no means do all of the possible causal linkages seem defensible. The most likely ones are indicated in Figure 1. In it straight solid

lines stand for causal lines that are to be theoretically expected, dotted lines stand for possible but theoretically debatable causal lines, and curved lines represent unanalyzed correlations among variables which cannot be assigned causal priority in present data.

Commencing from the left of the diagram, we assume, as has often been found before (Sewell and Shah, 1967; Sewell et al., 1957), that a low positive correlation, r_{78} , exists between the youth's measured mental ability (MA) and his parents' socioeconomic status (SES). This is the case: $r_{78} = .21$. We anticipate the existence of substantial effect of MA on academic performance (AP). We theorize that significant others' influence (SOI) is controlled by AP, and by socioeconomic status, as well as by exogenous factors, that they exert profound effects on aspiration, and that the latter in turn influences later attainments. A more detailed examination of the theory follows.

Working with partial conceptions of SOI (and using different terminology), Bordua

(1960) and Sewell and Shah (1968) have shown that parents' expectations for the youths' attainments are important influences on later aspirations and attainment. Similarly, Cramer (1967), Alexander and Campbell (1964), Campbell and Alexander (1965), Haller and Butterworth (1960), and Duncan et al. (1968) have investigated peer influences on aspirations and attainments. Each of these sets of actors, plus some others, may be seen as a special case of reference group influence. Building on such thinking, we have concluded that the key variable here is significant others' influence. Significant others are the specific persons from whom the individual obtains his level of aspiration, either because they serve as models or because they communicate to him their expectations for his behavior (Woelfel, 1967). The term "significant others" is more appropriate than that of "reference group" because it eliminates the implication that collectivities such as one's friends, or work groups, or parents are necessarily the influential agents for all individuals. Experimental research, beginning with Sherif's work (1935), has shown the importance of other persons in defining one's own situation. One obtains his social behavior tendencies largely through the influence of others. Herriott (1963) has carried this line of thinking into the present area of research. He has shown that one's conception of the educational behavior others think appropriate to him is highly correlated with his level of educational aspiration. Thus, significant others' influence is a central variable in a social psychological explanation of educational and occupational attainment. It is obviously important to discover the causal paths determining SOI, as well as those by which it exerts its effects on attainment. We hypothesize a substantial direct path (p_{57}) from socioeconomic status (SES) to SOI. We also hypothesize a substantial effect of mental ability on SOI. This is because we expect that the significant others with whom the youth interacts base their expectations for his

educational and occupational attainments in part on his demonstrated abilities. In turn, this implies that the path from mental ability (MA) to SOI is indirect by way of academic performance (AP). Thus, we hypothesize the existence of a pronounced path from MA to AP (p_{68}) and another from AP to SOI (p_{56}). So far we assume that one's grades in school are based on the quality of his performance. A strong undercurrent in the literature seems to have held, however, that the youth's family's SES has a direct influence on his grades (Havighurst and Neugarten, 1957:236–237). To our knowledge, this has not been adequately demonstrated, and in large high schools, often far removed from the youth's home and neighborhood, this may well be debatable. Nevertheless, since it is at least possible that school grades (the evidences of performance) are partly determined by teachers' desires to please prestigious parents or to reward "middle-class" behavior, we have drawn a dotted path (p_{67}) from SES to AP, allowing for the possibility of such an influence.

We hypothesize that the major effects of significant others' influence (SOI) on attainment are mediated by its effects on levels of aspiration. Thus, we have indicated a path (p_{35}) from SOI to level of occupational aspiration (LOA) and another (p_{45}) from SOI to level of educational aspiration (LEA). It is not inconsistent with this to suspect the possibility that SOI might have a direct influence on later educational attainment (EdAtt); we have thus included a dotted or debatable path (p_{25}) from SOI to EdAtt. Because we are here referring to SOI during late high school, it must necessarily refer largely to college education. There is, therefore, no reason to include such a path from SOI to occupational attainment.

Levels of educational aspiration (LEA) and occupational aspiration (LOA) are known to be highly correlated, since education is widely, and to some extent validly, considered to be a necessary condition for high occupational attainment (Haller and Miller,

1963:30, 39–42, 96). But LOA and LEA are not identical. (In these data, $r_{34.5} = r_{WX} = .56$.) We expect that LEA will have a pronounced effect on EdAtt (p_{24}), and that its entire effect on level of occupational attainment will be expressed through EdAtt. On the other hand, we do not hypothesize any effect of LOA on EdAtt which is not already contained in its correlation with LEA. Hence, there is no hypothetical path for LOA to EdAtt. A direct effect of LOA on OccAtt (p_{13}) is hypothesized, however.

There are 26 possible paths, given the sequence laid out above. As one can see by counting the paths (straight lines) in Figure 1, we hypothesize noteworthy effects for only eight of these—ten if the dotted lines are counted. If this were a rigorous theoretical model, path coefficients would be calculated only for these eight (or ten) supposed causal connections. We believe that because of the fact that it is not rigorous, and at this stage of our knowledge probably cannot be, it would be well to calculate all of the possible 26 path coefficients, using the calculated values as rough indicators of the influences operating in the system. If the theoretic reasoning is a fair description of the reality to which it is addressed, the path coefficients for the eight (or ten) predicted causal lines should be considerably greater than those for the remainder where no causal prediction was made. Also, it is entirely possible that some unhypothesized causal lines might turn out to be of importance. This, too, argues for calculating the whole set of 26. These data are presented in tabular form in Table 3.

Method

In 1957 all high school seniors in Wisconsin responded to an extensive questionnaire concerning their educational and occupational aspirations and a number of potentially related topics. In 1964 one of the authors (Sewell) directed a follow-up in which data on later educational and occupational attain-

ments were collected from an approximately one-third random sample of the respondents in the original survey.

This study is concerned with those 929 subjects for whom data are available at both times, in 1957 and 1964, and who (a) are males and (b) whose fathers were farmers in 1957. Zero-order correlations are computed on all 929 cases, using a computer program which accepts missing data. All higher order coefficients are based on 739 cases for whom data on each variable were complete. (The matrices of zero-order correlations between all eight variables for those two sets of cases are practically identical.)

Variables

Level of occupational attainment (X_1 —OccAtt) was measured by Duncan's (1961) socioeconomic index of occupational status.

Level of educational attainment (X_2 —EdAtt) was operationalized with data obtained in 1964 by dividing the sample into those who have had at least some college education and those who have not had any at all.¹

Level of occupational aspiration (X_3 —LOA) was determined by assigning Duncan's (1961) socioeconomic index scores to the occupation indicated by the respondent as the one he desired to fill in the future.

Level of educational aspiration (X_4 —LEA) is a dichotomous variable corresponding to the respondent's statement in 1957 of whether or not he planned to attend college after graduating from high school.

Index of significant others' influence (X_5 —SOI) is a simple summated score (range: zero to three) of three variables: (a) The youth's report of his parents' encouragement for college, dichotomized according to whether or not the respondent perceived direct parental encouragement for going to college. (b) The youth's report of his teachers' encouragement for college, dichotomized in a similar manner, according to whether or not direct teacher encouragement for college was perceived by the respondent. (c) Friends' college

Table 1. Zero-Order Correlations Between Indicators of Significant Others' Influence Regarding College

| | <i>Teachers' Influence</i> | <i>Friends' Influence</i> | <i>Index of Significant Others' Influence</i> |
|-------------------------------|----------------------------|---------------------------|---|
| Parental Influence | .37 | .26 | .74 |
| Teachers' Influence | -- | .32 | .72 |
| Friends' Influence | -- | -- | .68 |
| Significant Others' Influence | -- | -- | -- |

plans, dichotomized according to the respondent's statement that most of his close friends planned or did not plan to go to college. These variables, all emphasizing education, were combined because they reflect the same conceptual dimension, and that dimension is theoretically more relevant than any of its component parts. That the three components do in fact measure the same dimension is attested by the positive correlations among them and a subsequent factor analysis. These correlations and the correlation of each with the summated variable, significant others' influence, are shown in Table 1. It may be relevant to point out the composition of this significant others' index in the light of Kelley's distinction (1952). Clearly, the perceptions of direct parental and teacher pressures toward college conform to the classic case of normative reference groups. The educational plans of close friends, on the other hand, may be thought of as having mixed functions. First, close peer groups may exercise pressure toward conformity, and second, friends' plans also serve for the individual's cognitive comparison of himself with "people like himself." Therefore, though the main character of the dimension indicated by this index is clearly normative, it can be thought of as containing some elements of an evaluative function as well.

Quality of academic performance (X_6 —AP) is measured by a reflected arc sine transformation of each student's rank in his high school class.

Socioeconomic status (X_7 —SES) is measured by a factor-weighted combination of the education of the respondent's father and

mother, his perception of the economic status of the family, his perception of possible parental support should he choose to go to college and the approximate amount of such support, and the occupation of his father.²

Measured mental ability (X_8 —MA) is indexed by Henmon-Nelson test scores (1942). The data were taken when the youths were in the junior year of high school. The scores, originally recorded as percentile-ranks, were treated with an arc sine transformation to approximate a normal distribution.³

Results

The zero-order correlation coefficients among eight variables are presented in Table 2. A complete path diagram would involve too many lines to be intelligible, because path coefficients presented in Figure 1 were calculated for all 26 possible lines implied in the causal order specified above. With the exception of the theoretically dubious direct path from SES to AP, which turned out to be $p_{67} = .01$, each of the path coefficients for causal lines hypothesized in Figure 1 is larger than those not hypothesized. Both sets of standardized beta (or path) coefficients are presented in Table 3.

This table shows that the reasoning presented in the above section, offering a social psychological explanation for educational and occupational attainment, cannot be too far off the mark. We had hypothesized that SOI (significant others' influence) was of central importance. In fact, it has notable direct effects on three subsequent variables, each of which bears ultimately on prestige

Table 3. Standardized Beta Coefficients for Hypothesized and Non-Hypothesized Causal Paths

| <i>Dependent Variables</i> | <i>Independent Variables</i> | | | | | | |
|----------------------------|------------------------------|-----------------------|-----------------------|-----------------------|----------------------|-----------------------|----------------------|
| | X ₂ EdAtt | X ₃ LOA | X ₄ LEA | X ₅ SOI | X ₆ AP | X ₇ SES | X ₈ MA |
| X ₆ AP | -- | -- | -- | -- | -- | <i>(.01)</i> | .62 |
| X ₅ SOI | -- | -- | -- | -- | .39 | .21 | .13 |
| X ₄ LEA | -- | -- | -- | .45 | .18 | .07 | .08 |
| X ₃ LOA | -- | -- | -- | .42 | .12 | -.02 | .16 |
| X ₂ EdAtt | -- | .07 | .34 | <i>(.23)</i> | .17 | .05 | .03 |
| X ₁ OccAtt | .38 | .19 | -.10 | .11 | .06 | .00 | .04 |

Note: Numbers in italics are coefficients for paths hypothesized in Figure 1. Numbers in parentheses refer to theoretically debatable causal lines.

level of occupational attainment. Both theory and data agree that SOI has direct effects on levels of educational and occupational aspiration, as well as educational (i.e., college) attainment. In turn, each aspiration variable appears to have the predicted substantial effects on its respective attainment variable. Looking at its antecedents, we note theory and data again agree that SOI is affected directly by SES and indirectly by measured mental ability through the latter's effect on the youth's academic performance. The latter variable is crucial because it provides (or is correlated with) palpable evidence that significant others can observe and, thus to a degree, align their expectations for the youth with his demonstrated ability.

None of the unpredicted paths is very strong, but we must recognize that there may be more operating in such a system than we were able to anticipate from previous thinking. There is a pair of perhaps consequential direct paths from academic performance to educational aspiration ($p_{46} = .18$) and to educational attainment ($p_{26} = .17$). There are several possibilities. The data might imply the existence of a mediating factor, such as one's self-conception of his ability, a factor which could influence both educational aspirations and attainment. They also suggest that not all of the effect of ability on educational aspiration and attainments is mediated by SOI. Finally, one's ability may exert a continuing effect on his educational attainments

quite apart from the mediation of either significant others or aspirations—and therefore apart from one's conception of his ability. Arguments such as these, however, should not be pressed too far because the figures are small. Another unexpected but noteworthy path links mental ability directly to level of occupational aspiration. We offer no speculation regarding it.

So far we have seen that a consistent and plausible social psychological position is at least moderately well borne out by the analysis of lines of apparent influence of its variables when they are arranged in causal order. How well does the total set of independent variables work in accounting for variance in the attainment variables? In brief, $R^2_{1,2345678} = .34$ and $R^2_{2,345678} = .50$. Thus, the variables account for 34 percent of the variance in level of occupational attainment and 50 percent of the variance in level of educational attainment. Obviously, variables X₃ through X₈ are much more effective in accounting for educational attainment than in accounting for occupational attainment. Indeed, educational attainment alone accounts for 27 percent of the variance in occupational attainment (from Table 3, $r^2_{12} = .522 = .27$). What we have here, then, is a plausible causal system functioning primarily to explain variation in educational attainment. This, in turn, has considerable effect on occupational attainment. The same set of variables adds a small but useful amount to the explanation

of occupational attainment variance beyond that contributed by its explanation of educational attainment.⁴

Discussion and Conclusions

Using father's occupational prestige, the person's educational attainment, and his first job level, Blau and Duncan (1967:165–172) were able to account for 33 percent of the variance in occupational attainment of a nationwide sample of American men. Neither our sample nor our variables are identical with theirs; so it is impossible to assess the total contribution of this study to the state of knowledge as reflected in their work. Educational attainment is strategic in both studies and in this regard the studies are fairly comparable. The present model adds a great deal to the explanation of the social psychological factors affecting that variable. The prospects seem good, too, that if the present model were to be applied to a sample coming from a wider range of the American stratification system with greater age variation, it might prove to be more powerful than it appears with our sample of young farm-reared men. In general, the present take-off on the Blau-Duncan approach to occupational attainment levels seems worthy of further testing and elaboration.

Several comments are appropriate regarding the social psychological position and data presented here. (1) Clearly, the variable we have called significant others' influence is an important factor. The present evidence appears to show that once formed its effects are far-reaching. Also, besides being a powerful explanatory factor, significant others' influence should be amenable to manipulation. It thus suggests itself as a point at which external agents might intervene to change educational and occupational attainment levels. This means that at least part of the system is theoretically amenable to experimental testing. The parts of the present model which are hypothetically dependent upon this vari-

able might be more securely tested if such experiments can be worked out. Also, practical change agents might be able to change levels of attainment, either by inserting themselves or others as new significant others or by changing the expectations existing significant others have for the individual. There may well be a substantial pay-off from more refined work with this variable.

(2) The results seem to indicate, too, that aspirations (a special class of attitudes) are in fact performing mediational functions in transmitting anterior factors into subsequent behaviors. This has been a subject of recent debate, much of which has in effect held that attitudinal variables are useless epiphenomena. This was recently discussed by Fendrich (1967).

Such encouraging results do not, however, mitigate the need for (a) general experimental determination of the supposed effects of attitudes on behaviors, and (b) specific experimental determination of the effects of aspirations on attainments.

(3) The question may be raised as to the extent to which this system is inherently culture-bound. One might wonder whether attainment behavior within an institutionalized pattern of "sponsored" rather than "contest" achievement (Turner, 1960) would change the path model. Besides this (and perhaps other institutionalized types of achievement patterns), there is also the question of the relevance of the model for ascribed occupational attainment systems. Obviously we do not have data bearing on these questions but we may at least discuss them. Let us suppose that the same eight variables are measured on youth in a "sponsored" achievement context. We speculate that if measured mental ability is the basis of selection of those who are to be advanced, then the direct path from mental ability to significant others' influence would increase because sponsors are significant others. (This would require a more general measure of significant others' influence than was used

here.) If a variable other than mental ability or socioeconomic status is important to the sponsors, then the residual effect of unmeasured variables on significant others' influence would increase. Since one's sponsors presumably influence one's aspirations and aspirations in turn mediate attainment, the rest of the model probably would not change much.

Consider the case of ascribed attainment. Here one's parents' position determines what one's significant others will expect of one; mental ability is either irrelevant or controlled by family position; and one's aspirations are controlled by the family. The importance of higher education may vary among basically ascribed systems: in one it may be unimportant, in another it may merely validate one's status, or in still another it may train ascribed elites to fulfill the key social roles in the society. If educational attainment is important within the social system, aspirations will mediate the influence of significant others upon it, and it in turn will mediate occupational attainment. If not, occupational aspirations will mediate occupational attainment and educational attainment will drop out of the path model. In short, by allowing for variations in the path coefficients, the same basic social psychological model might work well to describe attainment in stratification and mobility systems quite different from that of the present sample.

(4) The linear model used here seems to be an appropriate way to operationalize social psychological positions holding that the function of "intervening" attitudinal variables is to mediate the influence of more fundamental social structural and psychological variables on behavior. By assuming linear relations among variables and applying a path system to the analysis, we have cast the attainment problem in such a framework. It seems to have worked quite well. We are sufficiently encouraged by this attempt to recommend that a parallel tack might be made

on problems in which the overt behavior variables are quite different from educational and occupational attainment.

(5) Nonetheless, satisfactory as such a linear model and its accompanying theory seems to be, there is still the possibility that other techniques flowing from somewhat different social psychological assumptions might be better. It is possible that, in the action situation, enduring attitudes (such as educational and occupational aspirations) may function as independent forces which express themselves in relevant overt behaviors to the degree that other personality and situational variables permit. Linear models would thus be effective to the degree that the persons modify their aspirations to bring them in line with potentials for action offered by the latter variables. More importantly, the combined effects of aspirational and facilitational variables would produce nonlinear accelerating curves of influence on behavior variables. For the present types of data, this would imply that parental stratification position, mental ability, and significant others' influence not only produce aspirations, but also, to the extent to which these influences continue more or less unchanged on into early adulthood, they function as differential facilitators for the expression of aspirations in attainments. If this is true, a nonlinear system of statistical analysis handling interaction effects would be even more powerful than the one used in this paper.

(6) It should be remembered that the most highly educated of these young men had just begun their careers when the final data were collected. If the distance between them and the less educated widens, the occupational attainment variance accounted for by the model may well increase. The direct relations of some of the antecedents to occupational attainment may also change. In particular, mental ability may show a higher path to occupational attainment.

(7) Finally, although the results reported in this paper indicate that the proposed model

has considerable promise for explaining educational and early occupational attainment of farm boys, its adequacy should now be tested on populations with a more differentiated socioeconomic background. It is quite possible that in such populations the effects of socioeconomic status on subsequent variables may be significantly increased. The effects of other variables in the system may also be altered when the model is applied to less homogeneous populations.

The present research appears to have extended knowledge of the causal mechanism influencing occupational attainment. Most of this was accomplished by providing a consistent social psychological model which adds to our ability to explain what is surely one of its key proximal antecedents, educational attainment.

NOTES

Revision of paper originally prepared for delivery at the joint sessions of the Rural Sociological Society and the American Sociological Association, San Francisco, August 1967. The research reported here was supported by the University of Wisconsin Graduate School, by the Cooperative State Research Service, and the University's College of Agriculture for North Central Regional Research Committee NC-86, by funds to the Institute for Research on Poverty at the University of Wisconsin provided by the Office of Economic Opportunity pursuant to the provisions of the Economic Opportunity Act of 1964, and by a grant from the National Institute of Health, U. S. Public Health Service (M-6275). The writers wish to thank Otis Dudley Duncan for his careful reading and incisive criticisms and Vimal P. Shah for help in the statistical analysis. The conclusions are the full responsibility of the authors.

1. It is important to note that the timing of the follow-up was such as to allow most individuals to complete their education up to the bachelor's degree and beyond. It is unlikely that the educational attainment of the sample as a whole will change much in the years to come. On the other hand, while the span of seven years allowed those individuals who did not continue their education to find a stable position in the occupational structure and even im-

prove upon it, there was not enough time for those who continued their education to do the same. A few of the latter were still in school; most had just begun their occupational careers. It is therefore possible that a follow-up taken five or ten years from now would show greater differentiation in attainments as the educated group gathers momentum and moves up in the occupational world.

2. Naturally, father's occupation is a constant in this subsample of farm-reared males. It is important to note that the SES mean and standard deviations for this subsample are considerably lower than for the total sample. The low and homogeneous SES levels of this subsample may yield atypical relations among the variables.

3. Our previous research (Sewell and Armer, 1966; Haller and Sewell, 1967) has led us to be skeptical of claims that local ecological and school class compositional factors influence aspirations and attainments. Nevertheless the zero-order intercorrelations of five such variables and their correlations with X_1 - X_8 are available (although they are not presented here). Two of these pertain to the county in which the youth attended high school: county level of living and degree of urbanization. Three pertain to his high school senior class: average SES of the class, percentage of the class members whose fathers attended college, and percentage of the class members whose fathers had professional-level occupations. Though substantially correlated with each other, the variables are uncorrelated with the variables in the above model.

4. Some readers will be interested in the path coefficients as calculated only for the lines hypothesized in the diagram. For this reason and because of the diagram's parsimony, we have calculated the values for each of its eight paths (or ten, including dubious ones). The restricted model explains 47 and 33 percent of the variance in X_2 and X_1 , respectively. Data not presented here show that the model reproduces the zero-order correlation matrix quite well. For this reason and because the model is an effective predictor of X_2 and X_1 , it may be considered to be fairly valid. Nonetheless, it seems more prudent to rest our case on the less presumptuous data already presented in Table 3. This is why the coefficients presented in the diagram are not discussed here.

REFERENCES

- Alexander, C. Norman, Jr., and Ernest Q. Campbell. 1964. "Peer influences on adolescent educational aspirations and attainments." *American Sociological Review* 29 (August):568-575.

- Blau, Peter M., and Otis Dudley Duncan. 1967. *The American Occupational Structure*. New York: Wiley.
- Bordua, David J. 1960. "Educational aspirations and parental stress on college." *Social Forces* 38 (March):262–269.
- Brookover, Wilbur B., Jean M. LePere, Don E. Hamachek, Shailer Thomas, and Edsel L. Erickson. 1965. *Self-Concept of Ability and School Achievement*. East Lansing: Michigan State University, Bureau of Educational Research Services.
- Campbell, Ernest Q., and C. Norman Alexander. 1965. "Structural effects and interpersonal relationships." *American Journal of Sociology* 71 (November):284–289.
- Cramer, M. R. 1967. "The relationship between educational and occupational plans of high school students." Paper presented at the meeting of the Southern Sociological Society, Atlanta (unpublished).
- Duncan, Otis Dudley. 1961. "A socioeconomic index for all occupations." Pp. 109–138 in Albert J. Reiss, Jr. (ed.), *Occupations and Social Status*. New York: Free Press.
- _____. 1968. "Ability and achievement." *Eugenics Quarterly* 15 (March):1–11.
- _____. 1969. "Contingencies in the construction of causal models." Edgar F. Borgatta, (ed.), *Sociological Methodology*. San Francisco: Jossey-Bass.
- Duncan, Otis Dudley, Archibald O. Haller, and Alejandro Portes. 1968. "Peer influences on aspirations: a reinterpretation." *American Journal of Sociology* 74 (September):119–137.
- Fendrich, James M. 1967. "Perceived reference group support: racial attitudes and overt behavior." *American Sociological Review* 32 (December):960–970.
- Gerth, Hans, and C. Wright Mills. 1953. *Character and Social Structure*. New York: Harcourt, Brace and World.
- Gross, Neal, Ward S. Mason, and Alexander W. McEachern. 1958. *Explorations in Role Analysis*. New York: Wiley.
- Haller, Archibald O. 1966. "Occupational choices of rural youth." *Journal of Cooperative Extension* 4 (Summer):93–102.
- Haller, Archibald O., and Charles E. Butterworth. 1960. "Peer influences on levels of occupational and educational aspiration." *Social Forces* 38 (May):289–295.
- Haller, Archibald O., and Irwin W. Miller. 1963. *The Occupational Aspiration Scale: Theory, Structure and Correlates*. East Lansing: Michigan Agricultural Experiment Station Bulletin 288.
- Haller, Archibald O., and William H. Sewell. 1967. "Occupational choices of Wisconsin farm boys." *Rural Sociology* 32 (March):37–55.
- Havighurst, Robert J., and Bernice L. Neugarten. 1957. *Society and Education*. Boston: Allyn and Bacon.
- Heise, David R. 1969. "Problems in path analysis and causal inference." Edgar F. Borgatta, (ed.), *Sociological Methodology*. San Francisco: Jossey-Bass.
- Henmon, V. A. C., and M. J. Nelson. 1942. *The Henmon-Nelson Test of Mental Ability*. Boston: Houghton Mifflin Company.
- Herriott, Robert E. 1963. "Some social determinants of educational aspiration." *Harvard Educational Review* 33 (Spring):157–177.
- Kelley, Harold H. 1952. "Two functions of reference groups." Pp. 410–414 in Guy E. Swanson *et al.* (eds.), *Readings in Social Psychology*. New York: Henry Holt and Company.
- Kuvlesky, William P., and George W. Ohlendorf. 1967. *A Bibliography of Literature on Status Projections of Youth: I. Occupational Aspirations and Expectations*. College Station: Texas A&M University, Department of Agricultural Economics and Sociology.
- Lipset, Seymour M., and Reinhard Bendix. 1959. *Social Mobility in Industrial Society*. Berkeley: University of California Press.
- Merton, Robert K. 1957. *Social Theory and Social Structure*. New York: Free Press.
- Miller, I. W., and Archibald O. Haller. 1964. "The measurement of level of occupational aspiration." *Personnel and Guidance Journal* 42 (January):448–455.
- Ohlendorf, George W., Sherry Wages, and William P. Kuvlesky. 1967. *A Bibliography of Literature on Status Projections of Youth: II. Educational Aspirations and Expectations*. College Station: Texas A & M University, Department of Agricultural Economics and Sociology.
- Parsons, Talcott. 1959. "The school class as a social system." *Harvard Educational Review* 29 (Summer):297–318.
- Portes, Alejandro, Archibald O. Haller, and William H. Sewell. 1968. "Professional executive vs. farming as unique occupational choices." *Rural Sociology* 33 (June):153–159.
- Sewell, William H. 1964. "Community of residence and college plans." *American Sociological Review* 29 (February):24–38.

- Sewell, William H., and J. Michael Armer. 1966. "Neighborhood context and college plans." *American Sociological Review* 31 (April): 159–168.
- Sewell, William H., Archibald O. Haller, and Murray A. Straus. 1957. "Social status and educational and occupational aspiration." *American Sociological Review* 22 (February):67–73.
- Sewell, William H., and Alan M. Orenstein. 1965. "Community of residence and occupational choice." *American Journal of Sociology* 70 (March):551–563.
- Sewell, William H., and Vimal P. Shah. 1967. "Socioeconomic status, intelligence, and the attainment of higher education." *Sociology of Education* 40 (Winter):1–23.
- _____. 1968. "Social class, parental encouragement, and educational aspirations." *American Journal of Sociology* 73 (March):559–572.
- Sherif, Muzafer. 1935. "A study of some social factors in perception." *Archives of Psychology* Number 187.
- Super, Donald E. 1957. *The Psychology of Careers*. New York: Harper.
- Turner, Ralph H. 1960. "Sponsored and contest mobility and the school system." *American Sociological Review* 25 (December):855–867.
- Woelfel, Joseph. 1967. "A paradigm for research on significant others." Paper presented at the Joint Session of the Society for the Study of Social Problems and the American Sociological Association, San Francisco (unpublished).
- Wright, Sewall. 1934. "The method of path coefficients." *Annals of Mathematical Statistics* 5 (September):161–215.
- _____. 1960. "Path coefficients and regression coefficients: alternative or complementary concept?" *Biometrics* 16 (June):189–202.

54 ■ Jay MacLeod

Ain't No Makin' It

Leveled Aspirations in a Low-Income Neighborhood

"Any child can grow up to be president." So maintains the dominant ideology in the United States. This perspective characterizes American society as an open one in which barriers to success are mainly personal rather than social. In this meritocratic view, education ensures equality of opportunity for all individuals, and economic inequalities result from differences in natural qualities and in one's motivation and will to work. Success is based on achievement rather than ascription. Individuals do not inherit their social sta-

tus—they attain it on their own. Because schooling mitigates gender, class, and racial barriers to success, the ladder of social mobility is there for all to climb. A favorite Hollywood theme, the rags-to-riches story resonates in the psyche of the American people. We never tire of hearing about Andrew Carnegie, for his experience validates much that we hold dear about America, the land of opportunity. Horatio Alger's accounts of the spectacular mobility achieved by men of humble origins through their own unremit-

Jay MacLeod. *Ain't No Makin' It: Leveled Aspirations in a Low-Income Neighborhood* (Boulder: Westview Press, 1987), pp. 1–2, 4–5, 8, 60–63, 69–75, 78–79, 81, 137–141, 162. Copyright © 1987 by Westview Press, Inc. Reprinted by permission of Westview Press, Inc., a member of the Perseus Books Group. The postscript to this selection is taken from *Ain't No Makin' It: Leveled Aspirations in a Low-Income Neighborhood* (Boulder: Westview Press, 1995), pp. 155, 169–170, 196, 240–241. Copyright © 1995 by Westview Press, Inc. Reprinted by permission of Westview Press, Inc., a member of the Perseus Books Group.

ting efforts occupy a treasured place in our national folklore. The American Dream is held out as a genuine prospect for anyone with the drive to achieve it.

"I ain't goin' to college. Who wants to go to college? I'd just end up gettin' a shitty job anyway." So says Freddie Piniella,¹ an intelligent eleven-year-old boy from Clarendon Heights, a low-income housing development in a northeastern city. This statement, pronounced with certitude and feeling, completely contradicts our achievement ideology. Freddie is pessimistic about his prospects for social mobility and disputes schooling's capacity to "deliver the goods." Such a view offends our sensibilities and seems a rationalization. But Freddie has a point. What of Carnegie's grammar school classmates, the great bulk of whom no doubt were left behind to occupy positions in the class structure not much different from those held by their parents? What about the static, nearly permanent element in the working class, whose members consider the chances for mobility remote and thus despair of all hope? These people are shunned, hidden, forgotten—and for good reason—because just as the self-made man is a testament to certain American ideals, so the very existence of an "underclass" in American society is a living contradiction to those ideals.

Utter hopelessness is the most striking aspect of Freddie's outlook. Erik H. Erikson writes that hope is the basic ingredient of all vitality;² stripped of hope, there is little left to lose. How is it that in contemporary America a boy of eleven can feel bereft of a future worth embracing? This is not what the United States is supposed to be. The United States is the nation of hopes and dreams and opportunity. As Ronald Reagan remarked in his 1985 State of the Union Address, citing the accomplishments of a young Vietnamese immigrant, "Anything is possible in America if we have the faith, the will, and the heart."³ But to Freddie Piniella and many other Clarendon Heights young people who grow

up in households where their parents and older siblings are unemployed, undereducated, or imprisoned, Reagan's words ring hollow. For them the American Dream, far from being a genuine prospect, is not even a dream. It is a hallucination.

I first met Freddie Piniella in the summer of 1981 when as a student at a nearby university I worked as a counselor in a youth enrichment program in Clarendon Heights. For ten weeks I lived a few blocks from the housing project and worked intensively with nine boys, aged eleven to thirteen. While engaging them in recreational and educational activities, I was surprised by the modesty of their aspirations. The world of middle-class work was entirely alien to them; they spoke about employment in construction, factories, the armed forces, or, predictably, professional athletics. In an ostensibly open society, they were a group of boys whose occupational aspirations did not even cut across class lines. . . .

The male teenage world of Clarendon Heights is populated by two divergent peer groups. The first group, dubbed the Hallway Hangers because of the group's propensity for "hanging" in a particular hallway in the project [i.e., outside doorway #13], consists predominantly of white boys. Their characteristics and attitudes stand in marked contrast to the second group, which is composed almost exclusively of black youths who call themselves the Brothers. Surprisingly, the Brothers speak with relative optimism about their futures, while the Hallway Hangers are despondent about their prospects for social mobility. . . .

Before describing the boys' orientation toward work [in more detail], I would like to make an analytical distinction between aspirations and expectations. Both involve assessments of one's desires, abilities, and the character of the opportunity structure. In articulating one's aspirations, an individual weighs his or her preferences more heavily; expectations are tempered by perceived capabilities and available opportunities. Aspirations

are one's preferences relatively unsullied by anticipated constraints; expectations take these constraints squarely into account.⁴

The Hallway Hangers: Keeping a Lid on Hope

Conventional, middle-class orientations toward employment are inadequate to describe the Hallway Hangers' approach to work. The notion of a career, a set of jobs that are connected to one another in a logical progression, has little relevance to these boys. They are hesitant when asked about their aspirations and expectations. This hesitancy is not the result of indecision; rather it stems from the fact that these boys see little choice involved in getting a job. No matter how hard I pressed him, for instance, Jinks refused to articulate his aspirations: "I think you're kiddin' yourself to have any. We're just gonna take whatever we can get." Jinks is a perceptive boy, and his answer seems to be an accurate depiction of the situation. Beggars cannot be choosers, and these boys have nothing other than unskilled labor to offer on a credential-based job market.

It is difficult to gauge the aspirations of most of the Hallway Hangers. Perhaps at a younger age they had dreams for their futures. At ages sixteen, seventeen, and eighteen, however, their own job experiences as well as those of family members have contributed to a deeply entrenched cynicism about their futures. What is perceived as the cold, hard reality of the job market weighs very heavily on the Hallway Hangers; they believe their preferences will have almost no bearing on the work they actually will do. Their expectations are not merely tempered by perceptions of the opportunity structure; even their aspirations are crushed by their estimation of the job market. These generalizations may seem bold and rather extreme, but they do not lack ethnographic support.

The pessimism and uncertainty with which the Hallway Hangers view their fu-

tures emerge clearly when the boys are asked to speculate on what their lives will be like in twenty years.

(all in separate interviews)

STONE: Hard to say. I could be dead tomorrow. Around here, you gotta take life day by day.

BOO-BOO: I dunno. I don't want to think about it. I'll think about it when it comes.

FRANKIE: I don't fucking know. Twenty years. I may be fucking dead. I live a day at a time. I'll probably be in the fucking pen.

SHORTY: Twenty years? I'm gonna be in jail.

These responses are striking not only for the insecurity and despondency they reveal, but also because they do not include any mention of work. It is not that work is unimportant—for people as strapped for money as the Hallway Hangers are, work is crucial. Rather, these boys are indifferent to the issue of future employment. Work is a given; they all hope to hold jobs of one kind or another in order to support themselves and their families. But the Hallway Hangers believe the character of work, at least all work in which they are likely to be involved, is essentially the same: boring, undifferentiated, and unrewarding. Thinking about their future jobs is a useless activity for the Hallway Hangers. What is there to think about?

For Steve and Jinks, although they do see themselves employed in twenty years, work is still of tangential importance.

JM: If you had to guess, what do you think you'll be doing twenty years from now?

(in separate interviews)

STEVE: I don't fucking know. Working probably. Have my own pad, my own house. Bitches, kids. Fucking fridge full of brewskies. Fine wife, likes to get laid.

JINKS: Twenty years from now? Probably kicked back in my own apartment doing

the same shit I'm doing now—getting high. I'll have a job, if I'm not in the service, if war don't break out, if I'm not dead. I just take one day at a time.

Although the Hallway Hangers expect to spend a good portion of their waking hours on the job, work is important to them not as an end in itself, but solely as a means to an end—money.

In probing the occupational aspirations and expectations of the Hallway Hangers, I finally was able to elicit from them some specific hopes. Although Shorty never mentions his expectations, the rest of the Hallway Hangers have responded to my prodding with some definite answers. The range of answers as well as how they change over time are as significant as the particular hopes each boy expresses.

Boo-Boo's orientation toward work is typical of the Hallway Hangers. He has held a number of jobs in the past, most of them in the summer. During his freshman year in high school Boo-Boo worked as a security guard at school for \$2.50 an hour in order to make restitution for a stolen car he damaged. Boo-Boo also has worked on small-scale construction projects through a summer youth employment program called Just-A-Start, at a pipe manufacturing site, and as a clerk in a gift shop. Boo-Boo wants to be an automobile mechanic. Upon graduating from high school, he studied auto mechanics at a technical school on a scholarship. The only black student in his class, Boo-Boo was expelled early in his first term after racial antagonism erupted into a fight. Boo-Boo was not altogether disappointed, for he already was unhappy with what he considered the program's overly theoretical orientation. (Howard London found this kind of impatience typical of working-class students in the community college he studied.⁵) Boo-Boo wanted hands-on training, but "all's they were doing was telling me about how it's made, stuff like that." Boo-Boo currently is unemployed, but

he recently had a chance for a job as a cook's helper. Although he was not hired, the event is significant nevertheless because prior to the job interview, Boo-Boo claimed that his ambition now was to work in a restaurant. Here we have an example of the primacy of the opportunity structure in determining the aspirations of the Hallway Hangers. One job opening in another field was so significant that the opening prompted Boo-Boo to redefine totally his aspirations.

In contrast to the rest of the Hallway Hangers who are already on the job market, Steve wants to stay in school for the two years required to get his diploma. Yet he has a similar attitude toward his future work as do the other youths. He quit his summer job with the Just-A-Start program and has no concrete occupational aspirations. As for expectations, he believes he might enlist in the Air Force after graduation but adds, "I dunno. I might just go up and see my uncle, do some fuckin' construction or something."

Many of these boys expect to enter military service. Jinks and Frankie mention it as an option; Stoney has tried to enlist, but without success. Although Jinks refuses to think in terms of aspirations, he will say what he expects to do after he finishes school.

JM: What are you gonna do when you get out?

JINKS: Go into the service, like everybody else. The navy.

JM: What about after that?

JINKS: After that, just get a job, live around here.

JM: Do you have any idea what job you wanna get?

JINKS: No. No particular job. Whatever I can get.

Jinks subsequently quit school. He had been working twenty hours a week making clothes-racks in a factory with his brother. He left school with the understanding that he would be employed full-time, and he was

mildly content with his situation: “I got a job. It ain’t a good job, but other things will come along.” Two weeks later, he was laid off. For the past three months he has been unemployed, hanging full-time in doorway #13.

Shorty has worked construction in the past and has held odd jobs such as shoveling snow. Shorty, an alcoholic, has trouble holding down a steady job, as he freely admits. He was enrolled in school until recently. Ordered by the court to a detoxification center, Shorty apparently managed to convince the judge that he had attended enough Alcoholics Anonymous meetings in the meantime to satisfy the court. He has not returned to school since, nor has he landed a job. Given that Shorty is often on the run from the police, he is too preoccupied with pressing everyday problems to give serious thought to his long-term future. It is not surprising that my ill-timed query about his occupational aspirations met with only an impatient glare. . . .

The definitions of aspirations and expectations given [earlier] suggest that an assessment of the opportunity structure and of one’s capabilities impinge on one’s preferences for the future. However, the portrait of the Hallway Hangers painted in these pages makes clear that “impinge” is not a strong enough word. But are the leveled aspirations and pessimistic expectations of the Hallway Hangers a result of strong negative assessments of their capabilities or of the opportunity structure?

This is not an easy question to answer. Doubtless, both factors come into play, but in the case of the Hallway Hangers, evaluation of the opportunity structure has the dominant role. Although in a discussion of why they do not succeed in school, the Hallway Hangers point to personal inadequacy (“We’re all just fucking burnouts”; “We never did good anyways”), they look to outside forces as well. In general, they are confident of their own abilities.

(In a group interview)

JM: If you’ve got five kids up the high school with all A’s, now are you gonna be able to say that any of them are smarter than any of you?

SLICK: *(immediately)* No.

JM: So how’d that happen?

SLICK: Because they’re smarter in some areas just like we’re smarter in some areas. You put them out here, right? And you put us up where they’re living—they won’t be able to survive out here.

SHORTY: But we’d be able to survive up there.

FRANKIE: See, what it is—they’re smarter more academically because they’re taught by teachers that teach academics.

JM: Not even streetwise, just academically, do you think you could be up where they are?

FRANKIE: Yeah.

CHRIS: Yeah.

SHORTY: Yeah.

JM: When it comes down to it, you’re just as smart?

FRANKIE: Yeah.

SLICK: *(matter-of-factly)* We could be smarter.

FRANKIE: Definitely.

CHRIS: On the street, like.

FRANKIE: We’re smart, we’re smart, but we’re just smart [inaudible]. It’s fucking, y’know, we’re just out to make money, man. I know if I ever went to fucking high school and college in a business course. . . .

SLICK: And concentrated on studying. . . .

FRANKIE: I know I could make it. I am a businessman.

JM: So all of you are sure that if you put out in school. . . .

FRANKIE: Yeah! If I went into business, I would, yeah. If I had the fucking money to start out with like some of these fucking rich kids, I’d be a millionaire. Fucking right I would be.

Although these comments were influenced by the dynamics of the group interview, they jibe with the general sense of self-confidence the Hallway Hangers radiate and indicate that they do not have low perceptions of their own abilities.

If their assessments of their own abilities do not account for the low aspirations of the Hallway Hangers, we are left, by way of explanation, with their perceptions of the job opportunity structure. The dominant view in the United States is that American society is an open one that values and differentially rewards individuals on the basis of their merits. The Hallway Hangers question this view, for it runs against the grain of their neighbors' experiences, their families' experiences, and their own encounters with the labor market.

The Clarendon Heights community, as a public housing development, is by definition made up of individuals who do not hold even modestly remunerative jobs. A large majority are on additional forms of public assistance; many are unemployed. Like most old housing projects, Clarendon Heights tends to be a cloistered, insular neighborhood, isolated from the surrounding community. Although younger residents certainly have external points of reference, their horizons are nevertheless very narrow. Their immediate world is composed almost entirely of people who have not "made it." To look around at a great variety of people—some lazy, some alcoholics, some energetic, some dedicated, some clever, some resourceful—and to realize all of them have been unsuccessful on the job market is powerful testimony against what is billed as an open society.

The second and much more intimate contact these boys have with the job market is through their families, whose occupational histories only can be viewed as sad and disillusioning by the Hallway Hangers. These are not people who are slothful or slow-witted; rather, they are generally industrious, intelligent, and very willing to work. With mem-

bers of their families holding low-paying, unstable jobs or unable to find work at all, the Hallway Hangers are unlikely to view the job opportunity structure as an open one.

The third level of experience on which the Hallway Hangers draw is their own. These boys are not newcomers to the job market. As we have seen, all have held a variety of jobs. All except Steve are now on the job market year round, but only Stoney has a steady job. With the exceptions of Chris, who presently is satisfied with his success peddling drugs, and Steve, who is still in school, the Hallway Hangers are actively in search of decent work. Although they always seem to be following up on some promising lead, they are all unemployed. Furthermore, some who were counting on prospective employment have had their hopes dashed when it fell through. The work they have been able to secure typically has been in menial, dead-end jobs paying minimum wage.

Thus, their personal experience on the job market and the experiences of their family members and their neighbors have taught the Hallway Hangers that the job market does not necessarily reward talent or effort. Neither they nor their parents, older siblings, and friends have shared in the "spoils" of economic success. In short, the Hallway Hangers are under no illusions about the openness of the job opportunity structure. They are conscious, albeit vaguely, of a number of class-based obstacles to economic and social advancement. Slick, the most perceptive and articulate of the Hallway Hangers, points out particular barriers they must face.

SLICK: Out here, there's not the opportunity to make money. That's how you get into stealin' and all that shit.

(in a separate interview)

SLICK: That's why I went into the army—cuz there's no jobs out here right now for people that, y'know, live out here. You have to know somebody, right?

In discussing the problems of getting a job, both Slick and Shorty are vocal.

SLICK: All right, to get a job, first of all, this is a handicap, out here. If you say you're from the projects or anywhere in this area, that can hurt you. Right off the bat: reputation.

SHORTY: Is this dude gonna rip me off, is he. . . .

SLICK: Is he gonna stab me?

SHORTY: Will he rip me off? Is he gonna set up the place to do a score or something? I tried to get a couple of my buddies jobs at a place where I was working construction, but the guy says, "I don't want 'em if they're from there. I know you; you ain't a thief or nothing."

Frankie also points out the reservations prospective employers have about hiring people who live in Clarendon Heights. "A rich kid would have a better chance of getting a job than me, yeah. Me, from where I live, y'know, a high crime area, I was prob'ly crime-breaking myself, which they think your nice honest rich kid from a very respected family would never do."

Frankie also feels that he is discriminated against because of the reputation that attaches to him because of his brothers' illegal exploits. "Especially me, like I've had a few opportunities for a job, y'know. I didn't get it cuz of my name, because of my brothers, y'know. So I was deprived right there, bang. Y'know they said, 'No, no, no, we ain't havin' no Dougherty work for us.'" In a separate discussion, Frankie again makes this point. Arguing that he would have almost no chance to be hired as a fireman, despite ostensibly meritocratic hiring procedures, even if he scored very highly on the test, Frankie concludes, "Just cuz fuckin' where I'm from and what my name is."

The Hallway Hangers' belief that the opportunity structure is not open also emerges when we consider their responses to the

question of whether they have the same chance as a middle- or upper-class boy to get a good job. The Hallway Hangers generally respond in the negative. When pushed to explain why, Jinks and Steve made these responses, which are typical.

(in separate interviews)

JINKS: Their parents got pull and shit.

STEVE: Their fucking parents know people.

Considering the boys' employment experiences and those of their families, it is not surprising that the Hallway Hangers' view of the job market does not conform to the dominant belief in the openness of the opportunity structure. They see a job market where rewards are based not on meritocratic criteria, but on "who you know." If "connections" are the keys to success, the Hallway Hangers know that they are in trouble.

Aside from their assessment of the job opportunity structure, the Hallway Hangers are aware of other forces weighing on their futures. A general feeling of despondency pervades the group. As Slick puts it, "The younger kids have nothing to hope for." The Hallway Hangers often draw attention to specific incidents that support their general and vague feelings of hopelessness and of the futility of nurturing aspirations or high expectations. Tales of police brutality, of uncaring probation officers and callous judges, and of the "pull and hook-ups of the rich kids" all have a common theme, which Chris summarizes, "We don't get a fair shake and shit." Although they sometimes internalize the blame for their plight (Boo-Boo: "I just screwed up"; Chris: "I guess I just don't have what it takes"; Frankie: "We've just fucked up"), the Hallway Hangers also see, albeit in a vague and imprecise manner, a number of hurdles in their path to success with which others from higher social strata do not have to contend.

Insofar as contemporary conditions under capitalism can be conceptualized as a race by

the many for relatively few positions of wealth and prestige, the low aspirations of the Hallway Hangers, more than anything else, seem to be a decision, conscious or unconscious, to withdraw from the running. The competition, they reason, is not a fair one when some people have an unobstructed lane. As Frankie maintains, the Hallway Hangers face numerous barriers: “It’s a steeplechase, man. It’s a motherfucking steeplechase.” The Hallway Hangers respond in a way that suggests only a “sucker” would compete seriously under such conditions.

Chris’s perspective seems a poignant, accurate description of the situation in which the Hallway Hangers find themselves.

CHRIS: I gotta get a job, any fucking job. Just a job. Make some decent money. If I could make a hundred bucks a week, I’d work. I just wanna get my mother out of the projects, that’s all. But I’m fucking up in school. It ain’t easy, Jay. I hang out there [in doorway #13] ’til about one o’clock every night. I never want to go to school. I’d much rather hang out and get high again. It’s not that I’m dumb. You gimme thirty bucks today, and I’ll give you one hundred tomorrow. I dunno. It’s like I’m in a hole I can’t get out of. I guess I could get out, but it’s hard as hell. It’s fucked up.

The Brothers: Ready at the Starting Line

Just as the pessimism and uncertainty with which the Hallway Hangers view their futures emerges when we consider what they perceive their lives will be like in twenty years, so do the Brothers’ long-term visions serve as a valuable backdrop to our discussion of their aspirations. The ethos of the Brothers’ peer group is a positive one; they are not resigned to a bleak future but are hoping for a bright one. Nowhere does this optimism surface more clearly than in the Brothers’ re-

sponses to the question of what they will be doing in twenty years. Note the centrality of work in their views of the future.

(all in separate interviews)

SUPER: I’ll have a house, a nice car, no one bothering me. Won’t have to take no hard time from no one. Yeah, I’ll have a good job, too.

JUAN: I’ll have a regular house, y’know, with a yard and everything. I’ll have a steady job, a good job. I’ll be living the good life, the easy life.

MIKE: I might have a wife, some kids. I might be holding down a regular business job like an old guy. I hope I’ll be able to do a lot of skiing and stuff like that when I’m old.

CRAIG: I’ll probably be having a good job on my hands, I think. Working in an office as an architect, y’know, with my own drawing board, doing my own stuff, or at least close to there.

James takes a comic look into his future without being prompted to do so. “The ones who work hard in school, eventually it’s gonna pay off for them and everything, and they’re gonna have a good job and a family and all that. Not me though! I’m gonna have *myself*. I’m gonna have some money. And a different girl every day. And a different car. And be like this (*poses with one arm around an imaginary girl and the other on a steering wheel*).”

The Brothers do not hesitate to name their occupational goals. Although some of the Brothers are unsure of their occupational aspirations, none seems to feel that nurturing an aspiration is a futile exercise. The Brothers have not resigned themselves to taking whatever they can get. Rather, they articulate specific occupational aspirations (although these often are subject to change and revision).

Like all the Brothers, Super has not had extensive experience on the job market; he only has worked at summer jobs. For the past

three summers, he has worked for the city doing maintenance work in parks and school buildings through a CETA-sponsored summer youth employment program. During the last year, Super's occupational aspirations have fluctuated widely. His initial desire to become a doctor was met with laughter from his friends. Deterred by their mocking and by a realization of the schooling required to be a doctor, Super immediately decided that he would rather go into business: "Maybe I can own my own shop and shit." This aspiration, however, also was ridiculed. "Yeah, right," commented Mokey, "Super'll be pimping the girls, that kinda business." In private, however, Super still clings to the hope of becoming a doctor, although he cites work in the computer field as a more realistic hope. "Really, I don't know what I should do now. I'm kinda confused. First I said I wanna go into computers, right? Take up that or a doctor." The vagueness of Super's aspirations is important; once again, we get a glimpse of how little is known about the world of middle-class work, even for somebody who clearly aspires to it. Of one thing Super is certain: "I just know I wanna get a good job."

Although Super does not distinguish between what constitutes a good job and what does not, he does allude to criteria by which the quality of a job can be judged. First, a good job must not demand that one "work on your feet," a distinction, apparently, between white and blue-collar work. Second, a good job implies at least some authority in one's workplace, a point Super makes clearly, if in a disjointed manner. "Bosses—if you don't come on time, they yell at you and stuff like that. They want you to do work and not sit down and relax and stuff like that, y'know. I want to try and be a boss, y'know, tell people what to do. See, I don't always want people telling me what to do, y'know—the low rank. I wanna try to be with people in the high rank." Although Super does not know what occupation he would like to enter, he is certain that he wants a job that is

relatively high up in a vaguely defined occupational hierarchy. . . .

The Brothers display none of the cockiness about their own capabilities that the Hallway Hangers exhibit. Instead, they attribute lack of success on the job market exclusively to personal inadequacy. This is particularly true when the Brothers speculate about the future jobs the Hallway Hangers and their own friends will have. According to the Brothers, the Hallway Hangers (in Super's words) "ain't gonna get nowhere," not because of the harshness of the job market but because they are personally lacking. The rest of the Brothers share this view.

JM: Some of those guys who hang with Frankie, they're actually pretty smart. They just don't channel that intelligence into school, it seems to me.

CRAIG: I call that stupid, man. That's what they are.

JM: I dunno.

CRAIG: Lazy.

(in a separate interview)

SUPER: They think they're so tough they don't have to do work. That don't make sense, really. You ain't gonna get nowhere; all's you gonna do is be back in the projects like your mother. Depend on your mother to give you money every week. You ain't gonna get a good job. As you get older, you'll think about that, y'know. It'll come to your mind. "Wow, I can't believe, I should've just went to school and got my education."

(in a separate interview)

MOKEY: They all got attitude problems. They just don't got their shit together. Like Steve. They have to improve themselves.

In the eyes of the Brothers, the Hallway Hangers have attitude problems, are incapable of considering their long-term future, and are lazy or stupid.

Because this evidence is tainted (no love is lost between the two peer groups), it is significant that the Brothers apply the same criteria in judging each other's chances to gain meaningful employment. James thinks Mokey is headed for a dead-end job because he is immature and undisciplined. He also blames Juan for currently being out of work. "Juan's outta school, and Juan does *not* have a job (*said with contempt*). Now that's some kind of a senior. When I'm a senior, I'm gonna have a job already. I can see if you're gonna go to college right when you get out of school; but Juan's not doin' nothin'. He's just stayin' home." Juan, in turn, thinks that Mokey and Super will have difficulty finding valuable work because of their attitudes. He predicts that Derek and Craig will be successful for the same reason.

These viewpoints are consistent with the dominant ideology in America; barriers to success are seen as personal rather than social. By attributing failure to personal inadequacy, the Brothers exonerate the opportunity structure. Indeed, it is amazing how often they affirm the openness of American society.

(all in separate interviews)

DEREK: If you put your mind to it, if you want to make a future for yourself, there's no reason why you can't. It's a question of attitude.

SUPER: It's easy to do anything, as long as you set your mind to it, if you wanna do it. If you really want to do it, if you really want to be something. If you don't want to do it . . . you ain't gonna make it. I gotta get that through my mind: I wanna do it. I wanna be somethin'. I don't wanna be livin' in the projects the rest of my life.

MOKEY: It's not like if they're rich they get picked [for a job]; it's just mattered by the knowledge of their mind.

CRAIG: If you work hard, it'll pay off in the end.

MIKE: If you work hard, really put your mind to it, you can do it. You can make it.

This view of the opportunity structure as an essentially open one that rewards intelligence, effort, and ingenuity is shared by all the Brothers. Asked whether their chances of securing a remunerative job are as good as those of an upper-class boy from a wealthy district of the city, they all responded affirmatively. Not a single member of the Hallway Hangers, in contrast, affirms the openness of American society. . . .

Reproduction Theory Reconsidered

This basic finding—that two substantially different paths are followed within the general framework of social reproduction—is a major challenge to economically determinist theories. Two groups of boys from the same social stratum who live in the same housing project and attend the same school nevertheless experience the process of social reproduction in fundamentally different ways. This simple fact alone calls into question many of the theoretical formulations of Bowles and Gintis.⁶ If, as they argue, social class is the overriding determinant in social reproduction, what accounts for the variance in the process between the Brothers and Hallway Hangers? Bowles and Gintis, in considering a single school, maintain that social reproduction takes place primarily through educational tracking. Differential socialization through educational tracking prepares working-class students for working-class jobs and middle-class students for middle-class jobs. But the Hallway Hangers and the Brothers, who are from the same social class background and exposed to the curricular structure of the school in the same manner, undergo the process of social reproduction in substantially different manners. The theory of Bowles and Gintis cannot explain this difference.

Bourdieu's notion of habitus, however, can be used to differentiate the Hallway Hangers and the Brothers.⁷ The habitus, as defined by Giroux, is "the subjective dispositions which

reflect a class-based social grammar of taste, knowledge, and behavior inscribed in . . . each developing person.”⁸ According to Bourdieu, the habitus is primarily a function of social class. Bourdieu does not give an adequate sense of the internal structure of the habitus, but there is some precedent in his work for incorporating other factors into constructions of the habitus; for example, he differentiates people not only by gender and class, but also by whether they come from Paris or not. Although Bourdieu sometimes gives the impression of a homogeneity of habitus within the boundaries of social class, I understand habitus to be constituted at the level of the family and thus can include, as constitutive of the habitus, factors such as ethnicity, educational histories, peer associations, and demographic characteristics (e.g., geographical mobility, duration of tenancy in public housing, sibling order, and family size) as these shape individual action. Although Bourdieu never really develops the notion along these lines, he does allude to the complexity and interplay of mediations within the habitus. “The habitus acquired in the family underlies the structuring of school experiences, and the habitus transformed by schooling, itself diversified, in turn underlies the structuring of all subsequent experiences (e.g. the reception and assimilation of the messages of the culture industry or work experiences), and so on, from restructuring to restructuring.”⁹ When understood along the lines I have indicated, the concept of habitus becomes flexible enough to accommodate the interactions among ethnicity, family, schooling, work experiences, and peer associations that have been documented [here].

Although we may accept the notion of habitus as a useful explanatory tool, we must reject the inevitability of its *function* in Bourdieu’s theoretical scheme. According to Bourdieu, the habitus functions discreetly to integrate individuals into a social world geared to the interests of the ruling classes; habitus engenders attitudes and conduct

that are compatible with the reproduction of class inequality. The outstanding example of this process is the development by working-class individuals of depressed aspirations that mirror their actual chances for social advancement.

The circular relationship Bourdieu posits between objective opportunities and subjective hopes is incompatible with the findings [presented here]. The Brothers, whose objective life chances probably were lower originally than those available to the Hallway Hangers because of racial barriers to success, nevertheless nurture higher aspirations than do the Hallway Hangers. By emphasizing structural determinants at the expense of mediating factors that influence subjective renderings of objective probabilities, Bourdieu presumes too mechanistic and simplistic a relationship between aspiration and opportunity. This component of his theory fails to fathom how a number of factors lie between and mediate the influence of social class on individuals; Bourdieu cannot explain, for instance, how ethnicity intervenes in the process of aspiration formation and social reproduction.

Thus, the theoretical formulations of Bowles and Gintis and the deterministic elements of Bourdieu’s theory, although elegant and intuitively plausible, are incapable of accounting for the processes of social reproduction as they have been observed and documented in Clarendon Heights. These theories give an excellent account of the hidden structural and ideological determinants that constrain members of the working class and limit the options of Clarendon Heights teenagers. What the Hallway Hangers and the Brothers demonstrate quite clearly, however, is that the way in which individuals and groups respond to structures of domination is open-ended. Although there is no way to avoid class-based constraints, the outcomes are not predefined. Bowles and Gintis and Bourdieu pay too little attention to the active, creative role of individual and group

praxis. As Giroux maintains, what is missing from such theories “is not only the issue of resistance, but also any attempt to delineate the complex ways in which working-class subjectivities are constituted.”¹⁰

From Ethnography to Theory

Once we descend into the world of actual human lives, we must take our theoretical bearings to make some sense of the social landscape, but in doing so we invariably find that the theories are incapable of accounting for much of what we see. The lives of the Hallway Hangers and the Brothers cannot be reduced to structural influences or causes; although structural forces weigh upon the individuals involved, it is necessary, in the words of Willis, “to give the social agents involved some meaningful scope for viewing, inhabiting, and constructing their own world in a way which is recognizably human and not theoretically reductive.”¹¹ We must appreciate both the importance and the relative autonomy of the cultural level at which individuals, alone or in concert with others, wrest meaning out of the flux of their lives.

The possibilities open to these boys as lower-class teenagers are limited structurally from the outset. That they internalize the objective probabilities for social advancement to some degree is beyond question. The process by which this takes place, however, is influenced by a whole series of intermediate factors. Because gender is constant in the study discussed in these pages, race is the principal variable affecting the way in which these youths view their situation. Ethnicity introduces new structurally determined constraints on social mobility, but it also serves as a mediation through which the limitations of class are refracted and thus apprehended and understood differently by different racial groups. The Brothers comprehend and react to their situation in a manner entirely different from the response the Hallway Hangers make to a similar situation; ethnicity intro-

duces a new dynamic that makes the Brothers more receptive to the achievement ideology. Their acceptance of this ideology affects their aspirations but also influences, in tandem with parental encouragement, their approach to school and the character of their peer group, factors that in turn bear upon their aspirations.

If we modify the habitus by changing the ethnicity variable and altering a few details of family occupational and educational histories and duration of tenancy in public housing, we would have the Hallway Hangers. As white lower-class youths, the Hallway Hangers view and interpret their situation in a different light, one that induces them to reject the achievement ideology and to develop aspirations and expectations quite apart from those the ideology attempts to generate. The resultant perspective, which is eventually reinforced by the Hallway Hangers' contact with the job market, informs the boys' approach to school and helps us understand the distinctive attributes of this peer group. Thus, although social class is of primary importance, there are intermediate factors at work that, as constitutive of the habitus, shape the subjective responses of the two groups of boys and produce quite different expectations and actions.

Having grown up in an environment where success is not common, the Hallway Hangers see that the connection between effort and reward is not as clearcut as the achievement ideology would have them believe. Because it runs counter to the evidence in their lives and because it represents a forceful assault on their self-esteem, the Hallway Hangers repudiate the achievement ideology. Given that their parents are inclined to see the ideology in the same light, they do not counter their sons' rejection of the American Dream.

A number of important ramifications follow from the Hallway Hangers' denial of the dominant ideology: the establishment of a peer group that provides alternative means of

generating self-esteem, the rejection of school and antagonism toward teachers, and, of course, the leveling of aspirations. In schematizing the role of the peer group, it is difficult not to appear tautological, for the group does wield a reciprocal influence on the boys: It attracts those who are apt to reject school and the achievement ideology and those with low aspirations and then deepens these individuals' initial proclivities and further shapes them to fit the group. But at the same time, the peer subculture itself, handed down from older to younger boys, is the product of the particular factors that structure the lives of white teenagers in Clarendon Heights.

In addition to the peer group, the curricular structure of the school solidifies the low aspirations of the Hallway Hangers by channeling them into programs that prepare students for manual labor jobs. Low aspirations, in turn, make the Hallway Hangers more likely to dismiss school as irrelevant. Once on the job market, the Hallway Hangers' inability to secure even mediocre jobs further dampens their occupational hopes. Thus although each individual ultimately retains autonomy in the subjective interpretation of his situation, the leveled aspirations of the Hallway Hangers are to a large degree a response to the limitations of social class as they are manifest in the Hallway Hangers' social world.

The Brothers' social class origins are only marginally different from those of the Hallway Hangers. Being black, the Brothers also must cope with racially rooted barriers to success that, affirmative action measures notwithstanding, structurally inhibit the probabilities for social advancement, although to a lesser degree than do shared class limitations. What appears to be a comparable objective situation to that of the Hallway Hangers, however, is apprehended in a very different manner by the Brothers.

As black teenagers, the Brothers interpret their families' occupational and educational

records in a much different light than do the Hallway Hangers. Judging by the Brothers' constant affirmation of equality of opportunity, the boys believe that racial injustice has been curbed in the United States in the last twenty years. Whereas in their parents' time the link between effort and reward was very tenuous for blacks, the Brothers, in keeping with the achievement ideology, see the connection today as very strong: "If you work hard, it'll pay off in the end" (Craig). Hence, the achievement ideology is more compatible with the Brothers' attitudes than with those of the Hallway Hangers, for whom it cannot succeed against overwhelming contrary evidence. The ideology is not as emotionally painful for the Brothers to accept because past racial discrimination can help account for their families' poverty, whereas the Hallway Hangers, if the ideology stands, are afforded no explanation outside of laziness and stupidity for their parents' failures. The optimism that acceptance of the achievement ideology brings for the Brothers is encouraged and reinforced by their parents. Thus, we see how in the modified habitus ethnicity affects the Brothers' interpretation of their social circumstances and leads to acceptance of the achievement ideology, with all the concomitant results.

Postscript: The Hallway Hangers and Brothers Eight Years Later

"Hey, Jay, what the fuck brings you back to the Ponderosa?" Greeted by Steve in July 1991, I surveyed a Clarendon Heights that had changed considerably since 1983. Steve jerked his thumb over his shoulder at a group of African American teenagers lounging in the area outside doorway #13, previously the preserve of the Hallway Hangers. "How do you like all the new niggers we got here? Motherfuckers've taken over, man." I asked Steve about Frankie, Slick, and the other Hallway Hangers. "I'm the only one holding down the fort," he

answered. "Me and Jinks—he lives in the back. The rest of 'em pretty much cut loose, man."

In their mid-twenties, the seven Hallway Hangers should be in the labor force full time. Most of them aren't: They are unemployed or imprisoned, or are working sporadically either for firms "under the table" or for themselves in the drug economy. . . . The Hallway Hangers have been trapped in what economists call the secondary labor market—the subordinate segment of the job structure where the market is severely skewed against workers. Jobs in the primary labor markets provide wages that can support families and an internal career structure, but the rules of the game are different in the secondary labor market. Wages are lower, raises are infrequent, training is minimal, advancement is rare, and turnover is high.

When the legitimate job market fails them, the Hallway Hangers can turn to the underground economy. Since 1984 almost all of the Hallway Hangers have at least supplemented their income from earnings in the burgeoning, multibillion-dollar drug market. The street economy promises better money than does conventional employment. It also provides a work site that does not demean the Hallway Hangers or drain their dignity. As workers in the underground economy, they won't have to take orders from a boss's arrogant son, nor will they have to gossip with office colleagues and strain to camouflage their street identities. . . .

Although they have certainly fared better than the Hallway Hangers, the Brothers have themselves stumbled economically in the transition to adulthood. Even more so than the Hallway Hangers, the Brothers have been employed in the service sector of the economy. They have bagged groceries, stocked shelves, flipped hamburgers, delivered pizzas, repaired cars, serviced airplanes, cleaned buildings, moved furniture, driven tow trucks, pumped gas, delivered auto parts,

and washed dishes. They have also worked as mail carriers, cooks, clerks, computer operators, bank tellers, busboys, models, office photocopiers, laborers, soldiers, baggage handlers, security guards, and customer service agents. Only Mike, as a postal service employee, holds a unionized position. Although their experiences on the labor market have been varied, many of the Brothers have failed to move out of the secondary labor market. Instead, like the Hallway Hangers, they have been stuck in low-wage, high-turnover jobs. . . .

These results are depressing. The experiences of the Hallway Hangers since 1984 show that opting out of the contest—neither playing the game nor accepting its rules—is not a viable option. Incarceration and other less explicit social penalties are applied by society when the contest is taken on one's own terms. There is no escape: The Hallway Hangers must still generate income, build relationships, and establish households. Trapped inside the game, the Hallway Hangers now question their youthful resistance to schooling and social norms. Granted the opportunity to do it over again, the Hallway Hangers say they would have tried harder to succeed.

But the Brothers *have* always tried, which is why their experiences between 1984 and 1991 are as disheartening as the Hallway Hangers'. If the Hangers show that opting out of the contest is not a viable option, the Brothers show that dutifully playing by the rules hardly guarantees success either. Conservative and liberal commentators alike often contend that if the poor would only apply themselves, behave responsibly, and adopt bourgeois values, then they will propel themselves into the middle class. The Brothers followed the recipe quite closely, but the outcomes are disappointing. They illustrate how rigid and durable the class structure is. Aspiration, application, and intelligence often fail to cut through the firm figurations of structural inequality. Though

not impenetrable, structural constraints on opportunity, embedded in both schools and job markets, turn out to be much more debilitating than the Brothers anticipated. Their dreams of comfortable suburban bliss currently are dreams deferred, and are likely to end up as dreams denied.

NOTES

1. All names of neighborhoods and individuals have been changed to protect the anonymity of the study's subjects.

2. Erik H. Erikson, *Gandhi's Truth* (New York: Norton, 1969), p. 154.

3. Ronald Reagan, "State of the Union Address to Congress," *New York Times*, 6 February 1985, p. 17.

4. Kenneth I. Spenner and David L. Featherman, "Achievement Ambitions," *Annual Review of Sociology* 4 (1978):376–378.

5. Howard B. London, *The Culture of a Community College* (New York: Praeger, 1978).

6. Samuel Bowles and Herbert Gintis, *Schooling in Capitalist America* (New York: Basic Books, 1976).

7. See Pierre Bourdieu, *Outline of a Theory of Practice* (Cambridge: Cambridge University Press, 1977).

8. Henry A. Giroux, *Theory & Resistance in Education* (London: Heinemann Educational Books, 1983), p. 89.

9. Bourdieu, *Outline of a Theory of Practice*, p. 87.

10. Giroux, *Theory & Resistance*, p. 85.

11. Paul E. Willis, *Learning to Labor* (Aldershot: Gower, 1977), p. 172.

55 ■ Stephen L. Morgan

A New Social Psychological Model of Educational Attainment

The decision of whether or not to enter college is vexing for a substantial proportion of high school students. The alternative choices are clearly in view from early adolescence onward, but there is no simple normative guide for behavior. As a result, everyday commitment to preparation for college instruction varies across high school students, affecting resulting enrollment decisions and the potential gains from college instruction.

Consider two students, Max and Vinny, who have been raised in similar families. Their parents have identical educational and occupational profiles and value hard work to the same degree. By the end of middle school, Max and Vinny have equivalent academic records, and they both have access to equally strong opportunities in their respec-

tive high schools. In fact, at this point in their lives, Max and Vinny differ substantially in only one way, the amount of information that is available in their communities about the likelihood of graduating from college if enrolled and the typical life outcomes that college graduates and college dropouts receive.

In particular, Max attends a high school in which the majority of students carry on to college after high school graduation, and about half of whom then graduate from college. In contrast, Vinny attends a high school in which very few students per year enter college after high school graduation, even though they are also about equally likely to graduate from college after enrolling. As a result of this dearth of information, Vinny

sometimes wonders whether a college education is worth the investment of time and money. And, in fact, every once in a while when thinking through the attractiveness of college, he comes to the conclusion that a college education may not be worthwhile. In these moments, his attention to schoolwork wanes. Max, in contrast, remains diligent all the while. The abundant information that is available to him keeps him focused on ensuring that he is adequately prepared for the next phase of his educational career.

Now suppose that, in addition to differences in the amount of available information, Max's parents have college degrees while Vinny's parents have only high school degrees. In this case, status socialization theory (see Sewell, Haller, and Portes 1969) predicts that Max will have more ambitious college plans because, in general, students' own expectations of future behavior are shaped by the expectations that their parents, teachers, and peers have of them. Even if Max and Vinny receive the same amount of encouragement from their parents, it is likely that Max's teachers, who know his parents from parent-teacher conferences, are more likely to communicate to Max that they expect he will attend college just like his parents did.

Stepping back from the case of Max and Vinny, how important are differences in the amount of available information, on which students form their own expectations of future behavior, relative to social influence processes that shape expectations as well? And how do teachers and peers form their expectations? Do they transmit accurate information about a student's likely success in the educational system? Or do they systematically mislead students, so that some students who could succeed in college are discouraged while others who would likely fail are encouraged?

The stochastic decision tree model of commitment that I develop here is designed to enable the explicit modeling of questions

such as these, based on a belief formation framework that gives separable influence to the accuracy and amount of information available to students. Before laying out the primary features of the model, I first adopt a specific conception of commitment that generates everyday behavioral preparation. I then introduce a standard decision tree model of educational choice, which is then modified for use in the stochastic decision tree model of commitment, as justified by the literature on cognitive constraints and bounded rationality.

Commitment and Preparation

The stochastic decision tree model of commitment is grounded on the specific assumption that in the period leading up to the college entry decision:

- A1. Everyday courses of behavior are self-regulated by commitments toward alternative future courses of behavior.

For example, the strength of a high school student's commitment to the future course of behavior "Go to college immediately following high school" is the ease with which he or she is able to envision entering and ultimately graduating from a college degree program.

For modeling educational attainment, there are three types of commitment toward alternative futures: purposive, normative, and imitative. For the nonrepeatable decision "Go to college" versus "Do not go to college," these three types of commitment are set in response to three forward-looking prediction rules: "I will go to college if I perceive it to be in my best interest to do so," "I will go to college if my significant others perceive it to be in my best interest to do so," and "I will go to college if I expect other students similar to me will also go to college." These three generative dimensions are roughly analogous to the self-reflection, adoption, and

imitation mechanisms of status socialization theory (see Morgan 1998).

Although all three dimensions of commitment are important, purposive commitment has the most power to subsume the other two. I will therefore focus development of the stochastic decision tree model here on the construction of purposive commitment (see Morgan 2005, Chapters 5 and 6, for proposals for how to then introduce reinforcing and destabilizing normative and imitative sources of commitment into the model). Thus, I motivate the development of the stochastic decision tree model by assuming that:

A2. Individuals use a decision tree to identify and commit to the future course of behavior that they believe is in their best interest.

Having defined commitment as a cognitive attachment to a future course of behavior, the course of everyday behavior that positions an individual to realize their commitment can then be labeled preparatory commitment, or even more simply as just preparation. I assume that:

A3. An individual's level of preparation for a future course of behavior is a direct function of the strength of their commitment to that course of behavior.

Accordingly, a student with maximal commitment toward the future course of action "I will enroll in college" will enact all possible behavior that prepares the student for enrolling in college and then successfully obtaining a college degree. While in high school, such students will take college preparatory classes, complete their homework diligently, focus their attention when taking tests, sign up early for college entrance examinations, and investigate their range of realistic college alternatives. And, because the payoff to college instruction is a function of

the ability to benefit from opportunities for learning, commitment and preparation determine final levels of well-being, above and beyond simply pushing students over a decision threshold to enroll in college.

A Decision Tree Model of Educational Choice

Figure 1 presents a standard decision tree for choice under uncertainty. This specific tree is isomorphic with the set of equations used to model students' enrollment decisions by Manski (1989). For this model, a prospective college student must contemplate whether it is in their best interest to choose the course of action "Go to college" or "Do not go to college."

Assume that students consider two possible abstract life outcomes—a very good position in life denoted by "High" and a not very good position in life denoted by "Low." High school students must decide which among three alternative paths through the educational system will put them in the best possible position to obtain the life position High rather than Low: (1) Go to college and graduate with a degree, (2) Go to college but fail to graduate, (3) Forego college and enter the labor market immediately. For rational choice models of educational attainment, students are presumed to think of these paths as alternative lotteries that are controlled by success parameters π , α , β , and γ that are probabilities between 0 and 1.

In particular, each student's decision tree has a parameter π that represents each student's subjective belief about the probability of completing college if initially enrolled. Similarly, students maintain a set of beliefs about the relative likelihood of attaining High versus Low after traversing each of the three possible paths through the educational system. They expect that if they complete college, they will attain High with probability α and Low with probability $(1-\alpha)$. If they

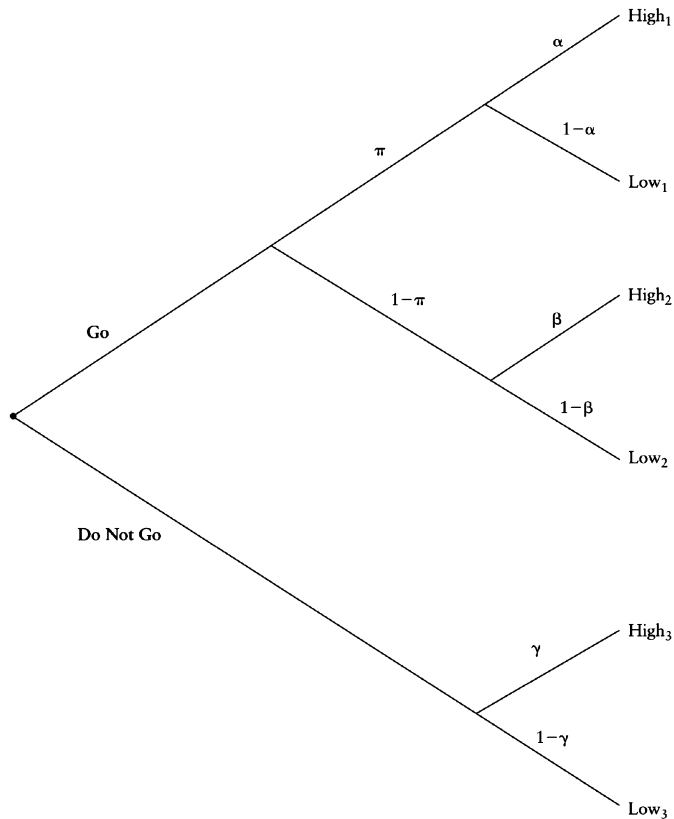


Figure 1. A Simple Decision Tree for the College Entry Decision

attempt but do not complete college, then they expect that they will attain High with probability β and Low with probability $(1-\beta)$. And finally, if they choose to forego college, they expect that they will attain High with probability γ and Low with probability $(1-\gamma)$.

Rational choice theory offers the prediction that, under specific assumptions about optimization and consistency of choice, prospective college students will choose to go to college if:

$$\pi\alpha[u(\text{High}_1)-u(\text{Low}_1)]+(1-\pi)\beta[u(\text{High}_2)-u(\text{Low}_2)] > \gamma[u(\text{High}_3)-u(\text{Low}_3)] \quad (1a)$$

where $u(\cdot)$ is a utility function that, in its most general sense, assigns subjective value

to the alternative payoffs High and Low. If an individual's utility function is process-independent so that $u(\text{High}_1)=u(\text{High}_2)=u(\text{High}_3)$ and $u(\text{Low}_1)=u(\text{Low}_2)=u(\text{Low}_3)$, then both sides of Equation (1a) can be divided by a common utility difference, $u(\text{High})-u(\text{Low})$, in order to obtain the simplified decision rule:

$$\pi\alpha+(1-\pi)\beta > \gamma \quad (1b)$$

The left-hand sides of Equations (1a) and (1b) represent the expected utility of enrolling in college, and the right-hand sides represent the expected utility of not enrolling in college. If the expected utility of enrolling is greater than the expected utility of not enrolling, rational choice theory assumes that a

high school student will choose to enroll in college.

After students choose either the upper or lower branch of the tree—and thus either enroll in college or enter the labor force—they are subject to a set of real-life lotteries. The probability of reaching High instead of Low is a function of more than simply a student's decision. Each lottery has a “true” probability, denoted respectively as π , α , β , and γ . Actual outcomes are therefore determined in large part by exogenous factors that structure these probabilities. For researchers, claims about the relative sizes of π , α , β , and γ are usually based on assumptions about institutional constraints on educational attainment and the distribution of High versus Low outcomes in society. It is generally assumed that there are positive returns to schooling so that $\alpha \geq \beta$ and $\beta > \gamma$. However, Manski (1989:307) argues that “analysis is trivial” unless, in the notation used here, for at least some students $\beta < \gamma$ and $\beta < \gamma$. If that were not the case, all students would choose to enter college.

When contemplating college entry decisions, do students maintain reasonable beliefs about the lotteries that they face? In other words, for each individual and for a decision tree such as the one depicted in Figure 1, is there a close correspondence between π , α , β , and γ and $\tilde{\pi}$, $\tilde{\alpha}$, $\tilde{\beta}$, and $\tilde{\gamma}$? The traditional assumption among rational choice researchers is that individuals, on average, have correct beliefs about $\tilde{\pi}$, $\tilde{\alpha}$, $\tilde{\beta}$, and $\tilde{\gamma}$. Accordingly, the values of π , α , β , and γ that students rely on when making college entry decisions are assumed to equal the true probabilities that students will face in their futures.

The model I develop in the next section provides a foundation for a more comprehensive approach. The model is based in part on rational choice theory but also embraces the core claim of the Wisconsin model of Sewell, Haller, and Portes (1969) that one's

current behavior is conditioned by socially structured beliefs about one's future.

A Stochastic Decision Tree Model of Commitment

The action “Go to college” can be thought of as a compound outcome of a series of underlying decisions, many of which must be enacted long before the first college tuition bill is due. Moreover, the payoff to obtaining a college degree is a function not just of having enrolled in college but of how seriously one has prepared to master the college curriculum before being exposed to it.

Decision Tree Structure

For the college entry decision, I assume that:

A4. The decision tree that individuals use to commit to a future course of behavior is more simple in structure than that which would be required to explicitly model all underlying decisions that are consequential for the college entry decision.

Equivalently, I assume that preparation cannot be modeled as a long series of fully conscious, forward-looking decisions. For the numerical demonstration of the framework that I offer in Morgan (2005), I use the simple decision tree in Figure 1, and hence, when referring to the college choice example, adopt the position of Manski (1989) that for enrollment decisions potential enrollees consider four basic lotteries. Trees with alternative and slightly more elaborate structures are permissible, but I stipulate with Assumption A4 that these must remain simple.

For example, if a student were contemplating a high school course enrollment decision, the decision tree in Figure 1 could be expanded to model this decision explicitly. In this case, the decision tree in Figure 1 could be specified as two alternative subtrees for

the pressing current decision: “Enroll in trigonometry” or “Do not enroll in trigonometry.” A more comprehensive tree of this form would be justified if beliefs such as π or the utility evaluations of High and Low depend on whether or not one enrolls in the trigonometry class.

Even though some very consequential short-run decisions may be approached in this sophisticated way, I presume that students have no desire to use a finely specified decision tree for every consequential underlying decision they must make, such as whether to study hard for an upcoming test or whether to do their homework on any given night. Were this the case, the large number of preparatory commitment decisions that adolescents must navigate in middle school and high school would render any such comprehensive decision trees massive and ever-changing as each intermediate decision is considered and then enacted.

Belief Distributions

Instead of requiring high school students to construct comprehensive decision trees for all intermediate decisions, I allow the complexity of the everyday decisions that constitute preparation to be modeled as uncertainty about the parameters of a simple decision tree for a future course of behavior. Thus, I assume that:

A5. The parameters of the decision tree that students use to commit to a future course of behavior are fundamentally stochastic.

This assumption is the major departure from standard practice in rational choice modeling of educational attainment. For the stochastic decision tree model of commitment, students recognize that the decision tree they are consulting is simply a rough approximation to the complex set of real-life lotteries to which they will ultimately be subjected. Thus, students forecast their futures

as if any ostensibly true parameters such as $\tilde{\pi}$, $\tilde{\alpha}$, $\tilde{\beta}$, and $\tilde{\gamma}$ have distributions. Accordingly, Assumption A5 stipulates that students maintain entire belief distributions for the parameters of a decision tree, such as π , α , β , and γ in Figure 1.

How are belief distributions constructed and revised? I address this question in substantial detail in Morgan (2005), but here I simply assume for brevity that individuals store belief distributions in memory based on observations of individuals whom they regard as payoff models. Furthermore, I assume that individuals with sparse information at their disposal will have uncertain belief distributions with large variances, and individuals with abundant information will have precise belief distributions with small variances. For example, consider two high school students contemplating the decision tree presented in Figure 1 when forecasting their own future decision of whether to enter college immediately. Student A observes twenty college graduates and ten college dropouts. In contrast, Student B observes only two college graduates and one college dropout. The framework I will adopt stipulates that Student A will have more confidence than Student B in eliminating extreme values from his or her belief distributions about the subjective probability of graduating from college if enrolled. Student A would consider values such as .66 much more likely than values such as .96 or .36. Student B would likewise consider .66 more likely than either .96 or .36 but would be less willing than Student A to discount the plausibility of .96 or .36. This position is consistent with the dominant literature on belief distributions, which for Bayesian decision theory (see Pratt, Raiffa, and Schlaifer 1995) and information-integration theory in psychology (see Davidson 1995) similarly allows uncertainty and imprecision of beliefs to be a function of the amount of available information.

Decision Tree Evaluation

How do students evaluate their stochastic decision trees? By principles of statistical decision theory, the traditional answer would be that they calculate the means of their belief distributions, declare them as their best estimates of their point-value beliefs, and use them to solve a decision rule such as Equation (1a). Instead, I accept the position from the bounded rationality literature that individuals are subject to substantial cognitive constraints. Accordingly, they cannot numerically evaluate density functions in order to quickly recover arbitrarily exact estimates of the means of their belief distributions. Rather, individuals must rely on “fast and frugal” strategies to evaluate their decision trees (see Gigerenzer and Selten 2001). Based on this position, I assume that:

A6. Students evaluate their stochastic decision trees by solving a decision rule with parameter values that are simple averages of a few randomly sampled candidate values drawn from their belief distributions.

In the numerical simulations of the framework offered in Morgan (2005), I allow students to randomly draw only a few candidate parameter values for a decision tree of the structure presented in Figure 1 and then average over these values to form a set of candidate point-value beliefs $\{\pi', \alpha', \beta', \gamma'\}$. Assuming that utility evaluations are equivalent across paths and scaled to values of 1 for High and 0 for Low, an individual will set a dichotomous variable E equal to 1 if:

$$\pi' \alpha' + (1 - \pi') \beta' > \gamma' \quad (1c)$$

Otherwise, E will be set equal to 0. E signifies an instantaneous commitment to the future affirmative decision “Go to college” (for example, the choice Go in Figure 1).

Over a set of J evaluations of the same stochastic decision tree with stable belief distri-

butions, the average level of commitment to future college attendance is then:

$$\Pr(E=1) = \frac{1}{J} \sum_{j=1 \dots J} E_1 \quad (2)$$

$\Pr(E=1)$ is therefore equal in expectation to the probability that E will be evaluated as 1 for a given set of beliefs and prespecified total number of candidate parameter draws selected for decision tree evaluation. Thus, whereas the binary orientation E represents the instantaneous commitment that guides each preparation decision, $\Pr(E = 1)$ represents the average level of commitment that is consistent with a set of stable beliefs and a procedure to analyze those beliefs in the process of constructing forward-looking commitment. The distinction implies that for two students who at a specific point in time both set E equal to 1, the student whose beliefs imply a lower value for $\Pr(E = 1)$ will be less likely to set E equal to 1 when contemplating subsequent preparation decisions. $\Pr(E = 1)$ is therefore the more fundamental quantity of interest.

Why would students use this form of decision evaluation? I assume that students generally wish to orient their current behavior to a long-run plan that is in their best interest, but doing so in a way that is as rigorous as stipulated by statistical decision theory is too costly, especially since students themselves must recognize that they cannot form comprehensive decision trees to capture all consequential intermediate decisions. Thus, students seek to avoid short-run mistakes in judgment by adopting a relatively frugal process of planning for their futures.

The simulations and the analytic results in Chapter 4 of Morgan (2005) demonstrate that when the uncertainty of beliefs is allowed to explicitly enter a decision evaluation, preparation for a future course of behavior such as college entry is sensitive to the amount and type of uncertainty in one’s beliefs as well as the effort expended to analyze

those beliefs. Thus, even in a world where the returns on investments in higher education are massive, if beliefs are the least bit uncertain, and if students have limited information processing and analysis capacities, some high school students some of the time will perceive it to be in their future best interest to forego a college education. In these episodic and contrarian instances, they will adjust their current behavior accordingly, failing to enact preparation decisions that their beliefs nonetheless imply are genuinely in their best interest.

Conclusions

To construct the stochastic decision tree model of commitment, I have borrowed and then extended the strongest pieces of both status socialization and rational choice models of educational attainment. The key innovation of the framework is the specification of a simple decision tree with fundamentally stochastic parameters. When a simple, boundedly rational decision rule is invoked to specify the way in which individuals draw forecasts of future behavior from their stochastic decision trees, commitment and preparation can be shown to be functions of the accuracy and precision of beliefs about future courses of behavior. If, as assumed by status socialization theory, current behavior is a function of beliefs about the future, then everyday forecasts of future behavior operate as self-fulfilling prophecies by regulating preparation for the course of behavior that is thought to be in one's best interest.

The conception of educational attainment that I have developed here can be read as a strategic evasion of the question that is the title of Diego Gambetta's 1987 book, *Were They Pushed or Did They Jump? Individual Decision Mechanisms in Education*. From the

perspective developed here, students wander through a series of daily decisions, adjusting current behavior to anticipation of future behavior. If belief formation has its own inertia, is easily contaminated through social influence, and regulates current behavior, then it would seem best not to conceive of students as if they are standing on the edge of a cliff, contemplating whether to jump. It should suffice to see students as if they are on the edge of commitment to alternative futures. Although less dramatic, this metaphor is sufficient to motivate the new generation of theoretically guided empirical analysis that is now needed.

REFERENCES

- Davidson, Andrew R. 1995. "From Attitudes to Actions to Attitude Change: The Effects of Amount and Accuracy of Information." In *Attitude Strength: Antecedents and Consequences*, edited by R. E. Petty and J. A. Krosnick. Mahwah, NJ: Erlbaum.
- Gambetta, Diego. 1987. *Were They Pushed or Did They Jump? Individual Decision Mechanisms in Education*. Cambridge: Cambridge University Press.
- Gigerenzer, Gerd, and Reinhard Selten, eds. 2001. *Bounded Rationality: The Adaptive Toolbox*. Cambridge, MA: MIT Press.
- Manski, Charles F. 1989. "Schooling as Experimentation: A Reappraisal of the Postsecondary Dropout Phenomenon." *Economics of Education Review* 8: 305–312.
- Morgan, Stephen L. 1998. "Adolescent Educational Expectations: Rationalized, Fantasized, or Both?" *Rationality and Society* 10: 131–162.
- _____. 2005. *On the Edge of Commitment: Educational Attainment and Race in the United States*. Stanford, CA: Stanford University Press.
- Pratt, John W., Howard Raiffa, and Robert Schlaifer. 1995. *Introduction to Statistical Decision Theory*. Cambridge, MA: MIT Press.
- Sewell, William H., Archibald O. Haller, and Alejandro Portes. 1969. "The Educational and Early Occupational Attainment Process." *American Sociological Review* 34: 82–92.

■ LABOR MARKETS ■

56 ■ Michael J. Piore

The Dual Labor Market

Theory and Implications

The central tenet of [my] analysis is that the role of employment and of the disposition of manpower in perpetuating poverty can be best understood in terms of a dual labor market. One sector of that market, which I have termed elsewhere the primary market,¹ offers jobs which possess several of the following traits: high wages, good working conditions, employment stability and job security, equity and due process in the administration of work rules, and chances for advancement. The secondary sector has jobs that are decidedly less attractive, compared with those in the primary sector. They tend to involve low wages, poor working conditions, considerable variability in employment, harsh and often arbitrary discipline, and little opportunity to advance. The poor are confined to the secondary labor market. Eliminating poverty requires that they gain access to primary employment.

The factors that generate the dual market structure and confine the poor to the secondary sector are complex. With some injustice to that complexity, they may be summarized: First, the most important characteristic distinguishing primary from secondary jobs appears to be the behavioral requirements they impose upon the work force, particularly that of employment stability. Insofar as secondary workers are barred from primary jobs

by a real qualification, it is generally their inability to show up for work regularly and on time. Secondary employers are far more tolerant of lateness and absenteeism, and many secondary jobs are of such short duration that these do not matter. Work skills, which receive considerable emphasis in most discussions of poverty and employment, do not appear a major barrier to primary employment (although, because regularity and punctuality are important to successful learning in school and on the job, such behavioral traits tend to be highly correlated with skills).

Second, certain workers who possess the behavioral traits required to operate efficiently in primary jobs are trapped in the secondary market because their superficial characteristics resemble those of secondary workers. This identification occurs because employment decisions are generally made on the basis of a few readily (and hence inexpensively) assessed traits like race, demeanor, accent, educational attainment, test scores, and the like. Such traits tend to be statistically correlated with job performance but not necessarily (and probably not usually) causally related to it. Hence, a number of candidates who are rejected because they possess the “wrong” traits are actually qualified for the job. Exclusion on this basis may

Michael J. Piore. “The Dual Labor Market: Theory and Implications,” in *The State and the Poor*, edited by Samuel H. Beer and Richard E. Barringer, pp. 55–59. Copyright © 1970 by Winthrop Publishers.

be termed *statistical discrimination*. In addition to statistical discrimination, workers are also excluded from primary employment by *discrimination pure and simple*.

Discrimination of any kind enlarges the labor force that is captive in the secondary sector, and thus lowers the wages that secondary employers must pay to fill their jobs. Such employers thus have an economic stake in perpetuating discrimination. Since it limits the supply of labor in the primary sector and raises the wages of workers who have access to jobs there, primary workers also have a stake in discrimination. Discrimination pure and simple is not generally of economic value to primary employers, since it forces them to pay higher wages without obtaining corresponding economic gains. In statistical discrimination, however, the higher wages are compensated by the reduced cost of screening job candidates, and here primary employers share the interest of secondary employers and primary workers in perpetuating such discrimination.

Third, the distinction between primary and secondary jobs is not, apparently, technologically determinate. A portion—perhaps a substantial proportion—of the work in the economy can be organized for either stable or unstable workers. Work normally performed in the primary sector is sometimes shifted to the secondary sector through subcontracting, temporary help services, recycling of new employees through probationary periods, and the like. Nor is the primary-secondary distinction necessarily associated with a given enterprise. Some enterprises, most of whose jobs constitute primary employment and are filled with stable, committed workers, have subsections or departments with inferior job opportunities accommodated to an unstable work force. Secondary employers generally have a few primary jobs, and some have a large number of them. Nonetheless, despite a certain degree of elasticity in the distribution of work between the primary and secondary sections, shifts in the distribution generally involve changes in the techniques of produc-

tion and management and in the institutional structure and procedures of the enterprises in which the work is performed. The investment necessary to effect these changes acts to strengthen resistance to antipoverty efforts.

Fourth, the behavioral traits associated with the secondary sector are reinforced by the process of working in secondary jobs and living among others whose life-style is accommodated to that type of employment. Hence, even people initially forced into the secondary sector by discrimination tend, over a period of time, to develop the traits predominant among secondary workers. Thus, a man who works in a world where employment is intermittent and erratic tends to lose habits of regularity and punctuality. Similarly, when reward and punishment in the work place are continually based upon personal relationships between worker and supervisor, workers forget how to operate within the impersonal, institutional grievance procedures of the primary sector. When such workers do gain access to primary jobs, they are frustrated by the system's failure to respond on a personal basis and by their own inability to make it respond on an institutional basis.

Finally, among the poor, income sources other than employment, especially public assistance and illicit activity, tend to be more compatible with secondary than with primary employment. The public assistance system discourages full-time work and forces those on welfare either into jobs that are part-time or into jobs that pay cash income which will not be reported to the social worker or can be quickly dropped or delayed when the social worker discovers them or seems in danger of doing so. The relationship between social worker and client builds upon the personal relationship that operates in the secondary sector, not on the institutional mechanisms that tend to operate in the primary sector. Illegitimate activity also tends to follow the intermittent work pattern prevalent in secondary employment, and the attractions

of such activity, as well as life patterns and role models it presents to those not themselves involved but associating with people who are, foster behavioral traits antagonistic to primary employment.

The dual market interpretation of poverty has some central implications: the poor do participate in the economy; the manner of their participation, not the question of participation as such, constitutes the manpower problem of the poor; and their current mode of participation is ultimately a response to a series of pressures—economic, social, and technical—playing upon individuals and labor market institutions. This suggests that a distinction can be drawn between policies that are designed to alleviate the pressures which generate the dual market structure and those that attempt to attack the problem directly by moving individuals from secondary to primary employment. The latter policies combat prevailing pressures but leave intact the forces that generate them. The thrust of [my] argument is that in concentrating upon training, counseling, and placement services for the poor, manpower policy has overemphasized direct approaches, and that more weight should be placed upon policies which affect the environment in which employment decisions are made and the pressures which the environment generates. Among such policies are antidiscrimination policy, occupational licensing reform, and the structure of public assistance.

Analysis of the dual labor market suggests a further implication: because the “poor” do participate in the economy, certain groups are interested in that participation and how it occurs. Policies aimed at moving the poor out of the secondary market work against the interests of these groups and therefore are in danger of being subverted by them. This danger is a major reason for concentrating on indirect approaches that are not susceptible to the same kind of subversion; in fact, because such approaches alleviate the pressures generating the dual market structure, they

reduce the resistance to policies that move directly against that structure. The dangers to which existing institutions subject programs designed to move the poor directly out of the secondary market are twofold. The new institutions created by these programs can be rejected by the prevailing economic system and isolated off to one side; a program, for example, would then recruit workers for training in skills that are little utilized in either the secondary or the primary market. Alternatively, the new institutions may be captured by the prevailing economic system and used to facilitate its operation; for example, neighborhood employment offices may recruit secondary workers for secondary jobs, and training may be provided in primary employment to workers who would have gotten it anyway in establishments that would have financed it themselves. The central problem in the design of direct approaches to manpower programs is to organize them in such a way that they can resist this two-fold threat of rejection on the one hand and capture on the other.

These conclusions follow directly from the dual market interpretation of the poverty problem but they are not uniquely dependent upon it. The dual labor market is one of a class of theoretical constructs which views poverty in the United States in terms of a dichotomy in the economic and social structures. Such a dichotomy is implicit in the concept of a “culture of poverty” and in the expression of public policy goals associated with poverty in terms of an income cutoff. Most such views of poverty entertain the idea that the dichotomy is a product of forces endogenous to the economy (or, more broadly, the society as a whole). It follows that attempts to eliminate poverty will tend to run counter to the natural operation of the economy, and that they will be resisted by existing institutions and are in danger of rejection. To say all this is perhaps to say simply that if poverty were easy to eliminate, it wouldn't be around in the first place. But it does at least

identify as a certain problem in the program design the task of equipping the institution which works with the poor to withstand the rejection pressures.

What the dual labor market interpretation implies that is not implicit in other dichotomous interpretations is that the poor are separated from the nonpoor not only in the negative sense of exclusion from activities and institutions to which the nonpoor have access, but also in the positive sense that they have economic value where they are; that, in

other words, *there are groups actively interested in the perpetuation of poverty*. It is this interest that makes new institutions created to work with the poor in the labor market subject to threats of capture as well as of rejection.

NOTE

1. See Michael J. Piore, "On-The-Job Training in the Dual Labor Market," in Arnold Weber, *et al.*, *Public-Private Manpower Policies* (Madison, Wisc.: Industrial Relations Research Association, 1969), pp. 101–132.

57 ■ Aage B. Sørensen and Arne L. Kalleberg

An Outline of a Theory of the Matching of Persons to Jobs

Much recent research in sociology has focused on labor market processes. These concerns include analysis of the processes that produce variation in individual earnings by characteristics of people and their jobs; the analysis of career patterns and job mobility processes; and the analysis of employment and unemployment patterns of various population groups. Sociologists share many of these concerns with economists, and there is much overlap in research topics among sociologists and economists.

Despite similarities in methodology and research design, the research traditions in sociology and economics have quite different intellectual backgrounds. Most empirical research on labor market processes in economics is guided by the dominant school of labor economics—the neoclassical theory of wage determination and labor supply, with mar-

ginal productivity theory accounting for the demand side and human capital theory taking care of the supply side. In contrast, sociological research on labor market phenomena has its origin in research describing socioeconomic attainment and social mobility processes for various population groups. Sociological research on attainment and mobility has not employed an explicitly stated conceptual apparatus that informs the choice of variables and the interpretation of parameters. Although there is a growing body of findings about the magnitude of the influences of various variables on the outcomes of labor market processes, particularly income attainment, there are few efforts by sociologists to identify the mechanisms that create the influences of personal and job characteristics on income and earnings or on the other labor market outcomes.

There is no need for sociologists to develop a unique theory of labor market processes if the neoclassical economic theory adequately accounts for the findings of empirical research. With respect to a favorite variable of both economists and sociologists—that is, education—human capital theory does provide an interpretation of results. However, the economic theory does not provide a rationale for the sociological concern for occupational attainment. Job characteristics, including those presumably captured by the Socioeconomic Index (SEI) or prestige scores of occupations, play little or no role in the orthodox economic theory. Still, occupational status accounts for a substantial fraction of the explained variance in sociological income attainment models.

The amount of variance added to income attainment models by occupation is not necessarily a strong argument for replacing or supplementing the economic theory. Sociologists have not been able to account for very much variance in income attainment. Research informed by human capital theory (e.g., Mincer, 1974) has in fact been able to do as well or better without including occupation. A measure of occupational status must necessarily show some relation to income, reflecting the between-occupation variance in income that it captures. An observed effect of job characteristics on income or earnings may be attributed to a misspecification of sociological models, both with respect to functional form and omitted variables, and need not be considered a challenge to the economic theory.

There are, however, other reasons for critically evaluating the neoclassical or orthodox economic theory. The economic theory is powerful, and numerous predictions can be derived from it regarding the earnings attainment process and other labor market processes, particularly labor supply. (A list of such predictions is presented by Becker, 1964.) Some of these predictions are borne out by empirical observations; some are not.

Thurow (1975, pp. 56–70) presents a list of deviations from the theory, pertaining to such issues as the relationship between wages and unemployment, changes in the distribution of earnings, and the relationship between the distribution of education and the distribution of income. Numerous others have identified features of the earnings attainment process and of labor markets that deviate from the assumptions and predictions of the neoclassical theory. A review of these challenges to orthodox theory has been presented by Cain (1976). Particularly important are those critiques that argue that labor markets are segmented and that stress the differences between either so-called primary and secondary jobs (cf. Doeringer & Piore, 1971); or monopoly, competitive, and state economic sectors (cf. Averitt, 1968; Bluestone, 1970; O'Connor, 1973); or wage competition and job competition sectors (Thurow, 1975); or internal and external markets (Doeringer & Piore, 1971; Kerr, 1954). These critiques all observe that jobs and job structures differ, contrary to the assumption about the homogeneous nature of labor markets made by the economic theory. They stress qualitative differences among jobs relevant for employment and earnings processes and claim to be able to account for the observations that deviate from the orthodox economic theory, as well as to provide different explanations for labor market processes that also can be explained by the orthodox theory. An example of such an alternative explanation is Thurow's (1975) interpretation of the relationship between education and earnings.

Most of the criticism comes from within economics, though there are examples of research and conceptual elaboration by sociologists pertaining to the issues raised by the segmented labor market theory (Sørensen, 1977; Spilerman, 1977; Stolzenberg, 1975). The issues are clearly relevant for sociological research, and more so since the alternatives to the neoclassical theory provide a rationale

for introducing job characteristics sociologists are likely to continue to emphasize.

The classical sociological theorists did not leave labor market analysis to economists. Marx and Weber spent lifetimes analyzing the relation between economy and society, and their concerns in many ways parallel the issues raised in recent controversies. Marx's analysis of capitalist society is an analysis of the implications of the fundamental condition of capitalist production: Labor is treated as a commodity bought and sold freely in a market. This conception of the labor market, we shall argue in the following pages, parallels the conception of the orthodox economic theory.

Marx treated labor in capitalist society as a homogeneous abstract category, and though there are occasional remarks concerning deviations from this model of labor as a commodity and their relevance for class conflict (e.g., Marx, 1961, Vol. 1, chap. 14), no systematic analysis of alternative labor market structures is presented. Weber's long analysis of the sociological categories of economic action (Weber, 1947, Pt. I, chap. 2) provides, in contrast, numerous concepts relevant for the analysis of labor market structures (including nonmarket relationships), particularly in the sections on the social division of labor. The concepts are highly relevant for the issues raised by the challenges to orthodox economic theory, and some of Weber's basic concepts will be used extensively in this chapter.

The following pages provide a conceptual framework for the analysis of labor markets. Labor markets are arenas for the matching of persons to jobs. The conditions that determine the earnings outcome of this matching process are of primary interest here, particularly the identification of what determines the influence of job and personal characteristics on earnings. The purpose of this chapter is not to show the neoclassical theory to be wrong, but rather to identify the conditions for the emergence of the matching process

associated with the labor market structure assumed in the orthodox economic theory. It will be argued that the conditions for the emergence of this matching process are not present in some segments of the labor market. The absence of these conditions leads to alternative matching processes, and a model of one important alternative matching process will be presented. The two contrasting matching processes will be shown to have very different implications for the earnings determination process and for other labor market processes.

Basic Concepts

The theory proposed in this chapter will rely on Weber's notion of open and closed social relationships (Weber, 1947, p. 139) to identify different job structures characterized by different matching processes.¹ The degree of closure, in turn, is seen as determined by the bargaining power of employers and employees. We shall, therefore, refer to the employment relationship as the crucial determinant of the notion of the matching process and its earnings outcome.

Employment relationships are social relationships created in the production of goods and services between an employer (or his agent) and an employee. We concentrate on employment relationships typical of capitalist production in which the employer appropriates the output from the production process and has complete possession over the nonhuman means of production. Our analysis will focus on the consequences for the earnings determination process and other labor market processes of variation in control over the job by the employer versus the employee. Two aspects of control over the job may be distinguished. One is control over the activities of the job, resulting in more or less autonomy for the employee; the other is control over access to the job, resulting in a more or less closed employment relationship. These two dimensions may vary

independently. Particularly, control over access to the job will be considered crucial, because it influences the nature of competition among employees.

The degree of control over access is a continuum. At one extreme, the employee “owns” the job and no one else can get access unless the current incumbent voluntarily leaves it and a vacancy is established. The length of the employment is then completely controlled by the employee, and the employment relationship is closed to outsiders. At the other extreme, the employer may replace the incumbent at any time. The employment contract is reestablished in every short interval of time, and the employment relationship is completely open to outsiders.

The employment relationship is established in a process assumed to involve purposive actors as employers and employees where both parties are attempting to maximize earnings. The earnings of the employer are determined by the value of the product of the job-person combination in relation to costs of production. The value of production is a question of prices of products and quantity produced. Quantity produced in turn reflects the performance of the employee and the technology used, including the technical division of labor adopted. For purposes of this analysis, the main variable of interest is the performance of the employee and the main costs of production of interest are the wages paid to the employee and the costs of supervision.

The performance of employees or the quantity of labor supplied will be taken as determined by such attributes of the employees as their skills, abilities, and effort. The employer’s return from production evidently depends on his or her ability to obtain the highest output at the lowest costs. While numerous factors may influence the overall level of wages, the employer’s ability to minimize costs of production depends not only on the overall level of wages but also on the ability to tie variations in wages

paid to variations in the employee’s productivity. The main argument of this chapter is that the mechanisms the employer can use to relate wages to performance depend on the employment relationship, particularly the employee’s control over access to the job, and that these different mechanisms identify important differences in labor market structures relevant also for labor market processes other than earnings.

The orthodox economic theory identifies a particular set of mechanisms for relating the productivity of employees to their earnings. We shall first consider these mechanisms and the employment relationships needed for these mechanisms to be effective.

The Neoclassical Theory of Earnings Determination

In the economic theory, a wage rate is generated by a labor market as a result of the demand and supply schedules of labor. Demand for labor varies with the derived demand for products, as reflected in their value. The link between wages and the value of products is established through the concept of marginal productivity, since profit-maximizing firms will be in equilibrium when the value of the marginal product equals the marginal cost or price of labor as a factor of production. This should produce different wage rates for identical labor supply because of differences in demand. However, the neoclassical theory emphasizes supply differences as a source of differences in wage rates and earnings, in particular those supply differences resulting from different skills and other individual characteristics related to an employee’s productive capacity.

Differences in skills, according to human capital theory, determine different levels of productive capacity resulting in different wage rates. If skills were acquired at no cost, those wage differentials would soon lead to equalizing skill acquisition. But skills are acquired at costs. These costs are partly direct

in the form of tuition and living expenses and partly opportunity costs in the form of earnings foregone. No one should undertake training if the returns from this training, in the form of increased earnings accumulated over the working life, are not at least equal to the costs of training.

If only skills acquired through training are relevant, earnings differentials would be exactly off-setting the differences in training costs. However, it is usually recognized that earnings differentials also capture variations in ability, where ability is used to refer to such characteristics as IQ, motivation, and creativity. Ability may be incorporated in the theory by recognizing that persons with different abilities have different investment costs and hence need different earnings to induce the undertaking of training. In addition, some aptitudes may be innate and scarce; these will command a rent because of their fixed supply. Finally, some variation in earnings can be attributed to different opportunities for financing training, particularly as a result of the unequal distribution of parental wealth in combination with the unwillingness of lenders to take collateral in human capital.

The basic proposition derived from the neoclassical theory is then that differences in earnings reflect differences in the productive capacity of persons as a result of their training, abilities, and training opportunities. There may be transient variations in earnings as a result of differences in derived demand in combination with market imperfections, but the basic source of inequality in earnings is unequal endowments in productive capacities among persons. In other words, identical persons are assumed to obtain almost identical earnings, regardless of the characteristics of the jobs they are in.

This theory can be used to account for a number of features of observed earnings attainment processes. Most importantly, it provides an explanation for the relation between education and earnings that inter-

prets education as a source of marketable skills. Also, the theory predicts growth patterns for earnings, where earnings increase rapidly in the younger years and then gradually reach a stable level, with growth after entry into the labor market explained by investment in on-the-job training. Empirically, the theory fares well in accounting for variations in earnings among persons, using schooling and time in the labor force (as a proxy for on-the-job training and experience) as the main independent variables (Mincer, 1974).

The economic theory also emphasizes supply in accounting for other market processes. Most importantly, unemployment is seen as mostly voluntary, except in certain population groups (youngsters, blacks) where minimum wage laws make it impossible for employers to pay the market wage.

The focus in human capital theory on the supply side—that is, on characteristics of persons—reflects the job structure assumed in the theory—that is, one of a competitive and perfectly functioning labor market. To distinguish the neoclassical theory of the earnings determination process from the alternative model of the matching process that will be formulated later in the chapter, we will refer to the neoclassical theory as the *wage competition* model (following Thurow, 1975) to emphasize the focus on competition among employees for wages. . . .

A competitive labor market that determines wage rates is one where employers make wage offers and workers bid for employment on the basis of their productivity. The match is made when the value of the marginal product demanded equals the wage rate of the employee. This presupposes that employees paid more than their value can be replaced by others who are willing to work at the wage rate that equals marginal productivity, whereas employees who are paid less than their value can get access to jobs where the wage rate reflects their productivity. Only when the employment

relationship is completely open will such a clearing of the market through wage rates be possible. Closed employment relationships, where new recruits can only get access if the incumbent leaves, insulate incumbents from competition. Employers cannot resolve discrepancies between productivity and wage rates by threatening to replace or actually replacing the current employee by someone who is more productive at the same wage rate or who is willing to work at a lower wage rate.

It could be argued that the existence of closed employment relationships does not prevent the employer from relating wages to performance, even in the absence of the ability to replace an employee. Most importantly, the employer can use promotion schemes to reward performance and in this way obtain efficient production. This is correct. Our argument is not that closed employment relationships necessarily prevent efficient production, but that promotion systems represent very different mechanisms for relating wages to performance than the use of competition among employees in open employment relationships where employers make wage offers and employees bid for employment on the basis of their productivity. Promotions can take place only when there is a vacancy in a higher level job and are meaningless as rewards for performance unless jobs at different levels provide different wages, so that wages become attributes of jobs rather than of people. Although a firm with closed employment relationships may operate efficiently because of the overall match between job assignments and performance of employees, the wages for individual employees will reflect the jobs they hold and therefore, not only their performance, but also the rate at which vacancies appear, the organization of jobs, and the seniority of employees. A very different labor market structure exists from the one assumed in the neoclassical theory when wages are tied

to jobs and not to individual variations in performance. . . .

Vacancy Competition

When employees have control over access to the job, others can only get access to the job when incumbents leave. Hence, a vacancy must exist for a person to get access to a job. We will refer to the resulting matching process as *vacancy competition*. We do not wish to argue that this is the only alternative matching process to the wage competition model described by neoclassical economics. At least one other alternative employment relationship can be identified: This is the often met arrangement when employees are directly involved in the disposition of goods to the market, and the “salesperson” is paid some fraction of total earnings. But such relationships presuppose that jobs are not highly interdependent and that the salesperson is primarily involved in the disposition, rather than in the production, of goods.² Vacancy competition in contrast is likely to emerge in closed employment relationships where jobs are interdependent in a technical and social division of labor around production.

In vacancy competition, as in wage competition, employers are assumed to be concerned about hiring the most productive employee at the least cost. But because of the indeterminate length of the employment relationship and the lack of competition among employees over wages, it will not be possible for the employer to link marginal productivity to the wage rate. This has important consequences for (a) the determination of who should be hired; (b) the determination of earnings; and (c) the organization of jobs in job ladders. These consequences all follow from the employer’s attempt to secure the highest possible return from production when faced with employee control over the job.

In wage competition, the employer can rely on the wage rate as a measure of a per-

son's productive capacity. The employer need only be concerned that the value of marginal productivity equals the wage rate and can be indifferent to the relationship between personal characteristics of employees and their performances. In contrast, in vacancy competition, the employer should be very much concerned about the relationship between personal characteristics and productive capacity, because once hired the employee cannot be easily dismissed. Furthermore, it is a person's potential performance that will be of concern, including the person's ability to fulfill the training requirements of jobs. Previous experience, education, and such ascriptive characteristics as race and sex will be used as indicators of potential performance; the main requirements are that the indicators chosen are visible and in the employer's experience show some relationship to performance. Based on the information provided by these indicators, the employer will hire the most promising candidate among those available for a job. In other words, access to a vacancy will be determined by a ranking of job candidates. As proposed by Thurow (1975), the situation may be conceived of as one where a queue of job candidates is established for vacant jobs. A person's position in the labor queue will be determined, not by his or her absolute level of productive capacity, but by the rank order in relation to other job candidates according to characteristics deemed relevant by employers.

As there is a queue of persons for jobs, there will be a rank order or a queue of vacant jobs, where the rank order is established by the earnings provided by vacant jobs, the career trajectories they imply, and such other characteristics as status, pleasantness, and convenience. The matching process, then, is a matching of the queue of persons to the queue of vacant jobs. The highest placed person in the labor queue will get the best job in the job queue. Changes in the supply of persons with certain characteristics (say a change in the distribution of education) and changes

in the availability of jobs at different levels of rewards will change the rank orderings. As a result, whenever there is a change in the labor and job queues, persons with similar characteristics will tend to be hired into different jobs and persons in similar jobs may have different personal characteristics. The organization of jobs into career trajectories (discussed later) will further reinforce these tendencies.

Wage rates in vacancy competition are characteristics of jobs, not of persons. Because employers have no effective way of enforcing a translation of productivity variations into wage rates other than by promotions, wages will tend to become heavily influenced by such institutional forces as collective bargaining and employee desire to preserve traditional relative wage differentials. Internally, wage differentials will reflect the organization of jobs into job ladders.

The creation of job ladders in internal labor markets is, as already mentioned, a way for the employer to create an incentive structure in the absence of open employment relationships. The organization of jobs into promotion schedules further acts as a screening device, inducing low-performance employees to leave on their own decision by denying or delaying promotion in relation to other employees. To be effective, jobs at the same level in a promotion schedule should provide identical earnings, whereas jobs at different levels should provide a differential large enough to induce employees to compete for promotion opportunities. This further reinforces the tendency in vacancy competition for earnings to become a characteristic of jobs so that similar jobs provide similar earnings regardless of characteristics of the incumbents.

Actual promotion opportunities are created when persons leave the firm or a new job is added, setting in motion chains of vacancies (White, 1970). The number of job levels, the distribution of jobs at various levels, the seniority distribution of employees,

and the demand for products influencing the creation of new jobs (or the elimination of jobs), all interact to produce promotion schedules governing the careers of employees. These promotion schedules will under certain conditions result in career lives that are similar to those predicted by human capital theory, even though the mechanisms are quite different (Sørensen, 1977).

In wage competition, employees can change their earnings only by changing their performance. In vacancy competition, changes in earnings are generated by moves in mobility regimes that are chains of vacancies in internal labor markets. There is, in vacancy competition, no automatic correspondence between the creation of promotion opportunities and whatever changes take place in a person's productive capacity. Employees may be promoted without a preceding change in productivity, and a change in productive capacity need not result in a promotion. This means that the cross-sectional association between personal characteristics and earnings will be attenuated, even though personal characteristics are crucial for access to jobs. (A formal derivation of this conclusion and an empirical illustration is presented by Wise, 1975.)

In vacancy competition, variations in earnings reflect variations in job characteristics and the organization of jobs in internal labor markets. This is in contrast to the situation in the neoclassical model of wage competition, where the primary source of variation is the variation in personal characteristics that determine a person's productive capacity.

Vacancy competition structures are likely to be similar to the job structures identified as primary jobs (e.g., Doeringer & Piore, 1971). However, the dualist literature has a very descriptive character, and there is also some confusion as to whether the labor market segmentation is a segmentation of jobs or of persons (blacks, poor, and women in the

secondary sector, white skilled workers in the primary sector). The main conclusion derived from this literature is that there are good jobs and bad jobs.

Constraints on Growth in Earnings

The two polar models of the matching process suggest different constraints on a person's ability to increase his or her earnings. In wage competition, earnings directly reflect performance and hence the skills and abilities of a person. Increases in earnings then are obtained by increasing the skill level of a person, and the major constraint on growth in earnings will be limitations on acquiring additional human capital. In wage competition markets, the amount of training that can be provided in jobs will be low, since on-the-job training is a major cause of the emergence of vacancy competition (Thurow, 1975). Hence, the major source of income inequality among persons lies outside the labor market—that is, in the educational and other training institutions that produce skill differentiation.

In vacancy competition sectors, the major constraint on the attainment of income is access to jobs. If no job is available, a person will not be able to obtain earnings. Growth in earnings is produced by the utilization of opportunities for mobility to better jobs, and this opportunity structure, not changes in skills, governs the earnings variations over time. The major source of variation in earnings is then the restriction of access to jobs and the level of derived demand that determines the availability of jobs.

The different constraints on growth in earnings in wage competition and vacancy competition jobs imply that quite different policies will have to be used in an attempt to increase pretransfer earnings of poverty groups. In wage competition sectors, policies aimed at increasing skill levels either through schooling or—for those already

having entered the labor market—through various off-the-job training programs would presumably be effective. In vacancy competition sectors such policies would be quite ineffective since such training would not make jobs available.

The rather limited success of worker training programs suggests that job vacancy competition indeed is predominant in the U.S. economy. More correctly, the fate of such programs suggests that it is indeed difficult to prepare low-skilled workers for jobs that demand high skill levels, since such jobs tend to be vacancy competition jobs.

NOTES

This research was supported in part by funds granted to the Institute for Research on Poverty at the University of Wisconsin by the Office of Economic Opportunity pursuant to the provisions of the Economic Opportunity Act of 1964. The conclusions expressed herein are those of the authors.

1. The definitions are given in paragraph 10 in the section on “Basic Concepts” in *Economy and Society*, Volume 1. “A social relationship . . . will be known as ‘open’ to those on the outside, if . . . participation . . . is . . . not denied to anyone who is inclined to participate and is actually in a position to do so. The relationship will be known as ‘closed’ [if] participation of certain persons is excluded, limited or subject to conditions [Weber, 1947, p. 139].” Weber argues that market relationships are open and gives as an example of a closed relationship the “establishment of rights to and possession of particular jobs on the part of the worker [Weber, 1947, p. 141].” This identification of open relationships with market relationships (for the exchange of labor for wages) and of closed relationships with control over the job by the worker (and the absence of market relationships) will be relied on heavily in this chapter.

2. A similar arrangement accounts for the apparent contradiction of the argument presented here exemplified by the existence of wage competition among faculty at elite universities despite tenure. Here the individual scholar, and not the employer (i.e., the university), disposes himself of the products (articles and other contributions) to a competitive

market and obtains himself the returns from this activity (i.e., prestige in the profession).

REFERENCES

- Averitt, R. T. 1968. *The Dual Economy*. New York: Norton.
- Becker, G. S. 1964. *Human Capital*. New York: National Bureau of Economic Research.
- Bluestone, B. 1970. “The tripartite economy: Labor markets and the working poor.” *Poverty and Human Resources Abstracts* (5 March–April):15–35.
- Cain, G. G. 1976. “The challenge of segmented labor market theories to orthodox theory: A survey.” *Journal of Economic Literature* 14: 1215–1257.
- Doeringer, P. B., and M. J. Piore. 1971. *Internal Labor Markets and Manpower Analysis*. Lexington, Massachusetts: Heath.
- Kerr, C. 1954. “The Balkanization of labor markets.” In E. W. Bakke, P. M. Hauser, G. L. Palmer, C. A. Myers, D. Yoder, and C. Kerr (eds.), *Labor Mobility and Economic Opportunity*. Cambridge, Massachusetts: Technology Press of MIT.
- Marx, K. 1961. *Capital* (Vol. 1–3). Moscow: Foreign Language Press. (Originally published in English, 1887)
- Mincer, J. 1974. *Schooling, Experience and Earnings*. New York: National Bureau of Economic Research.
- O’Connor, J. 1973. *The Fiscal Crisis of the State*. New York: St. Martin’s.
- Sørensen, A. B. 1977. “The structure of inequality and the process of attainment.” *American Sociological Review* 42:965–978.
- Spilerman, S. 1977. “Careers, labor market structure, and socioeconomic achievement.” *American Journal of Sociology* 83:551–593.
- Stolzenberg, R. M. 1975. “Occupations, labor markets and the process of wage attainment.” *American Sociological Review* 40:645–665.
- Thurow, L. C. 1975. *Generating Inequality*. New York: Basic.
- Weber, M. 1947. [*The Theory of Social and Economic Organization*] (A. M. Henderson and T. Parsons, Trans.). New York: Oxford Univ. Press.
- White, H. C. 1970. *Chains of Opportunity: System Models of Mobility in Organizations*. Cambridge, Massachusetts: Harvard Univ. Press.
- Wise, D. A. 1975. “Personal attributes, job performance and probability of promotion.” *Econometrica* 43:913–931.

58 ■ Arne L. Kalleberg

Nonstandard Employment Relations and Labour Market Inequality

Cross-national Patterns

In the past quarter-century, many industrial nations have witnessed a shift away from the standard work arrangement as the normative basis of the employment relationship and an increase in nonstandard work arrangements. This change in the nature of employment relations has made labour markets and organizations more flexible, and has enabled some people and their families to obtain greater flexibility to accomplish their goals. At the same time, this enhanced flexibility often has a ‘dark side’, sometimes leading to greater labour market inequality and excluding some workers from the social protections previously available to them through the standard employment relationship.

Cross-national research is essential for understanding these changes in the nature of employment relations and their impacts on labour markets and workers. For example, comparative research permits an assessment of the role of the state and economic and social institutions (such as laws and regulations governing trade union influence), as well as societal values, norms, and beliefs in shaping patterns of labour market inequality and exclusion.

This chapter discusses some important correlates and consequences associated with these changes occurring in the nature of employment relations in industrial societies. I first summarize some of the main cross-national

patterns in the use of the various types of nonstandard work arrangements, and suggest some of the reasons that nations and organizations have for using them. I then provide an overview of some issues related to labour market inequality associated with the use of these nonstandard employment arrangements in industrial societies; these issues constitute important components of a research agenda needed to understand cross-national differences in employment relations and patterns of labour market inequality and exclusion.

Standard and Nonstandard Work Arrangements

During the ‘long boom’ of the post-war period, many sociologists argued that an evolutionary process governed by a logic of industrialism would lead to a convergence among industrial societies towards a model of ‘pluralistic industrialism’. This model predicted a long-term historical shift towards the dominance of one standard pattern of production relationships,¹ despite a ‘wide and persisting degree of cross-national variation’.² The standard employment relationship (or *emploi total*³) had three main characteristics (see Table 1): (1) a personal relationship between a dominant master or employer and subordinate worker; (2) the work is done full-time; and (3) the employment relationship is generally

Arne L. Kalleberg, “Nonstandard Employment Relations and Labour Market Inequality: Cross-national Patterns,” in *Inequalities of the World*, edited by Göran Therborn, pp. 136–161. Copyright © 2006 by Verso.

Table 1. Characteristics of Standard and Nonstandard Work Arrangements

| <i>Type of work arrangement</i> | <i>Dimension of work arrangements</i> | | | | | | <i>Hours of Work</i> |
|--|---------------------------------------|-------------------------------------|--|---|--------------------------|-----------|----------------------|
| | <i>Who is the de jure employer</i> | <i>Who is the de facto employer</i> | <i>Assumption of continued employment by de jure employer?</i> | <i>Assumption of continued employment by de facto employer?</i> | <i>Who directs work?</i> | | |
| Standard | Organization A | Organization A | Yes | Yes | Organization A | Full-time | |
| Part-time | Organization A | Organization A | Sometimes | Sometimes | Organization A | Part-time | |
| On-call/day labour | Organization A | Organization A | No | No | Organization A | FT or PT | |
| Short-term | | | | | | | |
| Temporary help | Organization A | Organization A | No | No | Organization A | FT or PT | |
| agency | THA agency | Organization A | Sometimes | No | Organization A | FT or PT | |
| Contract company ^a | Contract company | Organization A | Yes | No | Contract company | FT or PT | |
| Independent contracting, self-employment | Self | Client(s) | Yes | No | Self | FT or PT | |

^aContract company employees may have a standard work arrangement with their de jure employer (the contract company), but from the point of view of Organization A their work arrangements are nonstandard.

assumed to continue for a substantial period or indefinitely.⁴ In addition, the work is generally done at the employer's place of business and usually under the employer's direction.

Standard work arrangements were the norm in many industrial nations for much of the twentieth century and were the normative basis of the framework within which labour law, collective bargaining, and social security systems developed.⁵

Changes beginning in the mid-1970s created conditions that led countries, organizations, and workers to adopt nonstandard employment relations that depart from one or more of the essential features of the standard work arrangement (see Table 1). First, some employment relations do not involve a direct relationship between an employer and employee. This occurs when administrative control and responsibility for the employee is 'externalized',⁶ creating 'detached' workers⁷ or 'triangular' employment relations where a worker establishes connections with several employers.⁸ Examples of such nonstandard employment relations include temporary help agency employment (in which the client organization directs the work of the temporary help agency's employee) and contract company employment (in which employees are directed by their contract company employers but often work at the client's site). In these cases, the workers' *de jure* employers differ from the *de facto* employers who use their labour most directly, and the meanings of the terms 'employer' and 'employee' become ambiguous. These terms are socially constructed products of negotiations: for example, the designation of temporary help agencies as employers rather than as employment agencies in the USA was the result of a campaign carried out over four decades by the temporary help agencies and their corporate backers.⁹

Second, some work deviates from the norm of full-time employment. The most common form of such nonstandard work is *regular part-time* employment, which is usually defined as regular wage employment in which the hours

of work are less than 'normal'.¹⁰ The definition of what is considered to be part-time work varies across countries (see below).

Third, some work does not involve the assumption of continued employment. In standard employment relations, the expectation still exists that employment is permanent or at least non-temporary for an indefinite period (even if in reality it is not 'permanent'). By contrast, short-term temporary employment is usually designated as such, as is the case with temporaries who are directly hired by the organization for a short term or on an on-call basis. Fixed-term contracts have been particularly important in countries where employers have a difficult time terminating contracts of indefinite duration, such as France, Germany, Italy, and Spain.¹¹

Finally, some work arrangements do not involve an employer at all, as in the case of self-employment or independent contracting. In these situations, workers direct their own activities and are 'employed' by clients who pay them for services performed or goods provided, not for their labour power.

Nonstandard employment relations are not new. Employers have a choice between organizing work in markets or hierarchies,¹² and there have always been work arrangements that did not fit the model of full-time work: history is replete with examples of peripheral labour forces and flexible labour markets in which work is unstable and temporary.¹³ For instance, in the inside contracting system that existed in the USA in the nineteenth century, management provided machinery and space in the factory, supplied raw material and capital, and arranged for the sale of the product while contractors were responsible for production and hired the workers and paid and supervised them.¹⁴ Indeed, the efficiencies associated with organizing work in standard, hierarchical employment relations and internal labour markets in the post-Second World War period may have been more of an historical irregularity than is the use of nonstandard employment relations.

There are several major reasons why nonstandard work arrangements became attractive to industrial societies and organizations in the past quarter-century, although the causes of the growth of nonstandard work arrangements differ by type of nonstandard work arrangement (e.g., part-time vs. temporary work) as well as by country.

First, the growth in nonstandard work arrangements was due in part to the need for greater labour market flexibility as a result of the economic crises of the post-1975 period. Sluggish economic growth triggered high unemployment that made it clear, especially in Europe, that economies were incapable of generating enough jobs to provide full-time wage employment for all workers.¹⁵ These nations sought greater flexibility to cope with unemployment and labour market rigidities that impeded efficiency in periods of reduced economic growth and lowered productivity rates. Global economic changes increased competition and uncertainty among firms and put pressure on them to push for greater profits and be more flexible in contracting with their employees and responding to consumers. These pressures encouraged them to seek greater flexibility to redeploy labour among tasks or to adjust the size of their workforces through the use of part-time (e.g., the Netherlands) or fixed-term (e.g., Spain) work arrangements.

This type of response to the economic crises of the mid-1970s may be termed dualism,¹⁶ a strategy adopted in countries such as Britain, West Germany, and France. This contrasts with the other major response, which was corporatism, a strategy adopted in countries such as Austria, Sweden, and Norway. Here, the high degree of union power led to institutional arrangements designed to be 'inclusionary' and to involve the interests of major economic groups (especially unions) in the development and carrying out of economic policy. The key actors here are unions and their members, who represent the interests of the working class as a whole.

The growth of nonstandard work arrangements in countries such as Britain, West Germany, and France represented an attempt to exclude groups of people from the protections afforded by the standard employment relationship and to expose them to market forces. This was accomplished by encouraging immigration, as migrant workers were employed to serve as a form of 'industrial reserve army', at least until reliance on immigrants entailed social costs that may outweigh the economic costs; and enlarging the portion of the indigenous population who is subjected to market forces and excluded from the protections of the standard employment relation. Employers and their managements are the key actors in this decentralized response to economic crisis.

A second reason for the adoption of nonstandard work arrangements was the shift—in all industrial societies—from a manufacturing to a service economy. The expansion of part-time work, in particular, accompanied the process of de-industrialization and the expansion of the service economy, which began in most OECD countries during the period 1965–70. In the Scandinavian, and especially the Swedish, case, the expansion of the welfare state, from the late 1960s and on, was accompanied by the growth in female part-time employment.

Third, the growth of nonstandard work arrangements was related to demographic changes in the composition of the labour force, such as the increase in married women workers and older workers, who often preferred the flexibility available through nonstandard work arrangements¹⁷ in order to balance work with family and other non-work activities. The increase in part-time and short-term temporary work arrangements also served to draw into the labour force those workers (such as young and older persons) who might not have otherwise sought standard employment.

The adoption of nonstandard work was facilitated by technological improvements in

Table 2. Percentage of Part-Time, Fixed-Term Temporary, and Self-Employed Persons, by Country

| Country | Percent part-time ^a | | Percent fixed-term temporary ^b | | Percent self-employed ^c | |
|----------------|--------------------------------|------|---|------|------------------------------------|------|
| | 1973 | 1998 | 1983 | 1998 | 1973 | 1993 |
| USA | 15.6 | 18 | — | 3.2 | 6.7 | 7.7 |
| Australia | 11.9 | 25.9 | 15.6 | 26.4 | 9.5 | 12.9 |
| Canada | 9.7 | 18.7 | 7.5 | 8.3 | 6.2 | 8.6 |
| Japan | 13.9 | 23.6 | 10.3 | 10.8 | 14 | 10.3 |
| Austria | 6.4 | 11.5 | — | 7.8 | 11.7 | 6.3 |
| Belgium | 3.8 | 16.3 | 5.4 | 7.8 | 11.2 | 13.3 |
| Denmark | 22.7 | 17 | 12.5 | 10.1 | 9.3 | 7 |
| Finland | 6.7 | 9.7 | 11.3 | 17.7 | 6.5 | 9.5 |
| France | 5.9 | 14.8 | 3.3 | 13.9 | 11.4 | 8.8 |
| Germany | 10.1 | 16.6 | 10 | 12.3 | 9.1 | 7.9 |
| Greece | — | — | 16.2 | 13 | — | — |
| Ireland | 5.1 | 15.2 | 6.1 | 7.7 | 10.1 | 13 |
| Italy | 6.4 | 11.8 | 6.6 | 8.5 | 23.1 | 24.2 |
| Luxembourg | — | — | 3.2 | 2.9 | — | — |
| Netherlands | 16.6 | 30 | 5.8 | 12.7 | 9.2 | 8.7 |
| Norway | 23 | 21 | — | — | 7.8 | 6.2 |
| Portugal | 7.8 | 7.7 | 14.4 | 17.4 | 12.7 | 18.2 |
| Spain | — | 13.5 | 15.7 | 32.9 | 16.3 | 18.7 |
| Sweden | 23.6 | 24.2 | 12 | 12.9 | 4.8 | 8.7 |
| United Kingdom | 16 | 23 | 5.5 | 7.1 | 7.3 | 11.9 |

^aPercentage of total employment: 1973 (ILO and OECD) estimates from Standing (1997), table 3; 1998 estimates from OECD Employment Outlook 1999, table E; US estimates from Bureau of Labor Statistics.

^bPercentage of total employment: 1983 (ILO and OECD) estimates from Standing (1997), table 3; and Campbell and Burgess (2001), table 1; 1998 estimates from Campbell and Burgess (2001), table 1.

^cPercentage of non-agricultural employees: 1973 and 1993 estimates from Standing (1997), table 3.

communication and information systems that made it easier for organizations to specialize their production, assemble temporary workers quickly for projects, and rely more on outside suppliers. Labour laws designed to protect permanent employees also fuelled the growth in nonstandard work by encouraging employers to avoid the mandates and costs associated with these laws.¹⁸

Cross-national Patterns in Nonstandard Employment Relations

There has been a general tendency for all industrial countries to expand their use of one or more types of nonstandard work arrangements and to shift away from the standard employment relationship as the *normative* model of employment that is the basis for labour law and other employment protections.¹⁹ Within

this general tendency, however, countries differ in the extent to which they use the various types of nonstandard work arrangements. This suggests that employers' strategies depend on institutional contexts such as worker control, institutional conditions related to skills and trust, and legal and regulatory systems that affect the relative costs of using standard employment relations, as well as different forms of flexibility such as part-time or temporary work. The extent to which a country uses nonstandard work arrangements and who and how many persons are excluded from the standard employment relationship depends on the institutional features of the society as well as its unique cultural and social characteristics.

Table 2 presents information on country differences in the use of part-time work, temporary (short-term) contracts, and self-employment. It should be noted that these

comparisons are often hampered by imprecise measures and sometimes inconsistent definitions of nonstandard work arrangements across countries.

Part-time Work

The definition of what constitutes part-time work varies among countries. In the USA, part-time work is generally defined as less than 35 hours a week, while Canada and the United Kingdom normally use 30 hours as the cut-off for part-time.²⁰ In France, part-time is defined as at least 20 per cent below the statutory level of working hours (which became 35 hours on 1 January 2000), while in Germany it is less than 36 hours of work per week.²¹ By contrast, part-time employment in Japan is explicitly related to status within the firm and not to hours worked; indeed, recent Japanese surveys indicate that 20–30 per cent of those classified by their employers as ‘part-time’ actually work as many hours as ‘full-time’ workers.²² Other examples of less-than-full-time employment are various forms of short-time work such as zero-hours contracts in Britain or relay contracts in France, Italy, and Spain, where workers are hired to work a few hours a day to perform work previously done by older workers nearing retirement.²³

Nearly one in five workers in the USA currently works part-time, making it by far the most widely used form of nonstandard work in that country. Part-time work in Europe is slightly lower than in the USA, with an average of about 16 per cent of the European Union’s total labour force working part-time in 1996.²⁴ There is, however, wide variation among countries (see Table 2). In the Netherlands, about a third of the labour force—and over half of women—work part-time, leading a prominent economist²⁵ to characterize it as the ‘only part-time economy of the world, with a finger in the dike of unemployment’. The Netherlands illustrates a case where part-time employment was used to combat (successfully) high levels of unemployment.

Part-time employment is also relatively high in Scandinavia (with the exception of Finland), with about 20 per cent of the labour force working part-time. By contrast, part-time work constitutes relatively small proportions of the labour force in Portugal²⁶ and Finland. Unlike in the USA, part-time employment is increasing relatively rapidly in a number of countries in Europe, where it has often been used as a way to alleviate unemployment and is the major source of employment growth since the 1980s.²⁷ Again, there is wide variation among countries: part-time employment has increased particularly rapidly in the Netherlands, Belgium, France, and Ireland.

In all industrial countries, most part-time workers are women.²⁸ This reflects in part women’s greater responsibilities for housework and raising children. Women’s share of part-time employment in the USA (about 65 per cent) is much lower than in Germany and France (about 90 per cent) and in the UK and Japan (about 80 per cent). The lower proportion of women part-timers in the USA may reflect the greater incidence of part-time work among students, both male and female.

The growth in part-time work in the USA since 1979 appears to have been due to the expansion of industries that typically employ many part-timers (services, retail trade, finance, insurance, real estate) rather than to the substitution of part-time for full-time workers within industries, which occurred mainly in the 1970s.²⁹ Similarly, the growth of part-time employment has accompanied the expansion of the service sector in other industrial countries.³⁰

Short-term Employment

The number of people on fixed-term temporary contracts has increased in the European Union by some 25 per cent in the past decade, though this still represents only about 12 per cent of employees and there is significant country variation in this (see Table 2).³¹

Temporary work in Europe has generally grown less than part-time work and plays a lesser role in the overall labour market. Unlike part-time work, temporary employment does not seem to mitigate unemployment, being rather an indicator of weak worker labour market position (especially among youth) in periods of high unemployment.³²

Fixed-term contracts have been particularly important in countries where employers have difficulty terminating contracts of indefinite duration, suggesting that labour market rigidities can lead to the greater use of temporary workers. For example, the proportion of fixed-term contracts more than doubled between 1983 and 1998 in Spain and France (which had 33 and 14 per cent fixed-term contracts in 1998, respectively), two countries that have strong restrictions on dismissals of regular workers;³³ these highly protective labour laws may have made employers more adept at getting around them by using short-term hires. . . .

Holmlund and Storrie³⁴ maintain that there has been a 'remarkable' increase in fixed-term employment in Sweden in the 1990s, from 10 per cent in early 1990s to 16 per cent at the end of the century. This increase in fixed-term employment was accompanied by a large increase in unemployment. They argue that the growth in fixed-term employment in Sweden reflects mainly the greater incentives that employers have to offer temporary rather than permanent contracts during recessions, and the generally greater willingness of workers to accept temporary jobs when the labour market is depressed. They show that labour demand was a more valid explanation of the rise in temporary employment in Sweden during the 1990s than were accounts based on changes in regulation or the preferences of workers.

On the other hand, some countries (such as Greece and Luxembourg) experienced a decrease in fixed-term employment during the 1983–98 period (see Table 2). There is a

lack of good data on the incidence and trends in short-term employment in the USA. A rough estimate is that 3.3 million workers (a little more than 3 per cent of the employed) were direct-hire temporaries in 1995.³⁵ . . .

Temporary Help Agencies

Employment in the temporary help services industry in the USA has experienced explosive growth since the early 1970s, increasing at an annual rate of over 11 per cent, and its share of total US employment rose from under 0.3 per cent in 1972 to nearly 2.5 per cent in 1998. By contrast, total non-farm employment grew at an annual rate of 2 per cent during this period.³⁶ Changes in temporary work exhibit much greater variance than other forms of employment and are very sensitive to the business cycle, rising and falling with the state of the economy.³⁷

Temporary help agencies are also growing rapidly in Europe,³⁸ although there is considerable variation among countries. Some nations feel that temporary help agencies are useful as employment intermediaries, while others object to them for reasons such as the principle that job placement should be done by public, not profit-making, agencies. In the mid-1980s, temporary help agencies were authorized (subject to some restrictions) in Argentina, Belgium, Brazil, Denmark, France, Germany, Ireland, Japan, the Netherlands, Norway, Portugal, and Switzerland, among other countries. They were banned in Algeria, Costa Rica, Gabon, Greece, Italy, the Libyan Arab Jamahiriya, Madagascar, Mauritania, Senegal, Spain, Sweden, and Zaire.³⁹ In general, though, the restrictions on the use of temporary help agencies have tended to be eased in most countries.⁴⁰ The growth of temporary help agency employment in both the USA and Europe has been generally driven by employers' needs,⁴¹ as well as the entrepreneurial efforts of temporary help agencies themselves.⁴² Moreover, it is likely that tem-

porary staffing agencies will increasingly replace the direct hire of temporaries (especially for low-skilled work⁴³).

Contracting Out

Contract companies, in contrast to temporary help agencies, supervise their employees' work, although this distinction may not always be clear-cut. Until the mid-1980s or so, subcontractors were independent businesses that provided a product or service; since then, subcontractors increasingly provide employees as well.⁴⁴ While contract work has always existed in some industries, such as construction, there has been an increase in the purchase of services (especially business services and engineering and management services) by US firms since the 1970s.⁴⁵ As firms began to contract out services that were previously done in house, they gained a greater appreciation of the variety of services (such as advertising, consumer credit reporting and collection, mailing and stenography, maintenance and cleaning, personnel supply, computer and data processing, protection, research and development) that could be contracted and realized that business service organizations could often supply these services more cheaply and efficiently.⁴⁶

The trend towards greater subcontracting also characterizes all major west European countries.⁴⁷ France and Italy are two countries in which there are strong traditions of small-scale enterprise, and in which there are lots of opportunities for subcontracting.⁴⁸ In the United Kingdom, outsourcing grew substantially in the 1990s after agencies in the public sector were required in the 1980s to go through a process of competitive tendering for catering, laundering, and domestic services.⁴⁹ Benson and Ieronimo⁵⁰ compared the outsourcing of maintenance work between Australian-owned firms and Japanese manufacturing firms operating in Australia. They found that the Australian-owned firms sought to improve their flexi-

bility via externalization (subcontracting), while Japanese firms sought to enhance flexibility via internalization strategies. Industrial relations issues such as labour rigidity, restrictive work practices, and demarcation disputes were central to these firms' outsourcing decisions.

Independent Contractors

Independent contractors are self-employed in that they have neither an employer nor a wage contract and are responsible for their own tax arrangements. However, not all self-employed persons consider themselves to be independent contractors, and this makes it difficult to draw empirical conclusions as to trends in this form of nonstandard work. For example, small shopkeepers who work at a fixed location are not likely to call themselves independent contractors. Nevertheless, information on self-employment is the best indicator of the presence of independent contractors. These data suggest that there has been relatively little change in the proportion of the US labour force that is self-employed since 1973.⁵¹ Self-employment growth also correlates only modestly with aggregate employment growth.⁵²

The trends in self-employment in Europe are diverse. There has been a decline in self-employment in some countries (for example, Austria, Denmark, France, Germany, and Norway). In other countries (such as Finland, Ireland, Portugal, Sweden, and the United Kingdom), there has been an increase in self-employment.

Consequences of Nonstandard Work Arrangements for Labour Market Outcomes

The use of nonstandard work arrangements in industrial societies has important implications for labour market phenomena such as the degree of employment security, the quality of jobs, and opportunities for training and mobility. In this section, I provide an

overview of some of the implications of the growth of nonstandard work arrangements for labour market patterns of inequality and exclusion.

Increase in Job Insecurity?

The shift away from the standard employment relation has led to a general increase in job insecurity with the employer. Almost all workers—those in standard as well as non-standard employment relations—face greater job insecurity than in the past, due to increases in the practice of downsizing, layoffs, and other expressions of employers' willingness to treat labour as a variable cost of production. This increase in insecurity characterizes all industrial societies, and there is a convergence among nations in a shift away from a 'full employment' society characterized by standard employment toward a 'risk society' in which workers are increasingly called upon to bear the risks of their own employment.⁵³

The media as well as some social scientists often exaggerate the view that the standard employment relationship has given way to nonstandard work. For example, objective measures of job stability such as job tenure show little overall change over time. An analysis of European countries (as well as Japan and the USA)⁵⁴ found little if any decline in job stability (as measured by average job tenure) in the last two decades: there was little change for men and an improvement for women in average job tenure, resulting in a general stability or slight increase of average tenure in almost all countries examined (except for Ireland). At the same time, Auer and Cazes found that there was an increase in perceived job insecurity on the part of workers, along with the clear increase in perception of this by the media and many observers, accompanying the small (if any) decline of job stability. They suggest that this paradox of objective stability and perceived insecurity might reflect the continuation of labour market segmentation: that is,

the standard employment relationship still constitutes the core of the employment relationship, although the periphery has grown, creating the perception of greater insecurity for all.

The consequences of this greater insecurity for workers may differ from one country to another, depending on factors such as the welfare regimes available to provide workers with 'safety nets' to reduce the impacts of unemployment, underemployment, and unstable employment.

Growth of Bad Jobs?

In the USA, every nonstandard work arrangement is more likely to be associated with lower fringe benefits than standard work arrangements, and sometimes with lower wages. In particular, male and female temporary help agency employees, on-call workers and day labourers, and part-time workers, are consistently more likely than workers in the other nonstandard arrangements to have low pay and to lack insurance and pension benefits (an exception is some women who are self-employed, who are the most likely group of women to have jobs with low wages). Moreover, most workers in these arrangements (especially temporary help agency employees, on-call workers, and day labourers) prefer standard, full-time employment.⁵⁵ In view of these findings, the explosive growth of the temporary help industry, in particular, makes the strong negative effect of employment in temporary help agencies on job quality a matter of concern.

Part-time work in particular is generally low-paid and low-status (such as sales, catering, and cleaning) in the USA. This is also true in Europe, although there is also some growth of higher-level part-time jobs in some countries. An analysis by the OECD⁵⁶ found that the median hourly earnings of part-time workers were lower than those of full-timers for all countries for which data were available. Differences among countries

in labour law and employment regulations such as job security entitlements affect the extent to which part-time work is of significantly lower quality than full-time employment. Thus, countries differ in whether the use of part-timers represents a marginalization strategy that provides employers with a source of cheap labour or an integration strategy used to retain valued workers. In some countries (such as Sweden, France, Belgium, the Netherlands, Spain), labour law enforces equal treatment between full-time and part-time workers,⁵⁷ preventing the use of part-time workers as a cheap labour source. By contrast, in the UK, Germany, and Japan, part-time employees' hours or income generally fall below thresholds that exclude them from coverage under certain laws.⁵⁸ . . .

An important question when assessing the consequences of nonstandard work arrangements is the extent to which fringe benefits are tied to employment status. Fringe benefits such as health insurance and pensions are distributed on the basis of employment in the USA, and part-time workers are much less likely to receive benefits such as health insurance and retirement benefits, in addition to being paid low wages.⁵⁹ By contrast, these benefits are given out as rights of citizenship in many other countries. So, while working in nonstandard work arrangements might be a source of insecurity, one has a sense that one's needs for health insurance and retirement are taken care of and any disadvantage of part-time work is much less salient.

The growth in nonstandard work arrangements, particularly part-time and short-term temporary jobs, has increased the number of people who are excluded from social protections and subject to low-wage employment. The latter trend coincides with the movement in countries such as the USA, UK, and others towards workfare policies that tie receipt of welfare benefits to employment. In most cases, this

means employment in low-wage, 'contingent', nonstandard jobs, since workfare recipients are often low-skilled workers without a history of stable employment. While their underlying causes may be different, then, the expansion of contingent work is thus consistent with, and mutually reinforcing of, the growth of workfare.⁶⁰ . . .

Polarization of Job Quality and Increased Inequality?

Nonstandard jobs are not all bad: there are also high-quality, high-skilled nonstandard jobs. In the USA, some nonstandard workers—particularly self-employed men and contract-company employees—earn higher wages than regular full-time workers in standard jobs (although the contract company workers are less likely to receive health insurance and retirement benefits). Moreover, relatively few workers in these particular nonstandard arrangements express a preference for standard jobs.⁶¹ Over time, the tasks for which temporary workers are used have changed, and there has been an increase in their use in high-skilled as well as low-skilled jobs.⁶² For example, the wages of temporary help agency employees differ considerably: temporary help agency nurses earn wages that are typically higher than those of nurses who are regular employees of a hospital, while auto supply workers who are employees of temporary agencies typically earn less than regular employees.⁶³

There are also occupational differences in wages as well as fringe benefits within type of work arrangement. In the USA, Kalleberg, Reskin, and Hudson⁶⁴ found that women operatives, sales workers, and service workers—nearly one-third of all women—experience significantly more bad job characteristics than female managers do. Men and women in more complex (and thus more highly skilled and autonomous) occupations are less likely to obtain low wages, and more likely to obtain health insurance and pension benefits from their jobs. Moreover, within

part-time workers, men and women in low-skilled occupations (and low-end service and sales occupations in particular) earn less, and receive fewer health and fringe benefits, than members of other occupations, which supports the idea of a duality within part-time employment.⁶⁵

Careers and Mobility

An important indicator of whether nonstandard jobs are 'good' or 'bad' is the extent to which they are 'dead-end traps' or 'bridges' to more permanent jobs. Moreover, as discussed above, nonstandard work arrangements might be viewed as less problematic than otherwise if they are populated mainly by the young, women, and immigrants, who are likely to have relatively low levels of work commitment and/or low expectations with regard to work.

Mobility opportunities from nonstandard to standard employment are likely to vary both by type of nonstandard work arrangement as well as by country. Countries differ in the extent to which they have inter-firm institutions that are created by unions or occupational associations that enable skill development/training and processes of inclusion between firms; hence, workers who have been excluded from their firms may be included via processes of social closure in inter-firm institutions. An example of these inter-firm institutions is the industrial apprenticeship system found in Germany, which requires high-trust relations such as those maintained through occupational communities.⁶⁶ These occupational institutions provide workers with portable skills that make them employable in a variety of organizations.

The presence of such inter-firm institutions may help explain the results of an OECD study, which showed that fixed-term jobs in Germany are concentrated among first-time job seekers, while in Spain they are spread across a much broader range of workers and firms. A higher proportion of fixed-term workers in Germany obtain permanent positions compared to Spain, where only about 10 per

cent do so. These differences may be due to several things: the tradition of social partnership in Germany (but not in Spain) may encourage employers to regard standard employment relations as a positive asset and see fixed-term contracts as a way of screening workers for permanent positions, while in Spain employers see fixed-term employment as a way of obtaining numerical flexibility. Also, the greater availability of apprenticeship contracts in Germany may lead employers to see them as preferable alternatives to fixed-term contracts. It may also be the case that the potential firing costs of permanent employees remain higher in Spain, though it is unclear if this is as true today as in the late 1980s.⁶⁷

This also suggests that employers in different countries may use temporaries for different reasons, as Casey et al.⁶⁸ showed in their study of temporary workers in Britain and the (then) Federal Republic of Germany. British employers were more likely to use temporaries to obtain numerical flexibility, in order to increase or decrease their workforce size. German employers sought instead to obtain functional flexibility (the ability of employees to do a variety of tasks) by giving temporaries permanent contracts once they acquired the needed skills.

In Italy, fixed-term, temporary workers tend to be concentrated in relatively less-skilled occupations, and people who hold them are relatively young. This suggests that temporary work has been concentrated in segments of the labour force that might prefer transitory employment anyway.⁶⁹

An analysis by the OECD found that the rate of transition out of part-time work into full-time work is much higher for men than women, and more frequent for younger, more highly educated and more highly skilled workers. For France and the Netherlands, part-time workers working very short hours are less likely to move into full-time jobs than other part-timers.⁷⁰

Moreover, the OECD study showed that part-timers received less job-related training

than full-timers. Nonstandard workers are generally less likely to receive job-related training than standard workers, as showed by Hoque and Kirkpatrick's⁷¹ analysis of data from the United Kingdom. They found that managers and professionals on nonstandard (part-time and/or short-term) contracts had less access to training opportunities and were less often consulted about workplace practices than full-time permanent managers and professionals. This suggests that nonstandard employees are treated differently regardless of their occupational level, and that there is not a dualism with regard to training between high and low occupations on nonstandard contracts.

Triangular Employment Relations

Nonstandard work arrangements such as temporary help agencies, leasing companies, and contract companies alter the power dynamics between employers and employees. The separation of the legal (*de jure*) employer from the actual (*de facto*) employer creates a triangular employment relationship or network among organizations. These networks develop differently from one country to another, depending on its economic factors, customs, norms, and regulations. In the USA, the temporary help and contract company industry developed in response to market forces. In other countries (such as many countries in Europe), these triangular relations have been influenced considerably by institutional factors such as union organization and regulations governing these market-mediating institutions.⁷²

Conclusions

The growth in nonstandard work arrangements constitutes an important shift in employment relations that has pervasive consequences for societies, organizations, and workers. While there appears to be a convergence among industrial nations in the use of

nonstandard work arrangements, the consequences of these arrangements for labour market outcomes depend on the institutional context within each country. In this chapter, I have sought to provide an overview of the kinds of nonstandard work arrangements that are increasingly being used in industrial societies, and to indicate some of their implications for patterns of labour market inequality and exclusion.

NOTES

1. Clark Kerr, John T. Dunlop, Frederick H. Harbison and Charles A. Meyers, *Industrialism and Industrial Man*, Cambridge, MA: Harvard University Press, 1960.

2. John H. Goldthorpe, 'The end of convergence: corporatist and dualist tendencies in modern Western societies', in John H. Goldthorpe, ed., *Order and Conflict in Contemporary Capitalism*, Oxford: Oxford University Press, 1984, pp. 315–43 (quote from p. 332).

3. Efrén Córdova, 'From full-time wage employment to atypical employment: a major shift in the evolution of labour relations?', *International Labour Review*, Vol. 125, No. 6, 1986, pp. 641–57.

4. Clyde W. Summers, 'Contingent employment in the United States', *Comparative Labour Law Journal*, Vol. 18, No. 4, 1997, pp. 503–22.

5. Alain Supiot, *Beyond Employment: Changes in Work and the Future of Labour Law in Europe*, Oxford: Oxford University Press, 2001.

6. Jeffrey Pfeffer and James N. Baron, 'Taking the workers back out: recent trends in the structuring of employment', *Research in Organizational Behavior*, Vol. 10, 1988, pp. 257–303.

7. Summers, 'Contingent employment in the United States'.

8. Córdova, 'From full-time wage employment to atypical employment'; A. S. Bronstein, 'Temporary work in Western Europe: threat or complement to permanent employment?', *International Labour Review*, Vol. 130, No. 3, 1991, pp. 291–310.

9. George Gonos, 'The contest over "employer" status in the postwar United States: the case of temporary help firms', *Law and Society Review*, Vol. 31, No. 1, 1997, pp. 81–110.

10. J. E. Thurman and G. Trah, 'Part-time work in international perspective', *International Labour Review*, Vol. 129, No. 1, 1990, pp. 23–40.

11. Córdova, 'From full-time wage employment to atypical employment'.

12. Oliver E. Williamson, 'The organization of work: a comparative institutional assessment', *Journal of Economic Behavior and Organization*, Vol. 1, No. 1, 1980, pp. 5–38.
13. For examples from the USA, see Summers, 'Contingent employment in the United States'.
14. John Buttrick, 'The inside contracting system', *Journal of Economic History*, Vol. 12, No. 3, 1952, pp. 205–21.
15. Córdova, 'From full-time wage employment to atypical employment'.
16. Goldthorpe, 'The end of convergence'.
17. Pfeffer and Baron, 'Taking the workers back out'.
18. Peter Cappelli, Laurie Bassi, Harry Katz, David Knoke, Paul Osterman and Michael Useem, *Change at Work*, New York: Oxford University Press, 1997.
19. Leah F. Vosko, 'Legitimizing the triangular employment relationship: emerging international labour standards from a comparative perspective', *Comparative Labor Law Journal*, Vol. 19, No. 1, 1997, pp. 43–84.
20. Hilda Kahne, 'Part-time work: a hope and a peril', in B. D. Warne, K. P. Lundy and L. A. Lundy, eds, *Working Part-time: Risks and Opportunities*, New York: Praeger, 1992, pp. 295–309.
21. Susan N. Houseman, 'Part-time employment in Europe and Japan', *Journal of Labor Research*, Vol. 16, No. 3, 1995, pp. 249–62.
22. Susan N. Houseman and Machiko Osawa, 'What is the nature of part-time work in the United States and Japan?', in J. O'Reilly and Colette Fagan, eds, *Part-Time Prospects: An International Comparison of Part-Time Work in Europe, North America and the Pacific Rim*, New York: Routledge, 1998, pp. 232–51.
23. Córdova, 'From full-time wage employment to atypical employment'.
24. Colette Fagan, 'Non-standard work arrangements: the UK in European perspective', unpublished paper, Department of Sociology, University of Manchester, 1999.
25. Richard Freeman, 'War of the models: which labour market institutions for the 21st century?', *Labour Economics*, Vol. 5, No. 1, 1998, pp. 1–24.
26. Andries de Grip, Jeroen Hoevenberg and Ed Willems, 'Atypical employment in the European Union', *International Labour Review*, Vol. 136, No. 1, 1997, pp. 49–71; Olga Tregaskis, Chris Brewster, Lesley Mayne and Ariane Hegewisch, 'Flexible working in Europe: the evidence and the implications', *European Journal of Work and Organizational Psychology*, Vol. 7, No. 1, 1998, pp. 61–78.
27. Chris Brewster, Lesley Mayne and Olga Tregaskis, 'Flexible staffing in Europe', *Journal of World Business*, Vol. 32, No. 2, 1997, pp. 133–51.
28. Hans-Peter Blossfeld and Catherine Hakim, eds, *Between Equalization and Marginalization: Women Working Part-Time in Europe and the United States of America*, New York: Oxford University Press, 1997; Colette Fagan and J. O'Reilly, 'Conceptualizing part-time work: the value of an integrated comparative perspective', in J. O'Reilly and Colette Fagan, eds, *Part-Time Prospects: An International Comparison of Part-Time Work in Europe, North America and the Pacific Rim*, New York: Routledge, 1998, pp. 1–31.
29. Thomas Nardone, 'Part-time employment: reasons, demographics, and trends', *Journal of Labor Research*, Vol. 16, No. 3, 1995, pp. 275–92.
30. Houseman, 'Part-time employment in Europe and Japan'.
31. De Grip et al., 'Atypical employment in the European Union'; Klaus Schömann, R. Rogowski and T. Kruppe, *Labour Market Efficiency in the European Union: Employment Protection and Fixed-Term Contracts*, New York: Routledge, 1998.
32. DeGrip et al., 'Atypical employment in the European Union'.
33. Fagan, 'Non-standard work arrangements'.
34. Bertil Holmlund and Donald Storrie, 'Temporary work in turbulent times: the Swedish experience', Munich: Center for Economic Studies and Institute for Economic Research Working Paper No. 671(4), www.CESifo.de, 2002.
35. Anne E. Polivka, 'Are temporary help agency workers substitutes for direct hire temps? Searching for an alternative explanation of growth in the temporary help industry', paper presented at the Society of Labor Economist Conference, 3–4 May, Chicago, 1996.
36. Lewis M. Segal and Daniel G. Sullivan, 'The growth of temporary services work', *Journal of Economic Perspectives*, Vol. 11, No. 2, 1997, pp. 117–36.
37. Lonnie Golden and Eileen Appelbaum, 'What was driving the 1982–88 boom in temporary employment? Preference of workers or decisions and power of employers', *American Journal of Economics and Sociology*, Vol. 51, No. 4, 1992, pp. 473–93; Lewis M. Segal, 'Flexible employment: composition and trends', *Journal of Labor Research*, Vol. 17, No. 4, 1996, pp. 525–42.
38. Bronstein, 'Temporary work in Western Europe'.
39. Córdova, 'From full-time wage employment to atypical employment': 657, note 16.

40. *OECD Employment Outlook*, Paris: Organisation for Economic Co-operation and Development, June 1999.
41. De Grip et al., 'Atypical employment in the European Union'.
42. C. M. Ofstead, 'Temporary help firms as entrepreneurial actors', *Sociological Forum*, Vol. 14, No. 2, 1999, pp. 273–94.
43. See Alan Burton-Jones, *Knowledge Capitalism: Business, Work, and Learning in the New Economy*, Oxford: Oxford University Press, 1999.
44. Jonathan G. Axelrod, 'Who's the boss? Employer leasing and the joint employer relationship', *Labor Lawyer*, Vol. 3, No. 4, 1988, pp. 853–72.
45. Angela Clinton, 'Flexible labor: restructuring the American work force', *Monthly Labor Review*, Vol. 120, No. 8, 1997, pp. 3–17.
46. Katharine G. Abraham and Susan K. Taylor, 'Firms' use of outside contractors: theory and evidence', *Journal of Labor Economics*, Vol. 14, No. 3, 1996, pp. 394–424.
47. Brewster et al., 'Flexible staffing in Europe'; Tregaskis et al., 'Flexible working in Europe'.
48. Goldthorpe, 'The end of convergence'.
49. Gareth Rees and Sarah Fielder, 'The services economy, subcontracting and the new employment relations: contract catering and cleaning', *Work, Employment and Society*, Vol. 6, No. 3, 1992, pp. 347–68.
50. John Benson and Nick Ieronimo, 'Outsourcing decisions: evidence from Australian enterprises', *International Labour Review*, Vol. 135, No. 1, 1996, pp. 59–73.
51. John E. Bregger, 'Measuring self-employment in the United States', *Monthly Labor Review*, Vol. 119, Nos. 1 and 2, 1996, pp. 3–9.
52. Segal, 'Flexible employment'.
53. Ulrich Beck, *The Brave New World of Work* (translated from German by Patrick Camiller), Oxford: Polity Press, 2000.
54. Peter Auer and Sandrine Cazes, 'Stable or unstable jobs: untangling and interpreting the evidence in industrialized countries', paper presented at Conference on Nonstandard Work Arrangements in Japan, Europe, and the United States, W. E. Upjohn Institute for Employment Research, Kalamazoo, MI, 25–26 August.
55. Arne L. Kalleberg, Edith Rasell, Naomi Casriner, Barbara F. Reskin, Ken Hudson, David Webster and Eileen Appelbaum, *Nonstandard Work, Substandard Jobs: Flexible Work Arrangements in the U.S.*, Washington, DC: Economic Policy Institute and Women's Research and Education Institute, 1997.
56. OECE, 'Employment outlook 1999'.
57. Thurman and Trah, 'Part-time work in international perspective'.
58. Houseman, 'Part-time employment in Europe and Japan'; Fagan and O'Reilly, 'Conceptualizing part-time work'.
59. See Arne L. Kalleberg, Barbara F. Reskin, and Ken Hudson, 'Bad jobs in America: standard and nonstandard employment relations and job quality in the United States', *American Sociological Review*, Vol. 65, No. 2, 2000, pp. 256–78.
60. Jamie Peck, *Workfare States*, London: Guilford Press, 2001.
61. See Kalleberg et al., 'Nonstandard work, substandard jobs'.
62. Bronstein, 'Temporary work in Western Europe'.
63. See Susan N. Houseman, Arne L. Kalleberg, and George Erickcek, 'The role of temporary agency employment in tight labor markets', *Industrial and Labor Relations Review*, Vol. 57, No. 1, 2003, pp. 105–27.
64. Kalleberg, Reskin and Hudson, 'Bad jobs in America'.
65. Chris Tilly, *Half a Job: Bad and Good Part-Time Jobs in a Changing Labor Market*, Philadelphia: Temple University Press, 1996.
66. David Marsden, *A Theory of Employment Systems: Micro-Foundations of Societal Diversity*, Oxford: Oxford University Press, 1999.
67. OECD, 'Employment outlook 1999'.
68. Bernard Casey, Rudiger Dragendorf, Walter Heering and Gunnar John, 'Temporary employment in Great Britain and the Federal Republic of Germany', *International Labour Review*, Vol. 128, No. 4, 1989, pp. 449–66.
69. Cebrián et al., 'Atypical work in Italy and Spain'.
70. OECD, 'Employment outlook 1999'.
71. Kim Hoque and Ian Kirkpatrick, 'Dualism in nonstandard employment? Evidence from the 1998 WERS employee questionnaire', paper presented at 17th EGOS Colloquium, Lyon, France, 5–7 July, 2001.
72. See OECD, 'Employment outlook 1999', for a discussion of country differences in the regulation of market-mediating organizations and other forms of nonstandard work.

■ SOCIAL CAPITAL, NETWORKS, AND ATTAINMENT ■

59 ■ Mark S. Granovetter

The Strength of Weak Ties

Most intuitive notions of the “strength” of an interpersonal tie should be satisfied by the following definition: the strength of a tie is a (probably linear) combination of the amount of time, the emotional intensity, the intimacy (mutual confiding), and the reciprocal services which characterize the tie. Each of these is somewhat independent of the other, though the set is obviously highly intracorrelated. Discussion of operational measures of and weights attaching to each of the four elements is postponed to future empirical studies. It is sufficient for the present purpose if most of us can agree, on a rough intuitive basis, whether a given tie is strong, weak, or absent.

Consider, now, any two arbitrarily selected individuals—call them *A* and *B*—and the set, $S = C, D, E, \dots$, of all persons with ties to either *or* both of them. The hypothesis which enables us to relate dyadic ties to larger structures is: the stronger the tie between *A* and *B*, the larger the proportion of individuals in *S* to whom they will *both* be tied, that is, connected by a weak or strong tie. This overlap in their friendship circles is predicted to be least when their tie is absent, most when it is strong, and intermediate when it is weak.

The proposed relationship results, first, from the tendency (by definition) of stronger

ties to involve larger time commitments. If *A-B* and *A-C* ties exist, then the amount of time *C* spends with *B* depends (in part) on the amount *A* spends with *B* and *C*, respectively. (If the events “*A* is with *B*” and “*A* is with *C*” were independent, then the event “*C* is with *A* and *B*” would have probability equal to the product of their probabilities. For example, if *A* and *B* are together 60% of the time, and *A* and *C* 40%, then *C*, *A*, and *B* would be together 24% of the time. Such independence would be less likely after than before *B* and *C* became acquainted.) If *C* and *B* have no relationship, common strong ties to *A* will probably bring them into interaction and generate one. Implicit here is Homans’s idea that “the more frequently persons interact with one another, the stronger their sentiments of friendship for one another are apt to be” (1950, p. 133).

The hypothesis is made plausible also by empirical evidence that the stronger the tie connecting two individuals, the more similar they are, in various ways (Berscheid and Walster 1969, pp. 69–91; Bramel 1969, pp. 9–16; Brown 1965, pp. 71–90; Laumann 1968; Newcomb 1961, chap. 5; Precker 1952). Thus, if strong ties connect *A* to *B* and *A* to *C*, both *C* and *B*, being similar to *A*, are probably similar to one another, increas-

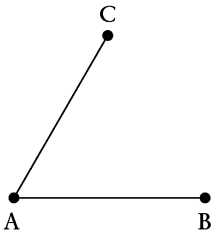


Figure 1. Forbidden Triad

ing the likelihood of a friendship once they have met. Applied in reverse, these two factors—time and similarity—indicate why weaker A - B and A - C ties make a C - B tie less likely than strong ones: C and B are less likely to interact and less likely to be compatible if they do. . . .

To derive implications for large networks of relations, it is necessary to frame the basic hypothesis more precisely. This can be done by investigating the possible triads consisting of strong, weak, or absent ties among A , B , and any arbitrarily chosen friend of either or both (i.e., some member of the set S , described above). A thorough mathematical model would do this in some detail, suggesting probabilities for various types. This analysis becomes rather involved, however, and it is sufficient for my purpose in this paper to say that the triad which is most unlikely to occur, under the hypothesis stated above, is that in which A and B are strongly linked, A has a strong tie to some friend C , but the tie between C and B is absent. This triad is shown in Figure 1. To see the consequences of this assertion, I will exaggerate it in what follows by supposing that the triad shown never occurs—that is, that the B - C tie is always present (whether weak or strong), given the other two strong ties. Whatever results are inferred from this supposition should tend to occur in the degree that the triad in question tends to be absent.

Some evidence exists for this absence. Analyzing 651 sociograms, Davis (1970, p. 845) found that in 90% of them triads consisting of two mutual choices and one non-

choice occurred less than the expected random number of times. If we assume that mutual choice indicates a strong tie, this is strong evidence in the direction of my argument. Newcomb (1961, pp. 160–65) reports that in triads consisting of dyads expressing mutual “high attraction,” the configuration of three strong ties became increasingly frequent as people knew one another longer and better; the frequency of the triad pictured in Figure 1 is not analyzed, but it is implied that processes of cognitive balance tended to eliminate it.

The significance of this triad’s absence can be shown by using the concept of a “bridge”; this is a line in a network which provides the *only* path between two points (Harary, Norman, and Cartwright 1965, p. 198). Since, in general, each person has a great many contacts, a bridge between A and B provides the only route along which information or influence can flow from any contact of A to any contact of B , and, consequently, from anyone connected *indirectly* to A to anyone connected indirectly to B . Thus, in the study of diffusion, we can expect bridges to assume an important role.

Now, if the stipulated triad is absent, it follows that, except under unlikely conditions, *no strong tie is a bridge*. Consider the strong tie A - B : if A has another strong tie to C , then forbidding the triad of Figure 1 implies that a tie exists between C and B , so that the path A - C - B exists between A and B ; hence, A - B is not a bridge. A strong tie can be a bridge, therefore, *only if* neither party to it has any *other* strong ties, unlikely in a social network of any size (though possible in a small group). Weak ties suffer no such restriction, though they are certainly not automatically bridges. What is important, rather, is that all bridges are weak ties.

In large networks it probably happens only rarely, in practice, that a specific tie provides the *only* path between two points. The bridging function may nevertheless be served *locally*. In Figure 2a, for example, the

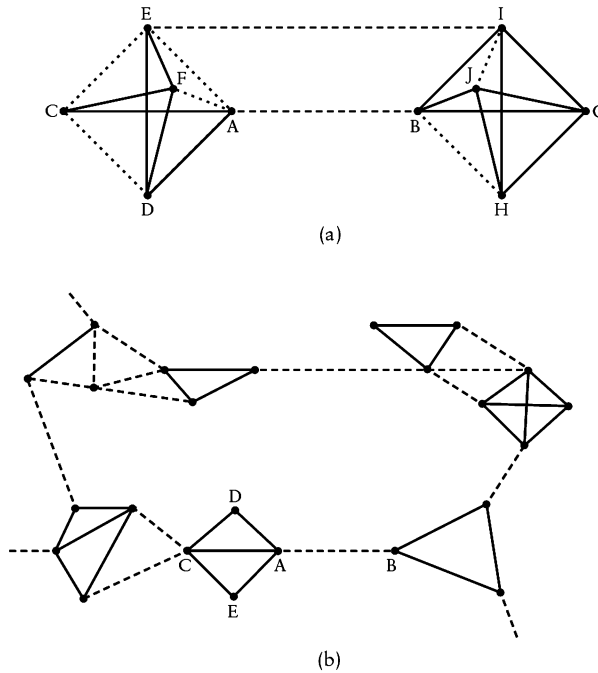


Figure 2. Local Bridges. (a) Degree 3; (b) Degree 13. Straight line = strong tie; dotted line = weak tie.

tie $A-B$ is not strictly a bridge, since one can construct the path $A-E-I-B$ (and others). Yet, $A-B$ is the shortest route to B for F , D , and C . This function is clearer in Figure 2b. Here, $A-B$ is, for C , D , and others, not only a local bridge to B , but, in most real instances of diffusion, a much more likely and efficient path. Harary et al. point out that “there may be a distance [length of path] beyond which it is not feasible for u to communicate with v because of costs or distortions entailed in each act of transmission. If v does not lie within this critical distance, then he will not receive messages originating with u ” (1965, p. 159). I will refer to a tie as a “local bridge of degree n ” if n represents the shortest path between its two points (other than itself), and $n > 2$. In Figure 2a, $A-B$ is a local bridge of degree 3, in 2b, of degree 13. As with bridges in a highway system, a local bridge in a social network will be more significant as a connection between two sectors to the extent that it is the only alternative for many people—

that is, as its degree increases. A bridge in the absolute sense is a local one of infinite degree. By the same logic used above, only weak ties may be local bridges.

Suppose, now, that we adopt Davis’s suggestion that “in interpersonal flows of most any sort the probability that ‘whatever it is’ will flow from person i to person j is (a) directly proportional to the number of all-positive (friendship) paths connecting i and j ; and (b) inversely proportional to the length of such paths” (1969, p. 549). The significance of weak ties, then, would be that those which are local bridges create more, and shorter, paths. Any given tie may, hypothetically, be removed from a network; the number of paths broken and the changes in average path length resulting between arbitrary pairs of points (with some limitation on length of path considered) can then be computed. The contention here is that removal of the average weak tie would do more “damage” to transmission probabilities than would that of the average strong one.

Intuitively speaking, this means that whatever is to be diffused can reach a larger number of people, and traverse greater social distance (i.e., path length), when passed through weak ties rather than strong. If one tells a rumor to all his close friends, and they do likewise, many will hear the rumor a second and third time, since those linked by strong ties tend to share friends. If the motivation to spread the rumor is dampened a bit on each wave of retelling, then the rumor moving through strong ties is much more likely to be limited to a few cliques than that going via weak ones; bridges will not be crossed. . . .

I will develop this point empirically by citing some results from a labor-market study I have recently completed. Labor economists have long been aware that American blue-collar workers find out about new jobs more through personal contacts than by any other method. (Many studies are reviewed by Parnes 1954, chap. 5.) Recent studies suggest that this is also true for those in professional, technical, and managerial positions (Shapiro, Howell, and Tombaugh 1965; Brown 1967; Granovetter 1970). My study of this question laid special emphasis on the nature of the *tie* between the job changer and the contact person who provided the necessary information.

In a random sample of recent professional, technical, and managerial job changers living in a Boston suburb, I asked those who found a new job through contacts how often they *saw* the contact around the time that he passed on job information to them. I will use this as a measure of tie strength. A natural a priori idea is that those with whom one has strong ties are more motivated to help with job information. Opposed to this greater motivation are the structural arguments I have been making: those to whom we are weakly tied are more likely to move in circles different from our own and will thus have access to information different from that which we receive.

I have used the following categories for frequency of contact: often = at least twice a week; occasionally = more than once a year but less than twice a week; rarely = once a year or less. Of those finding a job through contacts, 16.7% reported that they saw their contact often at the time, 55.6% said occasionally, and 27.8% rarely ($N = 54$). The skew is clearly to the weak end of the continuum, suggesting the primacy of structure over motivation.

In many cases, the contact was someone only marginally included in the current network of contacts, such as an old college friend or a former workmate or employer, with whom sporadic contact had been maintained (Granovetter 1970, pp. 76–80). Usually such ties had not even been very strong when first forged. For work-related ties, respondents almost invariably said that they never saw the person in a nonwork context. Chance meetings or mutual friends operated to reactivate such ties. It is remarkable that people receive crucial information from individuals whose very existence they have forgotten. . . .

From the individual's point of view, then, weak ties are an important resource in making possible mobility opportunity. Seen from a more macroscopic vantage, weak ties play a role in effecting social cohesion. When a man changes jobs, he is not only moving from one network of ties to another, but also establishing a link between these. Such a link is often of the same kind which facilitated his own movement. Especially within professional and technical specialties which are well defined and limited in size, this mobility sets up elaborate structures of bridging weak ties between the more coherent clusters that constitute operative networks in particular locations.

REFERENCES

- Berscheid, E., and E. Walster. 1969. *Interpersonal Attraction*. Reading, Mass.: Addison-Wesley.

- Bramel, D. 1969. "Interpersonal Attraction, Hostility and Perception." In *Experimental Social Psychology*, edited by Judson Mills. New York: Macmillan.
- Brown, David. 1967. *The Mobile Professors*. Washington, D.C.: American Council on Education.
- Brown, Roger. 1965. *Social Psychology*. New York: Free Press.
- Davis, James A. 1969. "Social Structures and Cognitive Structures." In R. P. Abelson et al., *Theories of Cognitive Consistency*. Chicago: Rand McNally.
- _____. 1970. "Clustering and Hierarchy in Interpersonal Relations." *American Sociological Review* 35 (October): 843–52.
- Granovetter, M. S. 1970. "Changing Jobs: Channels of Mobility Information in a Suburban Community." Doctoral dissertation, Harvard University.
- Harary, F., R. Norman, and D. Cartwright. 1965. *Structural Models*. New York: Wiley.
- Homans, George. 1950. *The Human Group*. New York: Harcourt, Brace & World.
- Laumann, Edward. 1968. "Interlocking and Radial Friendship Networks: A Cross-sectional Analysis." Mimeographed. Ann Arbor: University of Michigan.
- Newcomb, T. M. 1961. *The Acquaintance Process*. New York: Holt, Rinehart & Winston.
- Parnes, Herbert. 1954. *Research on Labor Mobility*. New York: Social Science Research Council.
- Precker, Joseph. 1952. "Similarity of Valuing as a Factor in Selection of Peers and Near-Authority Figures." *Journal of Abnormal and Social Psychology* 47, suppl. (April): 406–14.
- Shapiro, Albert, Richard Howell, and James Tombaugh. 1965. *The Structure and Dynamics of the Defense R & D Industry*. Menlo Park, Calif.: Stanford Research Institute.

60 ■ Nan Lin

Social Networks and Status Attainment

Status attainment can be understood as a process by which individuals mobilize and invest resources for returns in socioeconomic standings. These resources can be classified into two types: personal resources and social resources. *Personal resources* are possessed by the individual who can use and dispose of them with freedom and without much concern for compensation. *Social resources* are resources accessible through one's direct and indirect ties. The access to and use of these resources are temporary and borrowed. For example, a friend's occupational or authority position, or such positions of this friend's friends, may be ego's social resource. The friend may use his or her position or network to help ego to find a job. These resources are "borrowed" and

useful to achieve ego's certain goal, but they remain the property of the friend or his or her friends.

The theoretical and empirical work for understanding and assessing the status attainment process can be traced to the seminal study reported by Blau and Duncan (1967). Their major conclusion was that, even accounting for both the direct and indirect effects of ascribed status (parental status), achieved status (education and prior occupational status) remained the most important factor accounting for the ultimate attained status. The study thus set the theoretical baseline for further modifications and expansions. All subsequent theoretical revisions and expansions must be evaluated for their contribution to the explanation of status at-

tainment beyond those accounted for by the Blau-Duncan paradigm (Kelley 1990; Smith 1990). Several lines of contributions since then, including the addition of sociopsychological variables (Sewell and Hauser 1975), the recasting of statuses into classes (Wright 1979; Goldthorpe 1980), the incorporation of “structural” entities and positions as both contributing and attained statuses (Baron and Bielby 1980; Kalleberg 1988), and the casting of comparative development or institutions as contingent conditions (Treiman 1970), have significantly amplified rather than altered the original Blau-Duncan conclusion concerning the relative merits of achieved versus ascribed personal resources in status attainment.

In the last three decades, a research tradition has focused on the effects on attained statuses of social resources. The principal proposition is that social resources exert an important and significant effect on attained statuses, beyond that accounted for by personal resources. Systematic investigations of this proposition have included efforts in (1) developing theoretical explanations and hypotheses, (2) developing measurements for social resources, (3) conducting empirical studies verifying the hypotheses, and (4) assessing the relative importance of social resources as compared to personal resources in the process of status attainment. . . .

Contributions of social network analysis to status attainment can be traced to the seminal study conducted by Mark Granovetter (1974), who interviewed 282 professional and managerial men in Newton, Massachusetts. The data suggested that those who used interpersonal channels seemed to land more satisfactory and better (e.g., higher income) jobs. Inferring from this empirical research, substantiated with a review of job search studies, Granovetter proposed (1973) a network theory for information flow. The hypothesis of “the strength of weak ties” was that weaker ties

tend to form bridges that link individuals to other social circles for information not likely to be available in their own circles, and such information should be useful to the individuals.

However, Granovetter never suggested that access to or help from weaker rather than stronger ties would result in better statuses of jobs thus obtained (1995:148). Clues about the linkage between strength of ties and attained statuses came indirectly from a small world study conducted in a tricity metropolitan area in upstate New York (Lin, Dayton, and Greenwald 1978). The task of the participants in the study was to forward packets containing information about certain target persons to others they knew on a first-name basis so that the packets might eventually reach the target persons. The study found that successful chains (those packets successfully forwarded to the targets) involved higher-status intermediaries until the last nodes (dipping down in the hierarchy toward the locations of the targets). Successful chains also implicated nodes that had more extensive social contacts (who claimed more social ties), and yet these tended to forward the packets to someone they had not seen recently (weaker ties). The small world study thus made two contributions. First, it suggested that access to hierarchical positions might be the critical factor in the process of status attainment. Thus, the possible linkage between strength of ties and status attainment might be indirect: The strength of weak ties might lie in their accessing social positions vertically higher in the social hierarchy, which had the advantage in facilitating the instrumental action. Second, the study implicated behavior rather than a paper-and-pencil exercise, as each step in the packet-forwarding process required actual actions from each participant. Thus, the study results lend behavioral validity to those found in previous status attainment paper-pencil studies.

Based on these studies, a theory of social resources has emerged (Lin 1982, 1990). The theory begins with an image of the macro-social structure consisting of positions ranked according to certain normatively valued resources such as wealth, status, and power. This structure has a pyramidal shape in terms of accessibility and control of such resources: The higher the position, the fewer the occupants; and the higher the position, the better the view it has of the structure (especially down below). The pyramidal structure suggests advantages for positions nearer to the top, both in terms of number of occupants (fewer) and accessibility to positions (more). Individuals within these structural constraints and opportunities take actions for expressive and instrumental purposes. For instrumental actions (attaining status in the social structure being one prime example), the better strategy would be for ego to reach toward contacts higher up in the hierarchy. These contacts would be better able to exert influence on positions (e.g., recruiter for a firm) whose actions may benefit ego's interest. This reaching-up process may be facilitated if ego uses weaker ties, because weaker ties are more likely to reach out vertically (presumably upward) rather than horizontally relative to ego's position in the hierarchy.

REFERENCES

- Baron, J. N., and W. T. Bielby. 1980. "Bringing the Firm Back in: Stratification, Segmentation, and the Organization of Work." *American Sociological Review* 45:737-65.
- Blau, P. M., and O. D. Duncan. 1967. *The American Occupational Structure*. New York: Wiley.
- Goldthorpe, J. H. 1980. *Social Mobility and Class Structure in Modern Britain*. New York: Oxford University Press.
- Granovetter, M. 1973. "The Strength of Weak Ties." *American Journal of Sociology* 78:1360-80.
- Granovetter, M. 1974. *Getting a Job*. Cambridge, MA: Harvard University Press.
- Granovetter, M. 1995. *Getting a Job*. Rev. ed. Chicago: University of Chicago Press.
- Kalleberg, A. 1988. "Comparative Perspectives on Work Structures and Inequality." *Annual Review of Sociology* 14:203-25.
- Kelley, J. 1990. "The Failure of a Paradigm: Log-Linear Models of Social Mobility." Pages 319-46, 349-57 in *John H. Goldthorpe: Consensus and Controversy*, edited by J. Clark, C. Modgil, and S. Modgil. London: Falmer.
- Lin, N. 1982. "Social Resources and Instrumental Action." Pages 131-45 in *Social Structure and Network Analysis*, edited by P. V. Marsden and N. Lin. Beverly Hills, CA: Sage Publications.
- Lin, N. 1990. "Social Resources and Social Mobility: A Structural Theory of Status Attainment." Pages 247-71 in *Social Mobility and Social Structure*, edited by R. L. Breiger. New York: Cambridge University Press.
- Lin, N., P. Dayton, and P. Greenwald. 1978. "Analyzing the Instrumental Use of Relations in the Context of Social Structure." *Sociological Methods and Research* 7:149-66.
- Sewell, W. H., and R. M. Hauser. 1975. *Education, Occupation & Earnings: Achievement in the Early Career*. New York: Academic Press.
- Smith, M. R. 1990. "What Is New in New Structuralist Analyses of Earnings?" *American Sociological Review* 55:827-41.
- Treiman, D. J. 1970. "Industrialization and Social Stratification." Pages 207-34 in *Social Stratification: Research and Theory for the 1970s*, edited by E. O. Laumann. Indianapolis: Bobbs-Merrill.
- Wright, E. O. 1979. *Class Structure and Income Determination*. New York: Academic Press.

61 ■ Ronald S. Burt

Structural Holes

Some people enjoy higher incomes than others. Some are promoted faster. Some are leaders on more important projects. The human capital explanation is that inequality results from differences in individual ability. The usual evidence is on general populations, as is Becker's (1975) pioneering analysis of income returns to education, but the argument is widely applied by senior managers to explain who gets to the top of corporate America—managers who make it to the top are smarter or better educated or more experienced. But, while human capital is surely necessary to success, it is useless without the social capital of opportunities in which to apply it.

Social capital can be distinguished in its etiology and consequences from human capital (e.g., Coleman, 1990; Bourdieu and Wacquant, 1992; Burt, 1992; Putnam, 1993; Lin, 1998). With respect to etiology, social capital is a quality created between people, whereas human capital is a quality of individuals. Investments that create social capital are therefore different in fundamental ways from the investments that create human capital (Coleman, 1988, 1990). I focus in this paper on consequences, a focus in network analysis for many years (Breiger, 1995). With respect to consequences, social capital is the contextual complement to human capital. Social capital predicts that returns to intelligence, education, and seniority depend in some part on a person's location in the social structure of a market or hierarchy. While human capital refers to in-

dividual ability, social capital refers to opportunity. Some portion of the value a manager adds to a firm is his or her ability to coordinate other people: identifying opportunities to add value within an organization and getting the right people together to develop the opportunities. Knowing who, when, and how to coordinate is a function of the manager's network of contacts within and beyond the firm. Certain network forms deemed social capital can enhance the manager's ability to identify and develop opportunities. Managers with more social capital get higher returns to their human capital because they are positioned to identify and develop more rewarding opportunities.

The Network Structure of Social Capital

Structural hole theory gives concrete meaning to the concept of social capital. The theory describes how social capital is a function of brokerage opportunities in a network (see Burt, 1992, for detailed discussion). The structural hole argument draws on several lines of network theorizing that emerged in sociology during the 1970s, most notably, Granovetter (1973) on the strength of weak ties, Freeman (1977) on betweenness centrality, Cook and Emerson (1978) on the power of having exclusive exchange partners, and Burt (1980) on the structural autonomy created by network complexity. More generally, sociological ideas elaborated by Simmel (1955) and Merton (1968), on the autonomy

generated by conflicting affiliations, are mixed in the structural hole argument with traditional economic ideas of monopoly power and oligopoly to produce network models of competitive advantage. In a perfect market, one price clears the market. In an imperfect market, there can be multiple prices because disconnections between individuals, holes in the structure of the market, leave some people unaware of the benefits they could offer one another. Certain people are connected to certain others, trusting certain others, obligated to support certain others, dependent on exchange with certain others. Assets get locked into suboptimal exchanges. An individual's position in the structure of these exchanges can be an asset in its own right. That asset is social capital, in essence, a story about location effects in differentiated markets. The structural hole argument defines social capital in terms of the information and control advantages of being the broker in relations between people otherwise disconnected in social structure. The disconnected people stand on opposite sides of a hole in social structure. The structural hole is an opportunity to broker the flow of information between people and control the form of projects that bring together people from opposite sides of the hole.

Information Benefits

The information benefits are access, timing, and referrals. A manager's network provides access to information well beyond what he or she could process alone. It provides that information early, which is an advantage to the manager acting on the information. The network that filters information coming to a manager also directs, concentrates, and legitimates information received by others about the manager. Through referrals, the manager's interests are represented in a positive light, at the right time, and in the right places.

The structure of a network indicates the redundancy of its information benefits.

There are two network indicators of redundancy. The first is cohesion. Cohesive contacts—contacts strongly connected to each other—are likely to have similar information and therefore provide redundant information benefits. Structural equivalence is the second indicator. Equivalent contacts—contacts who link a manager to the same third parties—have the same sources of information and therefore provide redundant information benefits.

Nonredundant contacts offer information benefits that are additive rather than redundant. Structural holes are the gaps between nonredundant contacts (see Burt, 1992: 25–30, on how Granovetter's weak ties generalize to structural holes). The hole is a buffer, like an insulator in an electric circuit. A structural hole between two clusters in a network need not mean that people in the two clusters are unaware of one another. It simply means that they are so focused on their own activities that they have little time to attend to the activities of people in the other cluster. A structural hole indicates that the people on either side of the hole circulate in different flows of information. A manager who spans the structural hole, by having strong relations with contacts on both sides of the hole, has access to both information flows. The more holes spanned, the richer the information benefits of the network.

Figure 1 provides an example. James had a network that spanned one structural hole. The hole is the relatively weak connection between the cluster reached through contacts 1, 2, and 3 and the cluster reached through contacts 4 and 5. Robert took over James's job and expanded the social capital associated with the job. He preserved connection with both clusters in James's network but expanded the network to a more diverse set of contacts. Robert's network, with the addition of three new clusters of people, spans ten structural holes.

Information benefits in this example are enhanced in several ways. The volume is

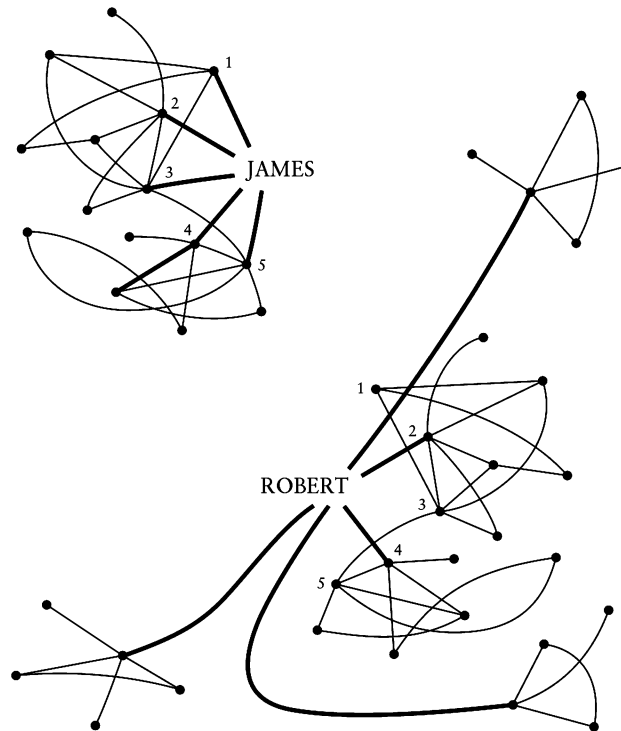


Figure 1. Illustrative Manager's Networks. Thick lines represent a manager's direct contacts.

higher in Robert's network simply because he reaches more people indirectly. Also, the diversity of his contacts means that the quality of his information benefits is higher. Each cluster of contacts is a single source of information because people connected to one another tend to know the same things at about the same time. Nonredundant clusters provide Robert with a broader information screen and, therefore, greater assurance that he will be informed of opportunities and impending disasters (access benefits). Further, since Robert's contacts are only linked through him at the center of the network, he is the first to see new opportunities created by needs in one group that could be served by skills in other groups (timing benefits). He stands at the crossroads of social organization. He has the option of bringing together otherwise disconnected individuals in the network when it would be rewarding. And because Robert's contacts are more di-

verse, he is more likely to be a candidate for inclusion in new opportunities (referral benefits). These benefits are compounded by the fact that having a network that yields such benefits makes Robert more attractive to other people as a contact in their own networks.

Control Benefits

The manager who creates a bridge between otherwise disconnected contacts has a say in whose interests are served by the bridge. The disconnected contacts communicate through the manager, giving the manager an opportunity to adjust his or her image with each contact, which is the structural foundation for managerial robust action (Padgett and Ansell, 1993). Simmel and Merton introduced the sociology of people who derive control benefits from structural holes: The ideal type is the *tertius gaudens* (literally, "the

third who benefits”), a person who benefits from brokering the connection between others (see Burt, 1992: 30–32, for review). As the broker between otherwise disconnected contacts, a manager is an entrepreneur in the literal sense of the word—a person who adds value by brokering the connection between others (Burt, 1992: 34–36; see also Martinelli, 1994). There is a tension here, but not the hostility of combatants. It is merely uncertainty. In the swirling mix of preferences characteristic of social networks, where no demands have absolute authority, the *tertius* negotiates for favorable terms. Structural holes are the setting for *tertius* strategies, and information is the substance. Accurate, ambiguous, or distorted information is strategically moved between contacts by the *tertius*. The information and control benefits reinforce one another at any moment in time and cumulate together over time.

Networks rich in structural holes present opportunities for entrepreneurial behavior. The behaviors by which managers develop these opportunities are many and varied, but the opportunity itself is at all times defined by a hole in the social structure around the manager. In terms of the structural hole argument, networks rich in the entrepreneurial opportunities of structural holes are entrepreneurial networks, and entrepreneurs are people skilled in building the interpersonal bridges that span structural holes.

Predicted Social Capital Effect

Managers with contact networks rich in structural holes know about, have a hand in, and exercise control over the more rewarding opportunities. They monitor information more effectively than it can be monitored bureaucratically. They move information faster, and to more people, than memos. These entrepreneurial managers know the parameters of organization problems early. They are highly mobile relative to people working through a

bureaucracy, easily shifting network time and energy from one solution to another. More in control of their immediate surroundings, entrepreneurial managers tailor solutions to the specific individuals being coordinated, replacing the boiler-plate solutions of formal bureaucracy. There is also the issue of costs: entrepreneurial managers offer inexpensive coordination relative to the bureaucratic alternative. Managers with networks rich in structural holes operate somewhere between the force of corporate authority and the dexterity of markets, building bridges between disconnected parts of the firm where it is valuable to do so. They have more opportunity to add value, are expected to do so, and are accordingly expected to enjoy higher returns to their human capital. The prediction is that in comparisons between otherwise similar people like James and Robert in Figure 1, it is people like Robert who should be more successful.

REFERENCES

- Becker, Gary. 1975. *Human Capital*, 2d ed. Chicago: University of Chicago Press.
- Bourdieu, Pierre, and Loïc J. D. Wacquant. 1992. *An Invitation to Reflexive Sociology*. Chicago: University of Chicago Press.
- Breiger, Ronald L. 1995. “Socioeconomic achievement and social structure.” In *Annual Review of Sociology*, 21: 115–136. Palo Alto, CA: Annual Reviews.
- Burt, Ronald S. 1980. “Autonomy in a social topology.” *American Journal of Sociology*, 85: 892–925.
- _____. 1992. *Structural Holes*. Cambridge, MA: Harvard University Press.
- Coleman, James S. 1988. “Social capital in the creation of human capital.” *American Journal of Sociology*, 94: S95–S120.
- _____. 1990. *Foundations of Social Theory*. Cambridge, MA: Harvard University Press.
- Cook, Karen S., and Richard M. Emerson. 1978. “Power, equity and commitment in exchange networks.” *American Sociological Review*, 43: 712–739.
- Freeman, Linton C. 1977. “A set of measures of centrality based on betweenness.” *Sociometry*, 40: 35–40.

- Granovetter, Mark S. 1973. "The strength of weak ties." *American Journal of Sociology*, 78: 1360–1380.
- Lin, Nan. 1998. *Social Resources and Social Action*. New York: Cambridge University Press (forthcoming).
- Martinelli, Alberto. 1994. "Entrepreneurship and management." In Neil J. Smelser and Richard Swedberg (eds.), *The Handbook of Economic Sociology*: 476–503. Princeton, NJ: Princeton University Press.
- Merton, Robert K. 1968. "Continuities in the theory of reference group behavior." In Robert K. Merton, *Social Theory and Social Structure*: 335–440. New York: Free Press.
- Padgett, John F., and Christopher K. Ansell. 1993. "Robust action and the rise of the Medici, 1400–1434." *American Journal of Sociology*, 98: 1259–1319.
- Putnam, Robert D. 1993. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton, NJ: Princeton University Press.
- Simmel, Georg. 1955. *Conflict and the Web of Group Affiliations*. Trans. by Kurt H. Wolff and Reinhard Bendix. New York: Free Press.

62 ■ Roberto M. Fernandez and Isabel Fernandez-Mateo

Networks, Race, and Hiring

It is common for scholars interested in race and poverty to invoke a lack of access to job networks as a reason why minorities face difficulties in the labor market (e.g., Royster 2003; Wilson 1996). Previous studies on this issue, however, have produced mixed results. Minorities have been found to be *more* likely to have obtained their job through networks than nonminorities (e.g., Elliott 1999). Yet, these jobs pay *less* than jobs obtained by other means (e.g., Falcon 1995). Rather than exclusion from white networks (e.g., Royster 2003), the emphasis in the literature has shifted to minorities' over-reliance on ethnic networks. Thus, the imagery that emerges from these studies is that minorities are stuck in the "wrong networks," that is, those that lead to low-wage jobs.

There is something slippery about the way these network arguments are currently being used, however. Because "wrong network" is defined in terms of the eventual outcome (a network is "good" if it leads to a good out-

come, otherwise, it is a "wrong network"), such explanations run the danger of circular reasoning. To give network accounts of minority underperformance analytical bite, we need to specify the mechanisms by which minorities are "excluded" from productive networks or "stuck" in unproductive ethnic networks.

We argue that being "stuck" in the "wrong network" can be produced by minority underrepresentation in *any* of a number of steps in the recruitment and hiring process. Using unique data from one employer, we illustrate the mechanisms by which minorities can be isolated from good job opportunities. To avoid circular reasoning, we form proper baselines of comparison using data on both networked and nonnetworked minorities and nonminorities at each stage of the hiring pipeline and identify the specific points in the process where network factors could lead minorities to have less access to these desirable jobs than nonminorities.

This is a commissioned chapter that draws heavily on material in a previous publication (Roberto M. Fernandez and Isabel Fernandez-Mateo, "Networks, Race, and Hiring," *American Sociological Review* 71 [2006], pp. 42–71).

Race and Networks in the Labor Market

A common argument in sociology is that jobs found through networks pay better and are of higher status than those found through formal channels (Lin 2002). Evidence on this issue, however, is mixed, especially for minority groups. Reingold (1999) suggests that social networks lead to racial insularity and contribute to the economic marginalization of minorities. However, since many of these studies analyze samples of job incumbents, they often suffer from causal ambiguity (do ties to higher status people cause superior labor market outcomes, or is it that people with superior labor market outcomes gain access to high-status people?). In order to avoid the causal ambiguity problem, a number of studies use samples of job seekers and examine the chances of obtaining employment for various search methods (e.g., Elliott 1999). Employer surveys (e.g., Holzer 1996) are an alternative way of studying this issue, by fleshing out the employer side of the hiring process.

Neither of these approaches, however, examines actual hiring processes and their role in social isolation from good jobs. Without baseline information on the presence or absence of social ties for the pool of competing applicants, some of whom are hired and others are not, it is impossible to identify the effect of social contacts on hiring *per se*. In order to address this issue, some authors have used single-firm screening studies (e.g., Fernandez, Castilla, and Moore 2000; Fernandez and Sosa 2005), but most of this research has not analyzed the role of race in hiring due to lack of appropriate data (a prominent exception is Petersen, Saporta, and Seidel 2000).

“Wrong Networks”

A key component of understanding whether minorities are cut off from employment op-

portunities is to understand why they may be underrepresented in networks that lead to *good* jobs. That is, in order to attribute logically the exclusion of minorities to the absence of network ties to good jobs, we would need to feel confident that minorities might plausibly have been hired except for the lack of the contact. It is critical to define the various processes whereby network factors could limit minorities' access to desirable jobs. In our conceptualization, the “wrong network” account is consistent with underrepresentation of minorities in any of a number of stages in the recruitment and hiring process. Figure 1 represents a conceptual map of the ways in which networks might affect the various stages. These are separated into two sets of processes, referring and screening.

The referral process may contribute to minorities' isolation either if there are no minority employees available to refer in the pool of workers (as in Kasinitz and Rosenberg's [1996] account), or if these employees are reluctant to pass on information about good jobs (see Smith 2005). Even if there are potential referrers who are willing to refer someone, minorities could still be cut off if these referrers were not to refer minority applicants (step 1c.). This could happen if job referral networks are less than perfectly homophilous by race (see Rubineau and Fernandez 2005). If all these conditions are met (steps 1a.–1c.), there will be a set of networked minority candidates in the applicant pool. At this point the screening stage on the demand side of the hiring interface begins.

The effect of screening processes on minorities' access to desirable jobs depends on the employer's attitude towards referrals. If firms prefer to recruit employee referrals (Fernandez, Castilla, and Moore 2000), they will tend to hire minority workers in proportion to their representation in the *networked* pool of applicants. In this case, minorities would thus be cut off from good jobs only if they are underrepresented in the networked applicant pool. However, if employers avoid

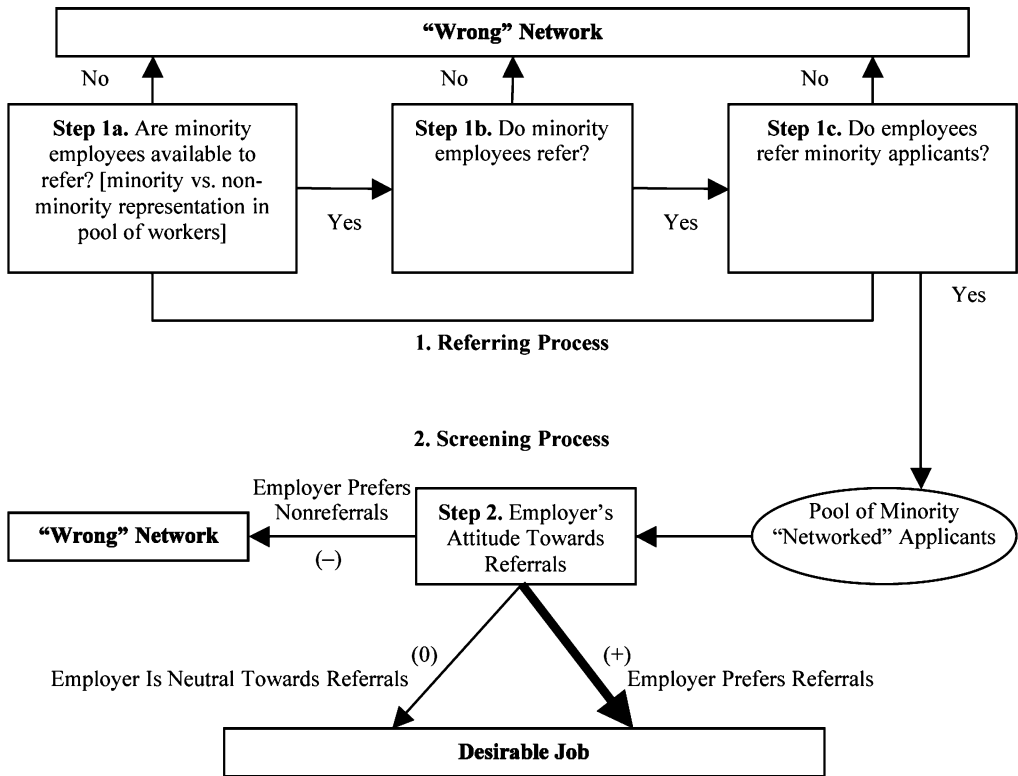


Figure 1. Steps in Network Processes that Connect Candidates to Jobs

hiring through networks (see Ullman 1966), being well represented in the pool of referred applicants is bad news for minorities. Finally, if the employer is neutral toward hiring referrals, access to desirable jobs for minorities will be provided in step 2 (screening) in proportion to the *overall applicant pool*, as opposed to the *networked applicant pool*. In such case the employers' preferences would have no effect in cutting off minorities from desirable jobs.

Analyses

We illustrate these processes using unique data on employees working at one company site and trace their networks of job contacts to applicants for desirable entry-level jobs. Minorities account for 50 percent of employees at this site, with Asian Americans and Hispanics being the largest groups. We

collected all 2,065 paper applications to the plant's entry-level production jobs from September 1997 to November 30, 2000. We coded data on applicants' education, work history, and other human capital characteristics and geocoded candidates to the addresses they listed on the application form. Most importantly, and unusually, we have data on the applicants' race. All applicants must turn in their form in person, and when they do so the receptionist at the company (who is the same one for the whole period) records the applicant's apparent race and gender. Race is thus not self-identified as in most surveys and is well suited to our purposes of understanding the role of race in the screening process.

As mentioned above, in order to ensure that we are not misattributing the effect of networks to the effect of human capital factors, we must show that the candidates for these

Table 1. Percentile Ranks of Company Offered Wages in Wage Distribution of All Non-College Population in Local Metro Area By Racial Group

| | <i>All Racial Groups</i> | | <i>Non-Hispanic White</i> | | <i>African American</i> | |
|--------------------------|--------------------------|---------------|---------------------------|---------------|-------------------------|---------------|
| | <i>Male</i> | <i>Female</i> | <i>Male</i> | <i>Female</i> | <i>Male</i> | <i>Female</i> |
| Starting Wages (N = 192) | | | | | | |
| Hired 9/97–4/98 | | | | | | |
| 8.3% (\$7.75) | 25 | 35 | 18 | 30 | 27 | 37 |
| Hired 5/98–11/00 | | | | | | |
| 91.7% (\$8.05) | 27 | 36 | 19 | 31 | 28 | 38 |
| Wages after 90-Day | | | | | | |
| Probation (N = 109) | | | | | | |
| 15.6% (\$8.05) | 27 | 36 | 19 | 31 | 28 | 38 |
| 62.4% (\$8.50) | 29 | 40 | 21 | 34 | 32 | 41 |
| 21.1% (\$10.05) | 37 | 51 | 27 | 45 | 40 | 48 |
| 0.1% (\$10.35) | 38 | 52 | 28 | 46 | 41 | 48 |
| | <i>Hispanic</i> | | <i>Asian American</i> | | <i>All Minorities</i> | |
| | <i>Male</i> | <i>Female</i> | <i>Male</i> | <i>Female</i> | <i>Male</i> | <i>Female</i> |
| Starting Wages (N = 192) | | | | | | |
| Hired 9/97–4/98 | | | | | | |
| 8.3% (\$7.75) | 35 | 43 | 32 | 39 | 33 | 41 |
| Hired 5/98–11/00 | | | | | | |
| 91.7% (\$8.05) | 36 | 44 | 35 | 41 | 35 | 42 |
| Wages after 90-Day | | | | | | |
| Probation (N = 109) | | | | | | |
| 15.6% (\$8.05) | 36 | 44 | 35 | 41 | 35 | 42 |
| 62.4% (\$8.50) | 39 | 49 | 37 | 45 | 37 | 46 |
| 21.1% (\$10.05) | 49 | 61 | 43 | 55 | 46 | 58 |
| 0.1% (\$10.35) | 50 | 62 | 45 | 57 | 48 | 59 |

Source: Persons in the U.S. Census Bureau's 5 percent 2000 Public Use Micro Sample for the local metro area who are at least 15 years of age and have fewer than 16 years of education with positive wage and salary income in 1999. Data are weighted to reflect the population.

jobs might plausibly have been hired even without network ties. In this setting, it is clear that these entry-level jobs are within reach for people with modest education and labor force experience. The median years of education and experience for the people hired into these jobs are 12 and 7.9, respectively. Moreover, the local labor market was experiencing high rates of unemployment during the period (between 6.76 and 14.1 percent). Data from the local labor market (see Table 1) shows that the wages offered by the firm were attractive—particularly for females and minorities. Starting wages were \$7.75/hr for the first eight months of the study, and \$8.05 afterwards.

These wages fell in the 25th and 27th percentiles of the overall wage distribution for males in the area (35th and 36th percentiles for females). For whites, these starting wages fell in the 18th and 19th percentiles of the male wage distribution, while they were more attractive for minority males (27th and 28th percentiles for African Americans; 35th and 36th for Hispanics; 32nd and 35th for Asian Americans). The pattern is similar for females, although for them the wages are even more attractive than for males in every case. Furthermore, most hires receive pay raises after a 90-day probation period (to \$8.50, \$10.05, or \$10.35), which makes wages even more at-

Table 2. Racial and Gender Distributions of Workers Employed During Hiring Window (September, 1997–November, 2000), and Persons Employed in Metropolitan Area

| | All Plant Employees ^a | | Entry Level Employees Only | | 2000 PUMS ^b | |
|--------------------|----------------------------------|-------------|----------------------------|-------------|------------------------|-------------|
| | <i>Female</i> | <i>Male</i> | <i>Female</i> | <i>Male</i> | <i>Female</i> | <i>Male</i> |
| Non-Hispanic White | 44.0 | 41.5 | 38.5 | 28.6 | 54.3 | 50.4 |
| African American | 3.1 | 5.9 | 3.1 | 5.6 | 5.0 | 4.0 |
| Hispanic | 28.7 | 24.9 | 30.1 | 28.6 | 29.8 | 35.3 |
| Asian American | 23.3 | 26.8 | 27.6 | 35.7 | 6.5 | 6.1 |
| Native American | 0.9 | 1.0 | 0.7 | 1.6 | 0.9 | 0.7 |
| Other, Multirace | — | — | — | — | — | — |
| Total | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |
| Total N | 352 | 205 | 286 | 126 | 174838 | 208174 |

^aEach person is counted equally if they were employed at any time during hiring window (September 1, 1997–November 30, 2000).

^bPersons in the 5 percent 2000 PUMS who are at least 15 years of age and less than 16 years of education with positive wage and salary income in 1999. Data are weighted to reflect the population.

tractive. In fact, the top wage approaches the median of the wage distribution for white and African American women in the area and well exceeds the median for Hispanic and Asian American females.

Are Minority Employees Available to Refer?

In order to study the role of job-referral networks in the application process, we distinguish networked candidates using data from the original application forms. Employee referrals made up 30.2 percent of the applications for which we could identify recruitment source (2,556 out of 2,605). Using the company's employee database, we have been able to link referrers with their referral for 83.7 percent of the "employee referral" applications. For studying who among company employees produces referrals, we use the firm's personnel records. Five hundred fifty-seven workers were employed at the focal plant at any point over the study period, and for these we coded their self-identified race and gender.

We address step 1a in Figure 1 by studying the race and gender distribution of employees who could have produced referral applications during the period. For both genders, over half the workers are minorities (this percentage increases to over 60 percent for

entry-level employees). Relatively large percentages of Asian Americans and Hispanics, but more modest percentages of African Americans, are represented. Women make up 69.4 percent of entry-level workers, and they are overrepresented irrespective of race.

We also compare these race and gender distributions of workers with the composition of persons employed in the local area (Table 2). We find that white workers are underrepresented in this factory compared to their proportion in the area. Asian Americans, however, are extremely overrepresented. While they make up 6.1 percent of employed males in the local labor market, they account for 26.8 percent of male employees. Hispanics are slightly underrepresented (24.9 percent of employees, 35.3 percent in the area), while the percentages of African Americans employed at the factory are quite similar to their proportion in the area (5.9 versus 4.0 percent for males). In sum, minorities are definitely available to refer in this setting.

Do Minority Employees Refer?

Of the 557 employees, 200 of them originated a total of 580 applications. Asian Americans refer the most (50.9 percent of male Asian Americans originated at least one referral applicant). Interestingly, whites show

the lowest rates of producing referrals (27.7 percent for females and 18.8 percent for males). To determine whether background factors might account for the observed race differences in referral behavior, we estimate a set of negative binomial regressions (see Fernandez and Fernandez-Mateo 2006, Table 5). These models show that there are no significant gender differences in the counts of referrals originated by employees at the company. There are race differences, however, as minorities generate more referrals than whites (with Asian Americans producing the most). These effects remain even when adding the extensive controls for individual background mentioned above. Thus, at step 1b in Figure 1, we have no evidence that minorities are less likely than whites to produce referral applicants.

Do Minority Employees Refer Minority Applicants?

The simplest criterion to assess if this is a point of disconnection for minorities is whether there are any minorities at all produced by the referral process, irrespective of the race of the referrer. We analyze the race distribution of applicants produced by referrers of different racial backgrounds. We find that there is a strong relationship between the race of the referrer and the race of the referral applicant (for both genders). Most important, 61.8 percent of male and 57.2 percent of female referred applicants are minorities. Clearly, by this first criterion, minorities are not cut off at step 1c.

A more stringent criterion would be to assess whether the referral process is reproducing the racial distribution of the referring population (the 200 employees identified above). We find that whites are the most insular of the groups in referring (76.9 percent of the referrals produced by white employees are white). The percentage of African Americans and Asian Americans in the referral pool matches the percentage of the referring population quite closely (5.6 vs. 4.0 percent for

African Americans; 34.6 vs. 32.5 percent for Asian Americans). Hispanics, however, constitute a third of referrers, but only 19 percent of referrals. Hispanics, therefore, seem to be somehow cut off, in relative terms, from the networks leading to these jobs.

A final criterion depends on whether minorities refer minority applicants less than they refer white applicants. We find, however, that irrespective of which minority group one considers, the percentage referring whites is lower than the percentage referring minorities. By this final criterion, we find no evidence that a lack of racial homophily in minorities' referring patterns is weakening their access to this company. In sum, there is little evidence that minorities in this setting are cut off from the job networks that lead to employment at this company due to the behavior of the originators of referral networks (the "referring" process).

Hiring Interface

Even if networked minorities are well represented in the application pool, this does not necessarily mean that they will be similarly represented at subsequent stages of the screening process. This step depends on the employer's racial biases and attitudes towards referrals ("screening process" in Figure 1). Our fieldwork and interviews with HR managers at this site suggest that the employer has neither a strong preference nor distaste for referred applicants. We assembled data on all applicants for entry-level jobs and tracked their progress through the hiring pipeline (from application, to interview, to offer, to hire).

For females (see Figure 2a), there is little evidence of a preference for networked candidates (33.9 percent of female applicants and 35.3 percent of female hires were referrals). Also, the race distribution did not change much across stages. Similarly for males (Figure 2b), the percentage of networked applicants does not increase across the various stages of the screening process. If

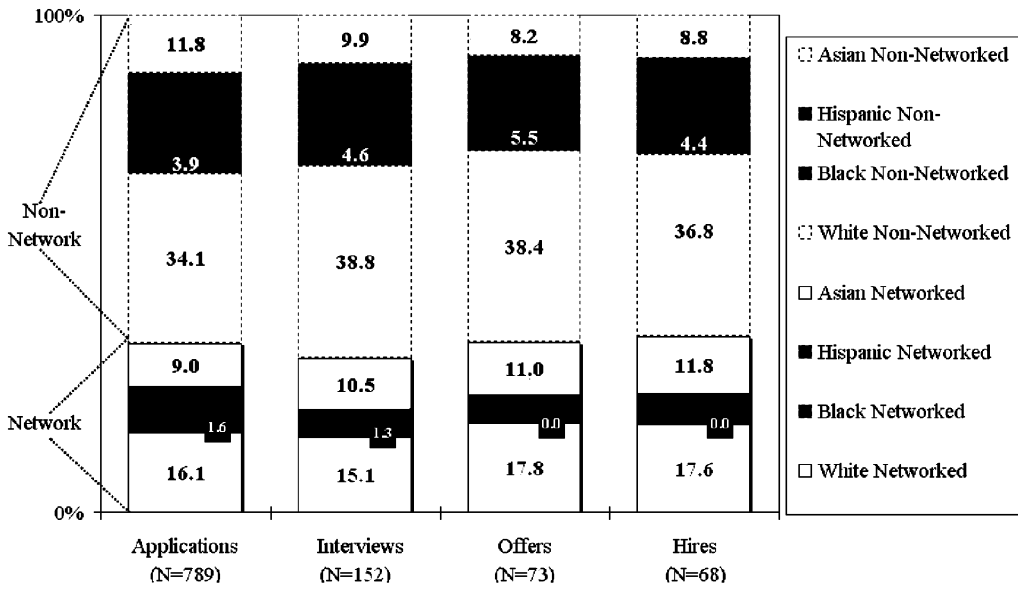


Figure 2a. Female Hiring Interface

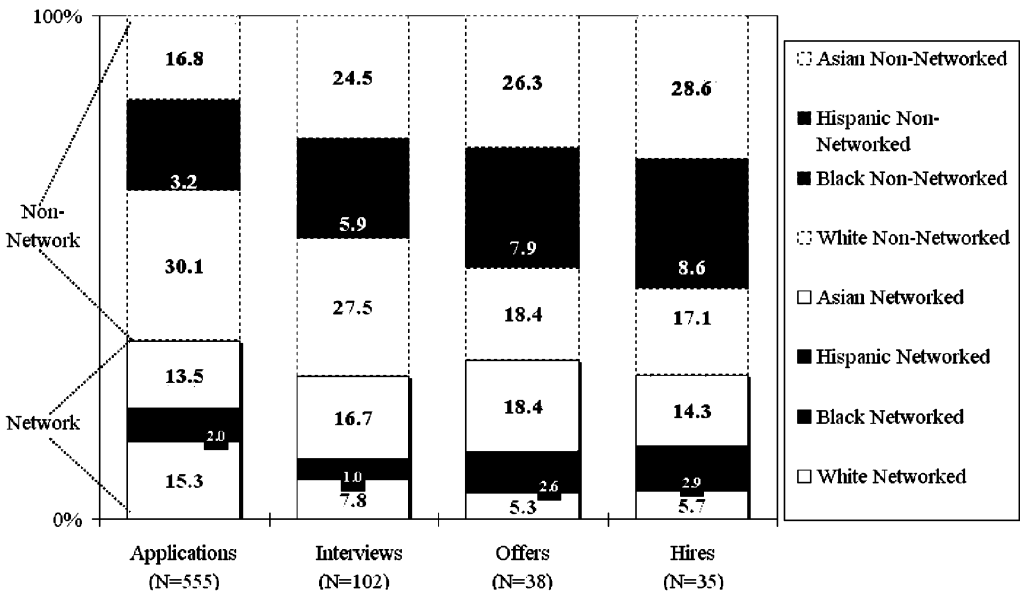


Figure 2b. Male Hiring Interface

anything, it decreases (networked candidates constitute 35.3 percent of the overall male applicant pool and 28.6 percent of hires).

We also performed multivariate (probit) analyses of who is hired versus not hired, in order to introduce controls into the analysis.

For female applicants, both recruitment source and race do not significantly influence the probability of hire. This does not change when human capital and controls are added to the model (although some of these controls such as “years of experience” are significant).

There is thus no evidence of a preference for networked candidates among females, and the race distribution of hires is not different from the distribution of applicants. For males, the results show that all minority racial groups are more likely to be hired than whites—irrespective of whether the person was a referral or not. The only significant pattern is that Asian American nonnetworked males are 8.5 percent more prevalent among hires than among applicants, even after everything else is controlled for.

In sum, the regression results show that there is little evidence of an employer's preference either for, or against, candidates who were referred to the company at the hiring interface. The one exception is the case of Asian American males, where nonnetworked candidates are more likely to be hired compared with networked applicants. We can only speculate whether this reflects a conscious effort to limit the number of Asian Americans—who are overrepresented in the application pool relative to the population in the local labor market. The hiring interface is thus a point of disconnection for Asian American males (i.e., the arrow labeled “(-)” in step 2 of Figure 1), but this employer is otherwise neutral with respect to networked candidates (i.e., the arrow labeled “0” in step 2 of Figure 1).

Summary and Conclusion

This paper makes a number of theoretical contributions to the study of racial inequality and networks. First, it contributes to specifying the mechanisms operating in network accounts of racial inequality in the labor market. Past accounts in this area often run the danger of circular reasoning because “wrong network” is defined in terms of the eventual outcome. We argue that much more analytical precision is needed to specify what it means to be “stuck” in the “wrong network,” as these stories are consis-

tent with minority underrepresentation in *any* of a number of steps in the recruitment and hiring process. Indeed, this study is the first to analyze comprehensively the racial implications of *both* the referring process and the screening mechanisms at the point of hire. We found that network factors operate at several stages of the recruitment process, but we found scant evidence that these factors serve to cut off minorities from employment at this company.

This study also has significant implications for policy. Since policies are often designed to target distinct steps in the recruitment process (National Research Council 2004), understanding each of these steps is crucial for crafting effective policy interventions. Affirmative action, for example, is focused on affecting the behavior of labor market screeners (Reskin 1998). Assessing the effect of these policies will be very difficult without data on both hires and nonhires. Other policy prescriptions recommend that companies open up their recruitment practices by broad advertising and use of formal recruitment systems, on the theory that informal, referral-based recruitment is inherently exclusionary (LoPresto 1986). The results presented here, however, suggest that this heuristic is too simple. Relying on referrals can help reproduce the distribution of the referring population. Therefore, in settings where the current workforce is racially diverse—as is the case in this setting—the referral processes can actually help *perpetuate* diversity (see Rubineau and Fernandez 2005).

Although the “wrong network” account does not fit the facts in this setting, this is not to say that in another, less diverse setting, where referrals may be preferred by screeners, the empirical results would not be markedly different. While we can make no claims of empirical generalizability, however, this study has important methodological implications. Its value is apparent in the light it sheds on mechanisms that are nor-

mally hidden from view. We suggest that the fine-grained processes that we have uncovered here need to be addressed to render the “wrong network” hypothesis testable in other settings. Moreover, it is important to realize that the race and network effects that are often reported in analyses of highly aggregated survey data are likely conflating the effects of the multiple mechanisms that we have delineated. Distinguishing among these steps should be a high priority in future research at the intersection of networks and race in the labor market.

REFERENCES

- Elliott, James R. 1999. “Social Isolation and Labor Market Insulation: Network and Neighborhood Effects on Less-Educated Urban Workers.” *Sociological Quarterly* 40: 199–216.
- Falcon, Luis M. 1995. “Social Networks and Employment for Latinos, Blacks, and Whites.” *New England Journal of Public Policy* 11: 17–28.
- Fernandez, Roberto M., Emilio J. Castilla, and Paul Moore. 2000. “Social Capital at Work: Networks and Employment at a Phone Center.” *American Journal of Sociology* 105: 1288–1356.
- Fernandez, Roberto M., and Isabel Fernandez-Mateo. 2006. “Networks, Race and Hiring.” *American Sociological Review* 71: 42–71.
- Fernandez, Roberto M., and M. Lourdes Sosa. 2005. “Gendering the Job: Networks and Recruitment at a Call Center.” *American Journal of Sociology* 111: 859–904.
- Holzer, Harry J. 1996. *What Employers Want: Job Prospects for Less-Educated Workers*. New York: Russell Sage Foundation.
- Kasinitz, Philip, and Jan Rosenberg. 1996. “Missing the Connection: Social Isolation and Employment on the Brooklyn Waterfront.” *Social Problems* 43: 180–93.
- Lin, Nan. 2002. *Social Capital: A Theory of Social Structure and Action*. New York: Cambridge University Press.
- LoPresto, Robert. 1986. “Recruitment Sources and Techniques.” Pp. 13–1 to 13–26 in *Handbook of Human Resource Administration*, 2nd ed., edited by J. J. Famularo. New York: McGraw-Hill.
- National Research Council. 2004. *Measuring Racial Discrimination*. Panel on Methods for Assessing Discrimination, edited by Rebecca M. Blank, Marilyn Dabady, and Constance F. Citro. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academies Press.
- Petersen, Trond, Ishak Saporta, and Marc-David Seidel. 2000. “Offering a Job: Meritocracy and Social Networks.” *American Journal of Sociology* 106: 763–816.
- Reingold, David A. 1999. “Social Networks and the Employment Problem of the Urban Poor.” *Urban Studies* 36: 1907–32.
- Reskin, Barbara F. 1998. *The Realities of Affirmative Action*. Washington, DC: American Sociological Association.
- Royster, Deirdre A. 2003. *Race and the Invisible Hand: How White Networks Exclude Black Men from Blue-Collar Jobs*. Berkeley: University of California Press.
- Rubineau, Brian, and Roberto M. Fernandez. 2005. “Missing Links: Referral Processes and Job Segregation.” MIT Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA. Unpublished manuscript.
- Smith, Sandra. 2005. “Don’t Put My Name on It’: Social Capital Activation and Job-Finding Assistance Among the Black Urban Poor.” *American Journal of Sociology* 111: 1–57.
- Ullman, Joseph C. 1966. “Employee Referrals: A Prime Tool for Recruiting Workers.” *Personnel* 43: 30–35.
- Wilson, William Julius. 1996. *When Work Disappears: The World of the New Urban Poor*. New York: Alfred A. Knopf.

■ SIBLING MODELS ■

63 ■ Dalton Conley

What Do Low (or High) Sibling Correlations Tell Us About Social Ascription?

How similar are the socioeconomic statuses of siblings and what does this mean for class analysis and theory? This chapter describes the degree of sibling resemblance in a number of components of socioeconomic status. It also documents *when* siblings most resemble each other and *which* siblings resemble each other to the greatest degree and discusses the implications of these results for how we conceptualize the intergenerational transmission of socioeconomic status.

This relationship of family background or socioeconomic status to the outcomes of offspring is one of the most central concerns of stratification research. Indeed, for nearly a century, behavioral geneticists have used sibling correlations as measures of the effects of genes and shared environment on behavioral traits (McGue and Bouchard 1998), and for more than three decades sociologists have used sibling correlations as “omnibus measures” (Solon, Page, and Duncan 2000: 383) of the effects of family and community on socioeconomic outcomes (Hauser, Sheridan, and Warren 1999; Jencks et al. 1972; Solon et al. 1991). Yet, as Hauser and Sewell have stated, “Nowhere has a research agenda of such substantive importance had to survive on

such meager scraps of data” (1976: quoted in Solon et al. 1991: 512). It is not just data that have been lacking. As Becker and Tomes (1986) write, “while sociologists have taken the lead in estimating the impact of family background on a number of outcomes—most notably educational and occupational attainment—they have lacked a clear explication of an underlying behavioral model.” The current chapter tries to address both of these concerns by engaging with economic models of the family and integrating this literature with theories of class inheritance.

What are we seeing when we see a sibling resemblance or sibling discordance model? The often-unspoken assumption is that a high degree of sibling resemblance reflects a more caste-like society, one in which life chances are highly stratified. A low degree of sibling resemblance, by contrast, is taken to mean that a society is less ascriptive and more meritocratic. However, how accurately do sibling correlations (or parent-child correlations or mobility matrices, for that matter) represent the openness of a society and the impact of family background? The short answer is “not very accurately”—that is, once we open up the possibility of within-family

This is an original article that draws heavily on an article published elsewhere (Dalton Conley, “Bringing Sibling Differences In: Enlarging Our Understanding of Advantage in Families,” Chapter 7 in *Social Class: How Does It Work?* by Amette Lareau and Dalton Conley, eds. New York: Russell Sage Foundation, 2008).

heterogeneity in parental investment, sibling effects upon one another, and differing responses to familial investment in households with multiple offspring. While it is trivially true that if there is a zero correlation between siblings (or between parent and child for that matter), family background can be said to have no impact, this is a descriptive account only. For example, it could be the case that family background matters enormously, but that within-family dynamics obscure this fact in sibling or parent-child associations of socioeconomic status.¹

For example, envision the case of a two-child family in which the elder child is expected to sacrifice for the benefit of the younger sibling. If such a dynamic were widespread in a given society and resulted in downward mobility for the sacrificing sibling and upward mobility for the sibling who benefited from the sacrifice, we could actually observe a negative sibling correlation and a zero parent-child correlation (since the upwardly mobile offspring would be cancelled out by the downwardly mobile one, assuming researchers would randomly survey one of the two). Negative assortative mating combined with specific parent-child socialization and investment patterns might also yield a zero (or negative) sibling (or parent-child) correlation when there are, in fact, quite strong background effects. This possibility is depicted graphically in Figure 1 below.

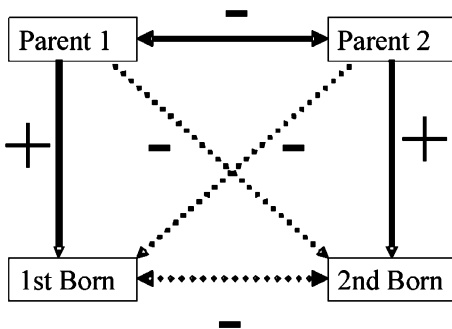


Figure 1. No Observed Family Effects I: Birth Order

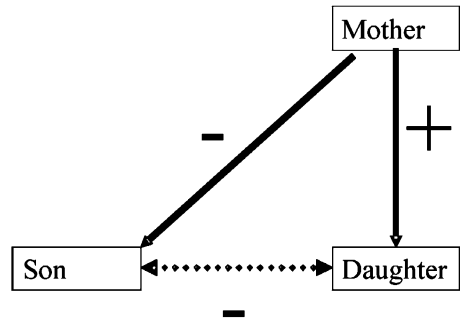


Figure 2. No Observed Family Effects II: Gender

If such dynamics were systematically stratified by a measurable variable such as gender or birth position, then we could accurately describe the intrafamily dynamics by observing correlations for within-family subgroups such as first borns or boys. (See Figure 2, an example with respect to gender in a single parent family.) But if the way that families generated outcomes among children was based on some unobservable factor—such as parental belief in child ability—then to the researcher, the apparent result may be randomness and a potentially faulty observation that family background means little. In fact, what it would mean is that the family would act as a primary queuing mechanism for socioeconomic opportunity. Take the case of height and weight. Figure 3 presents sibling correlations in height, weight, and body mass index (BMI—weight in kilograms divided by height in meters, squared). At just over .5, the correlation for brothers’ height is similar to that typically measured for most measures of socioeconomic status (income, for example). This significant within-family variation on a “relatively” ascriptive characteristic such as height could—in theory—play a salient role in the stratification system. If schools and employers, for example, favored taller individuals, then this factor would generate sibling inequality and appear to attenuate the effect of family background. This might give the false impression of a “meritocratic” society when social ascription was predominant. It would just happen

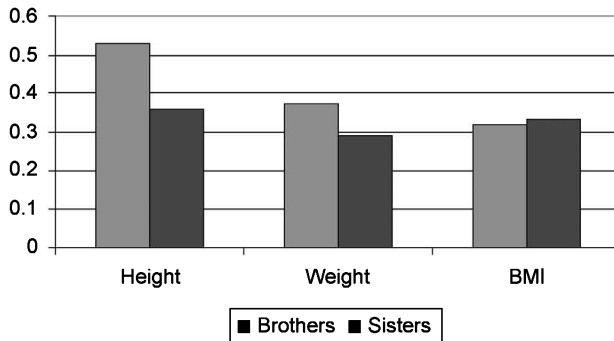


Figure 3. Sibling Correlations in Physical Status: 1999–2001

to be the case that the basis of assignment cuts within—as well as across—families. To take this case further, if the “discrimination” against shorter men took place primarily in the family home as the taller sibling received more parental investment, then this would be an important way that family background influenced life chances but ended up driving *down* the sibling correlation. Likewise, if through differential nutritional investment, the very height difference (or any other difference, even in “ability”) were generated by family dynamics, then sibling differences would reflect (rather than refute) background effects. The bottom line is that in order to make interpretive sense of sibling models, we must have a theory of whether, to what extent, and on what basis of assignment the family itself acts as a queuing mechanism for opportunity.

How to empirically resolve such ambivalent interpretations of sibling or parent-child associations? The best way would be to find an instrumental variable (IV) that affects individual success but that is otherwise unrelated to the success of the other children in the family. If we could identify and measure such a factor, then we might be able to identify its impact on the family distribution of resources. In other words, if an exogenous shock makes one kid thrive (or fail), we can then measure the impact of that shock on parental investment on the other offspring. In other words, when one kid is thriving, do parents invest more in that kid or in her brother? However, this strategy assumes perfect measurement of parental

investment. We would not, however, have identification to examine the effect on actual outcomes, since that discounts the possibility of direct effects of siblings on each other. In fact, this IV strategy has been used in an attempt to identify intersibling reciprocal effects—a dynamic that has been repeatedly hypothesized but has so far been impossible to document. Hauser and Wong (1989)

use friend’s IQ as an IV to identify sibling effects on one another’s IQ. But this assumes that one’s sibling’s friend’s IQ has no other effect on one’s own IQ other than through sibling socialization. To the extent that this assumption is violated—through, for example, socialization between the sibling friend and ego or through differential investment in each child by the parents based on their peer group—then the IV is invalid.

Birth weight might initially seem to provide just such an instrument, given its individual nature and the associations it has been shown to have with socioeconomic outcomes (see, e.g., Conley and Bennett 2000 or Behrman and Rosenzweig 2004). However, parents may invest differentially in their children by estimating the rate of return according to birth weight. So, it is unlikely that there will appear an instrument that satisfies these two criteria and, as a result, a fully specified model is by definition underidentified. (Moreover, if one adds in selective fertility—parity progression bias—based on parental perceptions of whether or not they got a “winner” with the first child or not, then the problem of understanding social transmission within the family gets even more complicated.)

Given the apparent intractability of modeling the true impact of family background(s), what are we to make of sibling correlations? The answer is simply that we can read a sibling correlation as a global effect of family background if we assume a model in which offspring are invested in equally (or at least

that any favoritism is randomly distributed) and in which siblings have only a mean-regressive effect on each other. That is, that they tend to cause each other to be more alike than they would in each other's absence. This is not an entirely unreasonable assumption, but it is an assumption nonetheless. It is for researchers—both ethnographic and quantitative—to determine whether this assumption (and that of nonfavoritism) is accurate or not. What follows is a discussion of how we might indirectly test this assumption and what variations on it might mean for our understanding of how family background impacts individual attainment.

Theoretical Arguments About Sibling Resemblance

Resemblance over the Life Course

In a series of works, Becker and Tomes (1976, 1986) hold that parents either have child-neutral preferences whereby they invest equally in both children, or they have child-equality preferences whereby they invest differentially in children to equalize their ultimate status attainment.

In the former, parents follow an “efficiency” model, that is, they don't care about ultimate differences in outcomes between offspring that result from ability differences—they merely provide the same investments in each and over time innate ability is revealed. With uncertainty as to ability, this approach would maximize overall family attainment. However, from the point of view of a stratification researcher, since the returns to investing in a child with greater innate endowments are greater than those that result from marginal investment in a less well endowed child, parents may reinforce inequality. A more extreme version of this approach would have parents investing more in offspring who “appear” to display the most ability since their return on investment would putatively be highest for these children. And thus, from the point of view of maximizing the total re-

turns from all offspring, this is even more efficient (assuming correct information about ability and returns) than investing equally in all children. However, this parenting strategy would result in an even more inequality-reinforcing result.

In the latter model, the “equalizing” paradigm, parents follow a “wealth” model and invest more in the human capital of their endowed child and make compensatory investments in the nonhuman capital (that is, wealth and income) of their less endowed child. Thus, they try to maximize returns while at the same time acting like a government transfer program, by providing direct material support (as opposed to human capital investment) for the less endowed children. Put crudely, they invest in the schooling of the smart ones and then transfer income or wealth to the dull ones.

A third alternative has parents less concerned with equalizing the ultimate wealth of their children and more concerned with equalizing the earnings of their children. In a series of studies on parental preference models, Behrman, Pollak, and Taubman (1982, 1989) provide a “separable earnings-bequest” model as an alternative to Becker and Tomes's wealth model. Whereas in the wealth model parents are concerned with the combination of earnings and wealth of their children as indicative of their ultimate economic wellbeing, in the separable earnings-bequest model, parents view earnings and wealth as separate factors—and parents attempt to maximize earnings over total income or wealth, since earnings are more socially valued than unearned income. Their analysis of fraternal white male twins shows that parents transfer wealth to their children equally, but transfer resources to their children unequally so as to increase the earnings of their less endowed child. In other words, they invest in exactly the opposite way as predicted by Becker and Tomes: they are inefficient, valuing equality so much as to invest more in the kid who will generate a lower rate of return.

In general, the (weak or strong) efficiency model would lead us to expect that siblings' social statuses "diverge" as differential ability—combined with either neutral (weakly efficient model) or ability-reinforcing (strongly efficient model) investment patterns—leads ability differences to result in cumulative differences over the life course.

There are other family dynamics that might result in a similar pattern of increasing divergence over the life course. For example, this might hold true if parental (and intersibling) influences are primarily made when offspring are children at home and decline in importance as children reach adulthood. While sociological status attainment and life course models hold that earlier events and accomplishments structure later life events and accomplishments (Blau and Duncan 1967), this is a stochastic process (like a Markov chain process) in which additional (random) error is induced at each life course stage, thereby diluting the observed sibling correlation as we move outward from the educational system to job choice to earnings, income, and wealth accumulation. Along these lines, research by Jencks et al. (1972) suggests that income is affected by many chance life events, or "luck." Following (any of) these theories, we would expect higher sibling correlations for education, followed by occupation (since this may, to a great extent, reflect the legacy of decisions and tradeoffs made relatively early in adulthood), followed by earnings, then total family income (since this introduces dynamics within the marriage market as well), and finally by net worth (which typically starts to vary most in later adulthood thanks to earlier investment decisions). Following these models, we provide alternative hypotheses:

1. H_1 : sibling correlations for education > occupation > earnings > income > wealth.
2. H_2 : sibling correlations for less than 30 years of age > sibling correlations for 40

years of age > sibling correlations for 50 years of age.

Indeed, some recent research provides evidence of diverging sibling resemblance across the life course (Hauser, Sheridan and Warren 1999; Warren et al. 2002). Warren et al. (2002) use longitudinal data from the Wisconsin Longitudinal Survey and find that siblings become less occupationally similar as they age.

However, if we believe that parents act more like a welfare state and transfer resources to children in an iterative fashion in an attempt to maintain equality over the life course as they observe ability differences and as initial differences in success emerge, then we should observe the exact opposite pattern. That is, if we assume that there are multiple routes to economic success and it is the downstream economic measures (the ultimate rewards) that are equalized by parents rather than initial investments, we might hypothesize:

1. H_{1a} : sibling correlations for education, occupation, and earnings < sibling correlations for total income and wealth.
2. H_{2a} : sibling correlations for less than 30 years of age (on all measures) < sibling correlations for 40 years of age (on the same measures) < sibling correlations for 50 years of age (on the same measures).

Sibling Resemblance Across Population Subgroups

So far, the discussion above of Becker and Tomes, on the one hand, and Behrman et al. on the other has been working under the assumption of zero credit or capital constraints for investment in children. However, Becker and Tomes (1986) posit that *with* capital constraints, resource-constrained parents may not be able to optimally invest in their children's human capital. Such underinvestment may lead to higher degrees of sibling resemblance at lower incomes

since “high-ability children from poor families may receive the same low level of education as a sibling with lower academic ability, compressing their earnings compared with similarly different siblings from a prosperous family” (Mazumder and Levine 2003: 16).² Following this capital constraint model (which, like the parental preference model, assumes either equal investment or equal outcome norms among parents), we hypothesize:

1. H_3 : sibling correlations among groups with fewer relative resources (large families, nonwhite families, and children of parents with lower educational attainment) > sibling correlations among groups with greater relative resources (small families, whites, and children of parents with higher educational attainment).

The alternative hypothesis to Becker and Tomes’s (1986) theory is that among families that are disadvantaged, we should observe greater sibling disparities (i.e., lower correlations). This “strategic investment” model sees parental investment not as seeking to maximize the human capital (and earnings) of all children or even to equalize them but rather to maximize the human capital (and earnings) of the most endowed child, who promises the greatest returns on investment. This strategic investment model could result from one or both of two dynamics: First, that opportunity-constrained parents may view status attainment as a “sticky” process that has a stochastic element that varies with offspring ability (meaning the error term for high-ability children is seen to be less than the error term for low-ability offspring, thus suggesting greater chance of a return on investment). Second, such strategies may also be the most efficient route to the highest living standards for all offspring when parents expect cross-sibling subsidies in adulthood. That is, resource constrained parents may

(need to) follow the “strongly efficient” model and invest in the human capital of the best-endowed offspring, hoping for the highest possible returns and that, in time, that child will make wealth transfers to less-endowed siblings. In this model the parents are the engine of offspring income and wealth growth and the children themselves are expected to act as the social transfer agents to insure a degree of equality or insurance against gross within-family economic status differences (perhaps enforced by informal norms or strong parental pressure).

Put another way, disadvantaged families may be behaving more efficiently (investing more in the offspring for whom they expect higher returns), thus reinforcing sibling differences, whereas better-off families could be behaving inefficiently (investing more in the child for whom they expect lower returns), thus compensating, i.e., trying to bring about more equity in the outcomes of offspring. It is as if equality is a luxury good that only better-off families have the means to purchase. Why would this difference emerge? There are two possibilities: First, if “success” is seen as a threshold, well-off families may (rightly) believe that their best-endowed children will certainly “make it” and then can invest in the high-risk children. Second, if there is a declining marginal utility of child success, then parents who expect their children to be at different points along that curve may invest differentially to maximize total utility. In other words, the *differential* returns to investment may be steeper at the early part of the curve where disadvantaged families’ offspring are more likely to be and be flatter farther out, where the offspring of advantaged families are more likely to be operating, making for different relative choices between siblings for two hypothetical families.

Previous qualitative work (Conley 2004) has suggested that among disadvantaged households, sibling disparities tend to increase, since limited opportunities and resources may

evince parenting strategies that accentuate sibling differences by directing family resources to the better-endowed sibling(s) for whom upward mobility is most likely. This research also suggested that among families that were well endowed with class resources (and were racially privileged as well), parents often invested more heavily in those offspring they saw as having the worst chances for success in the education system and/or labor market—in a compensatory fashion à la the “separable earnings-bequest” model (Behrman, Pollak, and Taubman 1989) or the “child equality” model of Becker and Tomes. Following these models, we posit an alternative hypothesis:

H_{3a} : sibling correlations among groups with fewer relative resources (larger families, non-white, and children of parents with lower educational attainment) < sibling correlations among groups with greater relative resources (smaller families, whites, and children of parents with higher educational attainment). This difference in correlations should be most distinct for education (where parental investments matter most) and perhaps occupational prestige and earnings, and less pronounced for total income and wealth (if indeed, siblings engage in compensatory transfers).

We might expect this outcome for other reasons as well: If the “price effect” of human capital investment trumps the “preference displacement effect.” The preference displacement effect suggests that the budget constraint moves the earnings (and educational) possibility frontier closer to the origin, resulting in greater similarity among the offspring (similar to the Becker and Tomes model) since parents have minimal levels of investment that they desire to maintain among all children. In contrast to this effect, the price effect suggests that when parents

are budget constrained—by virtue of low income and wealth or thanks to larger family size—students are more likely to be eligible for financial assistance, which itself may vary with children’s endowments (see Behrman, Pollak, and Taubman 1989). In other words, relying on outside financing accentuates endowment heterogeneity. This is really a public finance model of the “strategic investment” model at the family level.

Empirical Results and Implications

We are not the first to propose examining sibling correlations in social and economic status as a way of assessing the effect of measurable and unmeasurable family background characteristics (including shared environment and shared genes) and intersibling effects on socioeconomic outcomes. For example, Daphne Kuo and Robert Hauser (1995) analyze the Occupational Changes in a Generation (OCG) survey data and find that, for education, sibling differences (within-family variance components) among various age groups of black and white brothers range between 38 percent and 52 percent. Mary Corcoran, Roger Gordon, Deborah Laren, and Gary Solon (1990: 364) estimate a brother-brother correlation in permanent income of .45 using data from the Panel Study of Income Dynamics (PSID). Bhashkar Mazumder and David Levine (2003) examine the National Longitudinal Surveys (NLS) and the PSID and argue that between the 1960s and the 1970s, the correlation in earnings between brothers rose from .26 to .45. Sibling resemblance in such other outcomes as welfare usage, education, and occupation follow similar patterns and are sensitive to the specification deployed—particularly for nonlinear measures. For example, if a woman’s sister has received welfare, she is over three times more likely to receive it herself as compared to a woman whose sister did not receive public assistance (.66 versus .20 probability in their PSID sample).

When we reanalyze more recent waves of PSID data—in which the siblings are on average older and more stable economically—with a substantially larger sample size of person-years and sibling sets, we obtain similar estimates of the association between siblings on education, occupational prestige, and earnings. See Figure 4. Parental preference models are challenged when we look to results across socioeconomic outcome measures in that sibling correlations in wealth remain consistently lower than sibling correlations in income (as well as education and occupational prestige). Parental preference models would lead us to expect the opposite—that parents invest unequally in their children’s education so as to efficiently maximize the return on ability endowments, but then seek to equalize sibling economic status when it comes to non-human capital measures. Our findings also lend some support to stratification and life course models in that we find that siblings resemble each other more on socioeconomic outcomes pertaining to earlier life events (those with less exposure to chance factors). However, when we analyze age effects directly, we find that siblings remain steady in their occupational resemblance, fluctuate in their earnings similarity but tend to converge in income, and this would not be expected given Jencks’ (1972) and others’ theory of declining sibling resemblance in income due to chance events (see Figure 5). (We do not analyze wealth once we get into subgroups since it was surveyed relatively infrequently in the PSID and thus suffers from fewer person-years in the sample.) Rather, the findings of convergence in income could be due to a greater degree of measurement error (of a classic errors-in-variables type) income earlier in the life course; or it could be because of the emergent salience of shared genes; alternatively, it could be due to an increase in direct intersibling effects in adulthood as we age (though this seems improbable thanks to rel-

atively weak adult sibling relationships in American society); it could be the result of compensatory parental impacts later in life (though also this seems unlikely given norms of adult offspring economic independence); most likely it is the effect of family background rearing its head in the marriage market (and with respect to marital stability. This is a feature of the stratification system that merits more attention—especially given the increased importance of multiple earnings in American families. At any rate, it suggests that if studies analyze economic correlations earlier in adulthood, they may be underestimating the degree of sibling similarity (and thus overstating the amount of income mobility) in the United States. This dynamic, as we will see below, holds particularly true for blacks.

Our analyses further build on these previous studies by exploring two questions regarding sibling convergence or divergence across the life course and across population subgroup. These results can be found in Figures 6 and 7. Whereas Becker and Tomes (1986) predict and Mazumder and Levine (2003) find that siblings who have fewer economic resources tend to demonstrate greater concordance in their socioeconomic statuses due to budget constraints on optimal investment in well-endowed offspring, we find mixed evidence on this front. Sibling correlations in occupation are indeed significantly higher for mothers with lower educational attainment (lending support to their theories). However, the pattern for earnings and income is the opposite (confounding their expectations), while education correlations are not significantly different across mother’s educational status.

Further complicating the picture, we report race effects such that black siblings resemble each other less on education, earnings and income than non-black siblings—obliquely supporting previous sociological work that has suggested weaker family effects

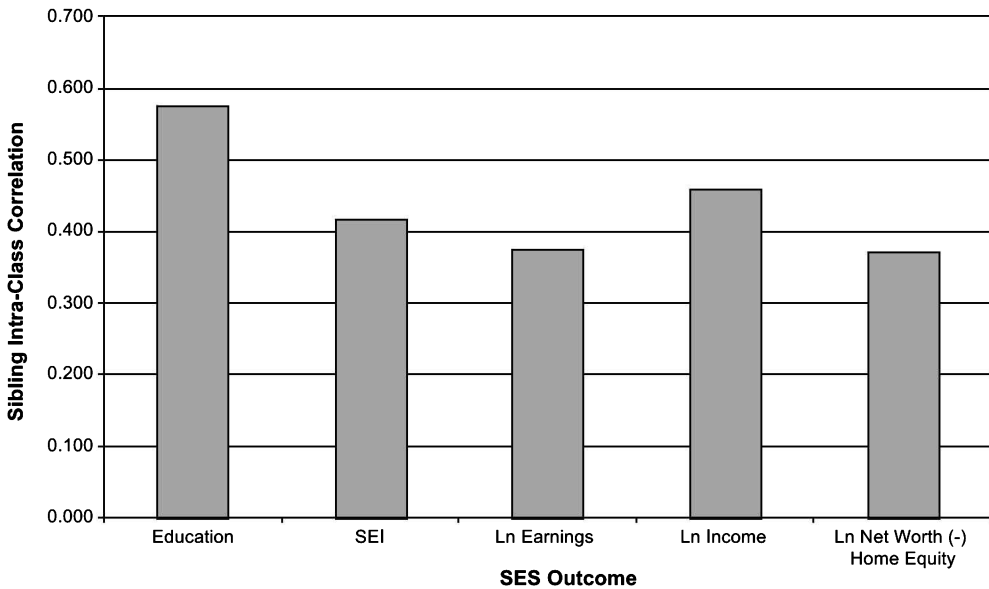


Figure 4. Sibling Correlations in SES Measures

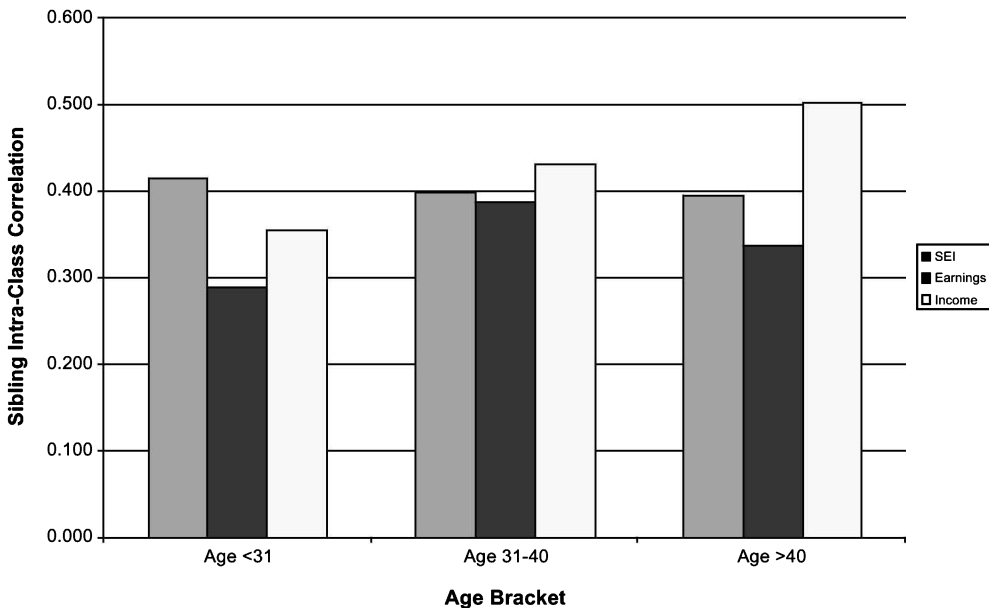


Figure 5. Sibling Correlations in SEI, Earnings, and Income, by Age

for blacks vis-à-vis whites with respect to occupation and education because of dynamics of discrimination and tokenism (see, for example, Blau and Duncan 1967; Featherman and Hauser 1978). (We say obliquely since the sign is opposite for SEI score, one of the

very measures upon which these previous researchers have relied.)

The story gets even more confusing when we interact age with class background or race. When we break out income correlations over the life course by maternal education level

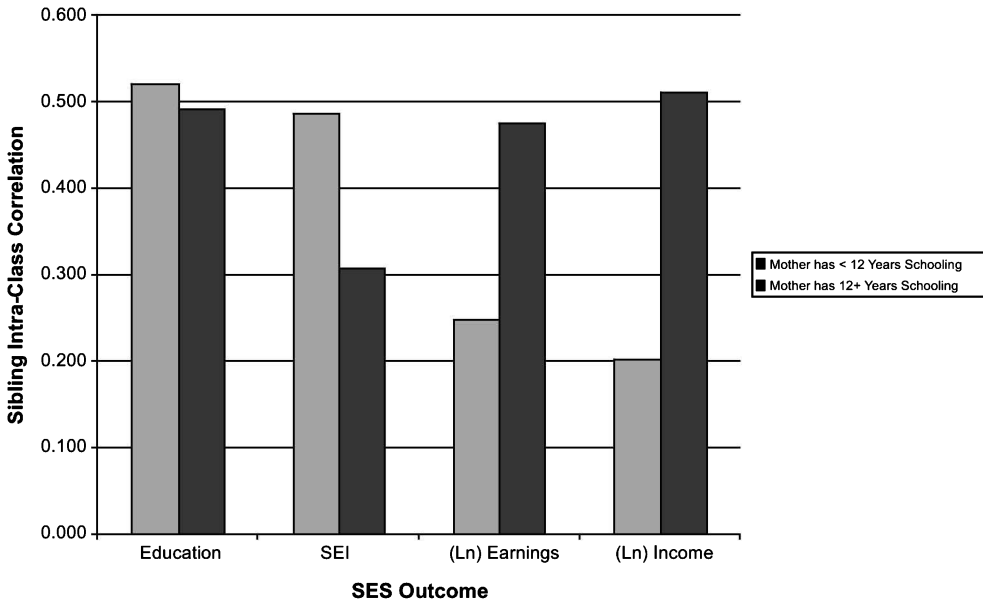


Figure 6. Sibling Correlations in Selected SES Measures by Mother's Education

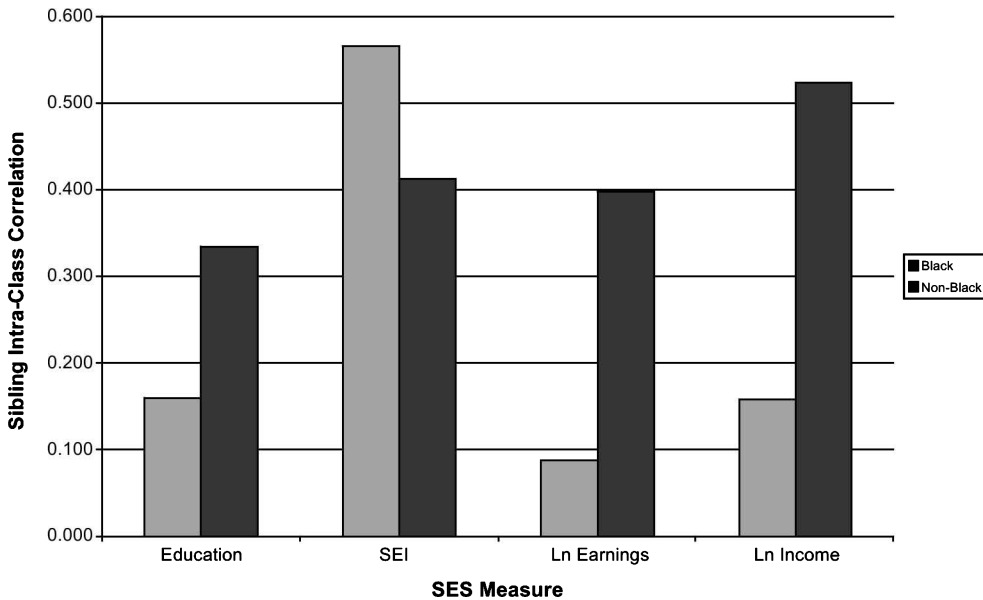


Figure 7. Sibling Correlations in SES Outcomes by Race (Black/Non-Black)

(again a proxy for the SES of the family of origin) we find that the similarity in income among siblings from more highly educated mothers tends to increase with age, peaking at a figure of .537 for the 40-year-old plus group (while declining [insignificantly] over

the life course among siblings born to mothers with less education, to a .329 level for the oldest group). (See Figure 8.) This pattern offers a challenge to the Becker-Tomes prediction about resource group differences. But when we examine the interaction of life

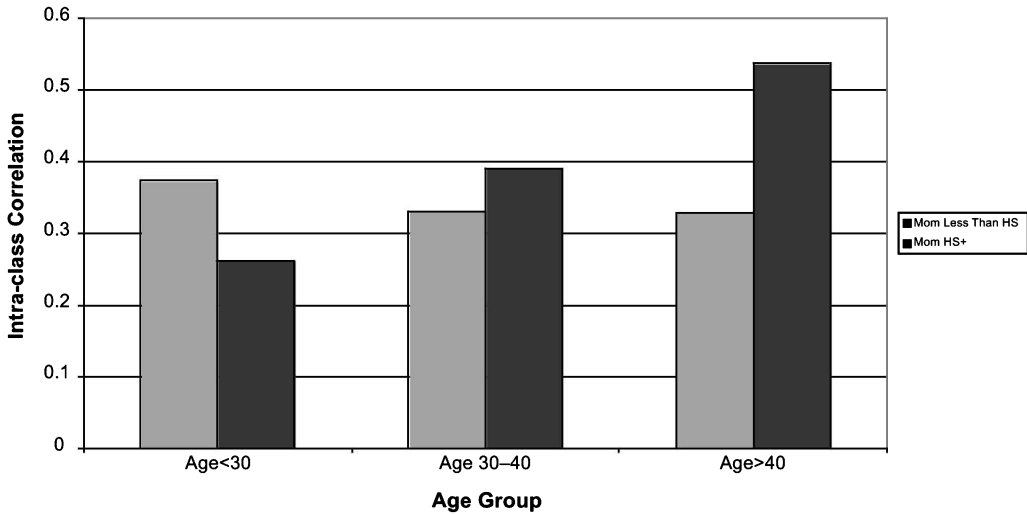


Figure 8. Sibling Correlations by Age and Class

course and racial differences we find that black siblings converge in income whereas non-black siblings do not converge in income as they age. Specifically, when we interact race and age in Figure 9, we contradict previous findings that black siblings have weaker family effects. Indeed, among the oldest age group, the black sibling correlation in income is a staggering .826—suggesting almost complete social reproduction in living standards by the fifth decade of life. (While statistically significant, we should keep in mind that this race-age analysis is based on a very small set of age 40 or older black siblings.) When contrasted to the opposite age-class interaction effect whereby the more advantaged group (as measured by maternal education, at least) experiences income convergence while the less-advantaged group does not, these results suggest, but by no means prove, that the American stratification system functions differently for blacks and whites in a way that is not reducible to resource issues alone—a proposition that begs for future research.

Conclusions

Though much more empirical work needs to be done, for now it is worth noting that re-

searchers may get very different answers to the seemingly simple descriptive question about how much income mobility there is in America depending on the age at which they measure economic status and the group(s) to which they are referring. Future researchers should perhaps refrain from using overall sibling correlations to make “global” assessments of the degree of openness in American society, as the answer appears to depend on the race-class group under study and the age group to which we are referring. Furthermore, taken together, results could imply that the simple model of equal investment in offspring and mean-regressive sibling effects may not hold equally for all groups in society. We are left where we started, with many questions about how class inheritance operates in American (or any) society and precious few solid answers.

Or, perhaps we are not left *quite* where we started. As mentioned, many researchers have tacitly or explicitly interpreted the sibling correlation as an omnibus measure of family background effects—that is, of the degree of ascription. In this vein, *when*, *where*, and for *what* measures sibling correlations are *low* is seen as an indication of social openness in a particular time and a specific

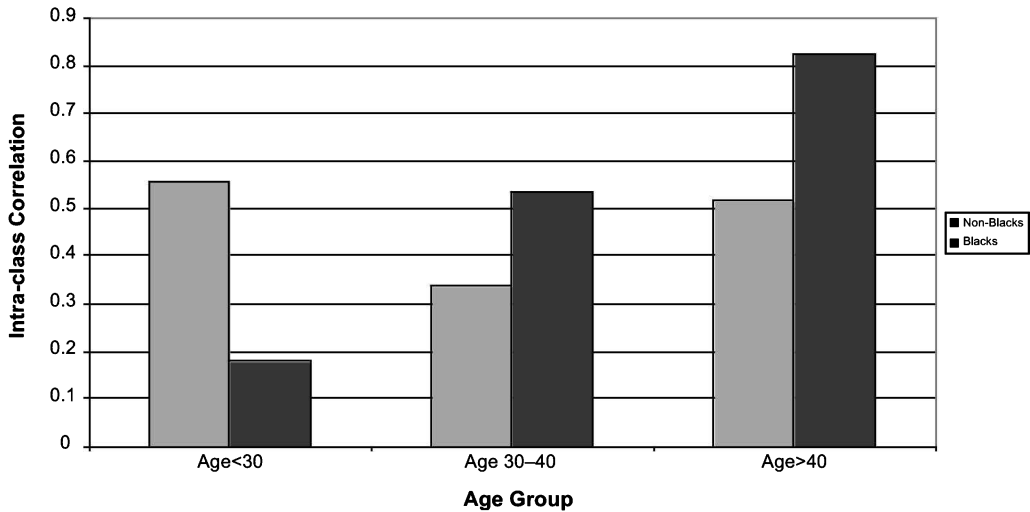


Figure 9. Sibling Correlations by Age and Race

place for a particular measure of socioeconomic status. By contrast, we tend to lament high degrees of sibling resemblance as indicative of a more caste-like system. However, such diagnoses may be premature. Envision the situation where there is an enormous amount of parent-child mobility; however, siblings move pair-wise away from their parents' status. For example, the two children of the manual worker both become doctors. From the point of view of sibling models, this scenario looks like a caste society (assuming the two children of the doctor end up similar in status as well). However, from the point of view of parent-child models, the picture painted is one of a completely open, fluid society with a high degree of mobility. Which is right? It is another instance of the blind men and the elephant.

The ultimate lesson is that when we consider the family as a cauldron and not as a rubber stamp, then we should not take any one piece of information—sibling *or* parent-child correlations—as indicative of the level of social ascription in a society.

It follows that low sibling correlations do not necessarily mean that family background—or the family more generally—does not matter for class attainment. Rather, the

family is a pivotal player in the assignment of social status and works not in a vacuum but in reaction to the resources within and the constraints and opportunities without.

A metaphor that may help illustrate this relationship between the individual, the family, and society writ large is a three-level chess board. At the first level is the individual; next is the family; and the top level is society. Envision a family where one sibling is recognized to be a musical prodigy from a young age and becomes the “star” of the household. Investment is directed to this talented child at the expense, perhaps, of others (perhaps due to limited overall resources). For example, the parents pay for daily private piano lessons at the expense of a house in a better school district (so the nonmusical sister suffers a worse public education as a result). Their investments appear to pay off, initially, as that kid gains admission to a prestigious conservatory. But classical music is what economist Robert Frank calls a “winner take all market” in which the rewards for a few individuals at the very top are enormous but where there is a steep drop to the next best performers (think professional sports, too). So, despite incredible investment and support from the parents, this

musically talented offspring ends up in the relatively low-status, low-pay position of a part-time music teacher to young children. Her sister, meanwhile, gets a regular college degree at the state university and works her way steadily up the corporate ladder. There is a high divergence in their earnings; and, ironically, it was the child who received less parental investment who economically outperforms the “golden child” in the end. While there was a good fit between the talents of the individual musician and the family’s values (they invested in her, after all), there was a poor fit between those two levels of the chess board and what the economy tends to reward at the present time. Does this case mean that family or class background didn’t matter?

Alternatively, imagine a family where one child is a math genius; but the parents are sports nuts and do not value achievement in the area of quantitative reasoning and thus do not invest much in this particular child. She may go on to outperform her siblings economically despite a lack of investment—given the high demand for math skills in the labor market. There was a good fit, in this case, between the individual’s talents and what society rewards, but a poor fit between the individual and the family, on the one hand, and the family and the society, on the other (which may have, in fact, disadvantaged *all* the offspring). Did class or family background matter in this case? Or do both these cases illustrate why and how family background can often have very little impact in a three-dimensional space of social attainment? You decide.

NOTES

1. For a good example of how these processes can be modeled, see, e.g., Behrman, Rosenzweig, and Taubman 1994.

2. Indeed, this is what Mazumder and Levine find: lower correlations among high-income siblings in both the 1968 and 1979 waves of the PSID. How-

ever, when they split the sample along the median, they end up with only a maximum of 185 multiple sibling sets in an income group (in the 1979 wave of the PSID, they have 1,086 cases from 901 families in all). Even if these were all two-sibling sets, then the maximum number of pairs for each income group would be 92—quite a small number. More important, however, is the fact that they do not split the sample based on the parental characteristics—as would be appropriate for the Becker-Tomes model—but rather by the incomes of the adult siblings themselves. This makes the sample split endogenous to their outcomes. In other words, what they may be observing could be a result of sibling decisions regarding tradeoffs between equity and efficiency. If certain sibships value equality, they may sacrifice the attainment of the better-endowed sibling, bringing the overall mean down, but resulting in a higher correlation between siblings. By contrast, we split the sample by parental measures, which are at least temporally anterior to the sibling outcomes.

REFERENCES

- Becker, Gary S., and Nigel Tomes. 1976. “Child Endowments and the Quantity and Quality of Children.” *Journal of Political Economy* 84: S143–S162.
- _____. 1986. “Human Capital and the Rise and Fall of Families.” *Journal of Labor Economics* 4: S1–S39.
- Behrman, Jere R., Robert A. Pollak, and Paul Taubman. 1982. “Parental Preferences and Provision for Progeny.” *Journal of Political Economy* 90: 52–73.
- _____. 1989. “Family Resources, Family Size, and Access to Financing for College Education.” *Journal of Political Economy* 97: 398–419.
- Behrman, Jere R., and Mark R. Rosenzweig. 2004. “The Returns to Birth Weight.” *Review of Economics and Statistics* 86: 586–601.
- Behrman, Jere R., Mark R. Rosenzweig, and Paul Taubman. 1994. “Endowments and the Allocation of Schooling in the Family and in the Marriage Market: The Twins Experiment.” *Journal of Political Economy* 102: 1131–1174.
- Blau, Peter, and Otis D. Duncan. 1967. *The American Occupational Structure*. New York: Free Press.
- Conley, Dalton. 2004. *The Pecking Order: Which Siblings Succeed and Why*. New York: Pantheon.
- Conley, Dalton, and Neil G. Bennett. 2000. “Is Biology Destiny: Birth Weight and Life Chances.” *American Sociological Review* 65: 458–467.
- Conley, Dalton, and Rebecca Glauber. 2005. “Sibling Similarity and Difference in Socioeconomic

- Status: Life Course and Family Resource Effects.” National Bureau of Economic Research Working Paper No. 11320.
- Corcoran, Mary, Roger Gordon, Deborah Laren, and Gary Solon. 1990. “Poverty and the Underclass: Effects of Family and Community Background on Economic Status.” *American Economic Review* 80: 362–366.
- Duncan, Greg J., and Daniel Hill. 1989. “Assessing the Quality of Household Panel Survey Data: The Case of the PSID.” *Journal of Business and Economic Statistics* 7(4): 441–451.
- Hauser, Robert M., and Hsiang-Hui Daphne Kuo. 1998. “Does the Gender Composition of Sibship Affect Educational Attainment?” *Journal of Human Resources* 33: 644–657.
- Hauser, Robert M., and William Sewell. 1986. “Family Effects in Simple Models of Education, Occupational Status, and Earnings: Findings from the Wisconsin and Kalamazoo Studies.” *Journal of Labor Economics* 4: S83–S115.
- Hauser, Robert M., Jennifer T. Sheridan, and John Robert Warren. 1999. “Socioeconomic achievements of siblings in the life course—New findings from the Wisconsin Longitudinal Study.” *Research on Aging* 21: 338–78.
- Hauser, Robert M., and R. Wong. 1989. “Sibling Resemblance and Intersibling Effects in Educational Attainment.” *Sociology of Education* 62: 149–171.
- Hill, Martha S. 1992. *The Panel Study of Income Dynamics: A User’s Guide*. Newbury Park, CA: Sage Publications.
- Jencks, Christopher S., M. Smith, H. Acland, M. J. Bane, D. Cohen, H. Gintis, B. Heyns, and S. Michelson. 1972. *Inequality: A Reassessment of the Effect of Family and Schooling in America*. New York: Basic Books.
- Kuo, H.-H. Daphne, and Robert M. Hauser. 1995. “Trends in Family Effects on the Education of Black and White Brothers.” *Sociology of Education* 68: 136–160.
- Mazumder, Bhashkar, and David I. Levine. 2003. “The Growing Importance of Family and Community: An Analysis of Changes in the Sibling Correlation in Earnings.” Federal Reserve Bank of Chicago, Working Paper 2003-24.
- McGue, Matt, and Thomas J. Bouchard. 1998. “Genetic and environmental influences on human behavioral differences.” *Annual Review of Neuroscience*. 21:1-24.
- Solon, Gary, Mary Corcoran, Roger Gordon, and Deborah Laren. 1991. “A Longitudinal Analysis of Sibling Correlation in Economic Status.” *Journal of Human Resources* 26: 509–534.
- Solon, Gary, Marianne E. Page, and Greg J. Duncan. 2000. “Correlations Between Neighboring Children in Their Subsequent Educational Attainment.” *Review of Economics and Statistics* 82: 383–392.
- Warren, John Robert, Jennifer T. Sheridan, and Robert M. Hauser. 2002. “Occupational stratification across the life course: Evidence from the Wisconsin Longitudinal Study.” *American Sociological Review*. 67: 432-455.

■ RATIONAL ACTION APPROACHES ■

64 ■ Richard Breen and John H. Goldthorpe

Explaining Educational Differentials

Towards a Formal Rational Action Theory

Introduction

In the light of recent research in the sociology of education, which has involved extensive over-time and cross-national analyses (see esp. Shavit and Blossfeld 1993; Erikson and Jonsson 1996b), it would seem that the following empirical generalizations can reliably be made and constitute *explananda* that pose an evident theoretical challenge.

Over the last half-century at least, all economically advanced societies have experienced a process of educational expansion. Increasing numbers of young people have stayed on in full-time education beyond the minimum school leaving age, have taken up more academic secondary courses, and have entered into some form of tertiary education.

Over this same period, *class* differentials in educational attainment, considered net of all effects of expansion per se, have tended to display a high degree of stability, i.e. while children of all class backgrounds have alike participated in the process of expansion, the pattern of association between class origins and the relative chances of children staying on in education, taking more academic courses or entering higher education has, in most societies, been rather little altered.

Children of less advantaged class origins have not brought their take-up rates of more ambitious educational options closer to those of their more advantaged counterparts.

It has, though, to be recognized that this latter generalization is not entirely without exception. In one national case at least, that of Sweden, there can be little doubt that class differentials in educational attainment have indeed declined over several decades (Erikson and Jonsson 1993); and, while some conflict of evidence remains, a similar decline has been claimed for The Netherlands (De Graaf and Ganzeboom 1993) and for Germany (Müller and Haun 1994; Jonsson, Mills and Müller 1996). Thus, any theory that is put forward in order to explain the more typical persistence of class differentials should be one that can at the same time be applied *mutatis mutandis* to such 'deviant' cases.

It would in addition be desirable that such a theory should be capable of yet further extension in order to account for a third regularity that has emerged from the research referred to.

Over a relatively short period—in effect, from the 1970s onwards—*gender* differen-

tials in levels of educational attainment, favouring males over females, have in nearly all advanced societies declined sharply and, in some instances, have been virtually eliminated or even reversed. In other words, while the process of educational expansion has not in the main led to children from less advantaged family backgrounds catching up with those from more advantaged backgrounds in their average levels of attainment, in families across the class structures of contemporary societies daughters have tended rather rapidly to catch up with sons.

In an earlier paper (Goldthorpe 1996), a theory of persisting class differentials in educational attainment, sensitive to the further requirements previously indicated, was developed from a 'rational action' standpoint. In the present [chapter], our aim is to refine this theory and to express it in a formal model. In this way we would hope to clarify its central arguments and in turn the wider implications that it carries. Since such attempts at the formalization of theory are still not very common in sociology, the [chapter] may also serve to stimulate discussion of the merits or demerits of this kind of endeavour. Readers interested in the more general *problematik* in the context of which the theory was initially conceived are referred to the earlier paper. In the remainder of this introductory section we set out certain 'background' assumptions of our subsequent exposition that will not be further discussed. The more specific assumptions on which our model rests will be introduced, and their significance considered, as the [chapter] proceeds.

We assume, to begin with, that class differentials in educational attainment come about through the operation of two different kinds of effect which, following Boudon (1974), we label as 'primary' and 'secondary'. Primary effects are all those that are expressed in the association that exists between children's class origins and their average levels of demonstrated academic ability. Children of more advantaged backgrounds are in fact

known to perform better, on average, than children of less advantaged backgrounds on standard tests, in examinations, etc. Primary effects, as will be seen, enter into our model but, fortunately, in such a way that we need not take up the vexed and complex question of the extent to which they are genetic, psychological or cultural in character. It is, rather, secondary effects that for us play the crucial role. These are effects that are expressed in the actual choices that children, together perhaps with their parents, make in the course of their careers within the educational system—including the choice of exit. Some educational choices may of course be precluded to some children through the operation of primary effects: i.e. because these children lack the required level of demonstrated ability. But, typically, a set of other choices remains, and it is further known that the overall patterns of choice that are made, are in themselves—over and above primary effects—an important source of class differentials in attainment.

We then further assume that, *in their central tendencies*, these patterns of educational choice reflect action on the part of children and their parents that can be understood as rational, i.e. they reflect evaluations made of the costs and benefits of possible alternatives—e.g. to leave school or to stay on, to take a more academic or a more vocational course—and of the probabilities of different outcomes, such as educational success or failure. These evaluations, we further suppose, will be in turn conditioned by differences in the typical constraints and opportunities that actors in different class positions face and in the level of resources that they command. However, what we seek to dispense with is any assumption that these actors will also be subject to systematic influences of a (sub)cultural kind, whether operating through class differences in values, norms or beliefs regarding education or through more obscure 'subintentional' processes. Not only do we thus gain in theoretical parsimony, but we

would in any event regard the ‘culturalist’ accounts of class differentials in educational attainment that have so far been advanced, as in various ways unsatisfactory (see further Goldthorpe 1996).

Finally, two other assumptions, regarding the structural context of action, should also be spelled out. On the one hand, we do of course suppose the existence of a class structure, i.e. a structure of positions defined by relations in labour markets and production units. And, in addition, we need to assume that within this structure classes are in some degree hierarchically ordered in terms of the resources associated with, and the general desirability of the positions they comprise. On the other hand, we suppose an educational system—i.e. a set of educational institutions that serve to define the various options that are open to individuals at successive stages in their educational careers. And here, too, we have a more specific requirement. That is, that this system should possess a diversified structure that provides options not just for more or less education but also for education of differing kinds, and that in turn entails individuals making choices at certain ‘branching points’ that they may not be able later to modify, or at least not in a costless way. It might be thought that this latter requirement will tend to limit the applicability of our model to educational systems of the more traditional European, rather than, say, the American variety, i.e. to ones where the type of school attended is likely to be more consequential than the total number of years spent in education. However, we would argue that, on examination, educational systems such as that of the USA, turn out to be more diversified than is often supposed, so that children do in fact face educational choices that involve considerations that go beyond simply ‘more’ or ‘less’: for example, in the American case, with the choice at secondary level between academic and vocational tracks. It is further of interest to note

how two American authors have specified in this regard the divergence between assumptions that we and they would share and those of most economists working within the ‘human capital’ paradigm. While for the latter education appears as a ‘fungible linear accumulation, like a financial investment’, a more realistic view would be that educational systems, the American included, ‘offer an array of choices and constraints that defy . . . simple linear formulations’ (Arum and Hout 1995, 1).

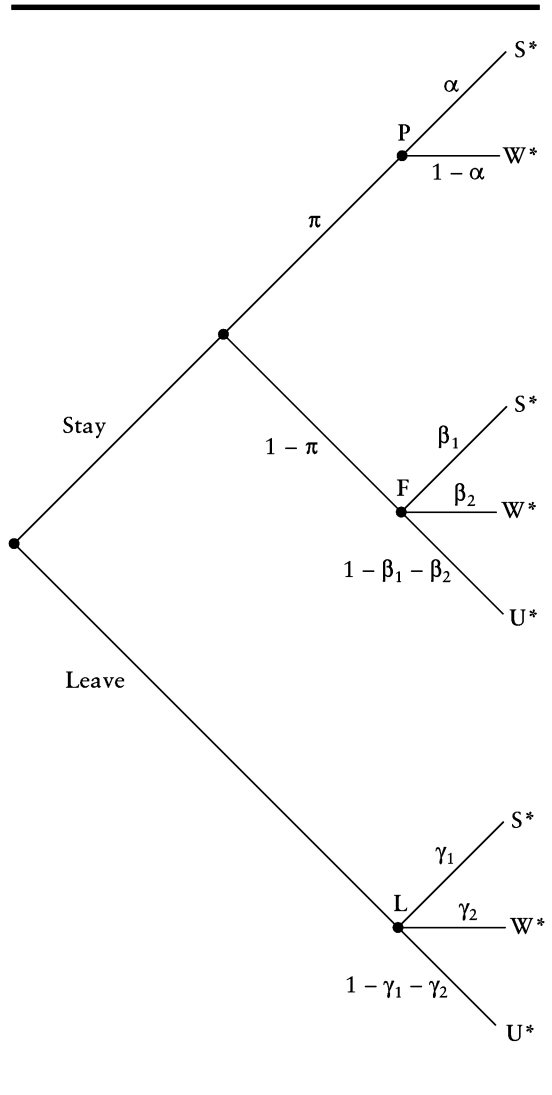
A Model of Educational Decisions

The model that we present is intended to be generic: that is, as one applicable in principle to the entire range of decisions that young people may be required to make over the course of their educational careers as regards leaving or staying on or as regards which educational option to pursue. However, in the interests of simplicity, we will here set out the model as it would apply just to the choice of leaving or continuing in education. The salient elements of the exposition are shown in Figure 1 by means of a decision tree. Here we assume that pupils must choose whether to continue in education—i.e. follow the ‘stay’ branch of the tree—to the completion of a further level (as, say, in the decision of whether or not to continue to A-level after GCSE) or to leave and enter the labour market, i.e. follow the ‘leave’ branch. Continuing in education has two possible outcomes, which we take to be success or failure. Because remaining at school often leads to an examination, we equate success with passing such an examination. This is indicated by the node labelled P in Figure 1, while failing the examination is indicated by the node labelled F. Leaving is then the third educational outcome in our model, that is, in addition to those of staying in education and passing and staying and failing, and is indicated by node L.

In deciding whether to continue in education or leave, parents and their children, we suppose, take into account three factors. The first of these is the cost of remaining at school. Continuing in full-time education will impose costs on a family which they would not have to meet were their child to leave school: these include the direct costs of education and earnings forgone. We can therefore express these costs relative to the costs of leaving by setting the latter to be zero and the former as $c > 0$. The second factor is the likelihood of success if a pupil continues in education. Since we distinguish only between success and failure, subjective beliefs about the chances of success at the next stage of education can be captured in our model by a single parameter, which we label π . This parameter measures the subjective conditional probability of passing the relevant examination given continuation. The third factor is then the value or utility that children and their families attach to the three educational outcomes represented by P, F and L in Figure 1. In our model this factor is expressed in terms of beliefs about the chances of access that each outcome affords to three possible destination classes.

For the purposes of our exposition, we take these classes as being the service class or salariat of professionals, administrators and managers (S^* in Figure 1), the working class (W^*) and the underclass, (U^*)—the class, say, of those with only a precarious place in the labour market and in only the lowest grades of employment if not unemployed. However, it should be emphasized that nothing of significance attaches to this choice of classes, except that, as earlier noted, we need to have a hierarchical ordering. Thus, the service class is regarded as comprising the most advantaged and most desirable positions and the underclass the least advantaged and least desirable, with the working class falling in-between. This

FIGURE 1
Single decision tree



ranking of classes is, moreover, assumed to be universally recognized or, at all events, not to vary across the population in any socially structured way.

As we have said, each of the three possible educational outcomes in our model has attached to it subjective probabilities of access to each of the three possible destination classes. So, as Figure 1 shows, for pupils who remain at school and pass their examination,

node P, the probability of access to the service class is given by α . There is no path linking this educational outcome to the underclass. This means that anyone who reaches this particular outcome is believed to be certain to avoid this class. It follows, therefore, that the probability of entering the working class, conditional on having been educationally successful, is given by $1 - \alpha$. At the other two outcome nodes, F and L, there is a positive probability of entering all three destination classes. So, for the outcome F (remaining at school and failing) the probability of access to the service class is given by β_1 , the probability of access to the working class by β_2 and the probability of access to the underclass by $1 - \beta_1 - \beta_2$. For the L outcome the corresponding probabilities are then given by the γ parameters.

We repeat that these are all subjective probabilities. Just as with π , the values for our various α , β and γ parameters reflect people's beliefs, in this case, about the returns to various educational outcomes conceptualized in terms of access to more or less desirable locations in the class structure. In principle, therefore, these parameters could vary widely across individuals and families. Again, though, we assume a societal consensus in regard to a set of beliefs that then serve as conditions on the parameters in question and that may be stated as follows:

- i. $\alpha > \beta_1$ and $\alpha > \gamma_1$. It is generally believed that remaining at school and succeeding affords a better chance of access to the service class than does remaining at school and failing or leaving school. Our model does not require that we make any assumptions about the relative magnitude of β_1 and γ_1 . It could, for example, be the case that a young person's chances of access to the service class are improved simply by acquiring more years of education, even if this does not lead to
- examination success. Alternatively, such time spent in education may be wasted in the sense that leaving school and embarking earlier on a career will yield a better chance of access to the service class.
- ii. $\gamma_1 + \gamma_2 > \beta_1 + \beta_2$. Remaining at school and failing increases the chances of entering the underclass. This means that there is a risk involved in choosing to continue to the next level of education.
- iii. $\gamma_2/\gamma_1 > 1$; $(\gamma_2/\gamma_1) \geq (\beta_2/\beta_1)$. Those who leave school immediately have a better chance of entry to the working class than to the service class. This may or may not be the case among those who remain at school and fail though, if it is, their odds of entering the working class rather than the service class are no greater than for those who leave school immediately.
- iv. $\alpha > 0.5$. Staying on at school and passing the examination makes entry to the service class more likely than entry to the working class.¹

In the interests of realism, especially as regards (ii) and (iii) above, it ought to be noted that 'leaving' and entering the labour market need not in most educational systems be equated with a definitive ending of the individual's educational career. Taking this option could in fact lead to further vocational courses pursued in conjunction with employment.

The Generation of Class Differentials

Given the model previously outlined, we can now turn to the question of explaining why differences exist across classes in the proportions of young people who make one kind of educational decision rather than another. For ease of exposition here we consider only two classes of origin, the service

class, S , and the working class, W . In all of what follows we assume that these classes differ in only two ways. First (and it is here that we give recognition to ‘primary’ effects) children of the two classes differ in their average ability. Ability is taken to be normally distributed within each class with means $a_s > a_w$ and variance given by σ_a^2 . Secondly, the two classes have different levels of resources, r , which they can use to meet c , the costs of education. Resources are taken to have a logistic distribution with mean values $r_s > r_w$ for the two classes and a common dispersion parameter, σ_r^2 . Throughout, we make no other assumptions about differences between the classes. In particular, and as earlier noted, we do not suppose any class-specific cultural values or social norms nor any class differences in the subjective α , β and γ parameters of our model.

We then propose three mechanisms through which class differentials in educational attainment may arise at the level of ‘secondary’ effects. Of these three, we would wish to stress the particular importance of the first, since this provides an account of how these differentials may be created and sustained through the apparently ‘free’ choices made by those in less advantaged classes. Our second and third mechanisms can be understood as accentuating the differing patterns of choice that derive from this initial source.

Relative Risk Aversion

We begin with an assumption regarding aspirations: that is, that families in both classes alike seek to ensure, so far as they can, that their children acquire a class position at least as advantageous as that from which they originate or, in other words, they seek to avoid downward social mobility. This means that the educational strategy pursued by parents in the service class is to maximize the chances of their children acquiring a position in this class. In terms of our model their

strategy is to maximize the probability of access to S^* . For working-class parents the implication is that they should seek for their children a place in either the working or the service class, since either meets the criterion of being at least as good as the class from which they originate. In terms of our model their strategy is then to maximize the probability of access to S^* or W^* , which is the same as minimizing the probability of access to U^* . This establishes families in both classes as having identical *relative* risk aversion: they want to avoid, for their children, any position in life that is worse than the one from which they start.

To see the consequence of these two strategies, maximize $pr(S^*)$ for those of service-class origins and minimize $pr(U^*)$ for those of working-class origins—assume, for the moment, that continuing in education is costless ($c = 0$). Then we find that whether or not a pupil believes it to be in his or her best interests to continue in education rather than leave depends on the value p_i (where i indicates the i^{th} pupil) given by

$$p_{is} = \frac{\pi_i \alpha + (1 - \pi_i) \beta_1}{\pi_i \alpha + (1 - \pi_i) \beta_1 + \gamma_1} \quad (1)$$

for the i^{th} service-class pupil and

$$p_{iw} = \frac{\pi_i + (1 - \pi_i)(\beta_1 + \beta_2)}{\pi_i + (1 - \pi_i)(\beta_1 + \beta_2) + (\gamma_1 + \gamma_2)} \quad (2)$$

for the i^{th} working-class pupil. Here we have allowed π to vary between pupils but we have assumed the values of α , β and γ to be common to all. If p takes a value greater than one-half this indicates that the expected returns to remaining at school exceed those of leaving. Thus, without taking account, as yet, of the costs of pursuing the former strategy, pupils for whom $p_i > 0.5$, can be said to prefer to remain in education. Even if subjective

expectations of future success, as captured by π , do not differ between the two classes it will nevertheless be the case that, given conditions (1) to (4), $p_{iS} > p_{iW}$ for any value of π less than one.²

Proof: $p_{iS} > p_{iW} \forall \pi \leq 1$ if and only if

$$\frac{\pi\alpha + (1 - \pi)\beta_1}{\gamma_1} > \frac{\pi + (1 - \pi)(\beta_1 + \beta_2)}{(\gamma_1 + \gamma_2)} \quad (3)$$

Taking the first term on the left hand side of (3) we have

$$\frac{\pi\alpha}{\gamma_1} = \frac{\pi}{\frac{1}{\alpha}\gamma_1} > \frac{\pi}{\gamma_1 + \gamma_2} \quad (3a)$$

by conditions (3) and (4). Taking the second term on the left hand side of (3) we have

$$\frac{(1 - \pi)\beta_1}{\gamma_1} \geq \frac{(1 - \pi)(\beta_1 + \beta_2)}{(\gamma_1 + \gamma_2)} \quad (3b)$$

by conditions (ii) and (iii). Together (3a) and (3b) imply (3) which in turn implies $p_i > p_w$ as required.

This result establishes that if continuing in education is costless and there are no class differences in the subjective probability parameters α , β , γ and π , children from middle-class backgrounds will more strongly 'prefer' (in the sense of perceiving it to be in their best interests) to remain in school to a further level of education rather than leave.

The proportions in each class who prefer to stay are derived as follows. Assume that p has an unspecified distribution with means in each class p_i and p_w and dispersion parameters σ_{pS} and σ_{pW} . Because $p_{iS} > p_{iW}$ for any common value of π_i , and assuming, for the moment, no class difference in the distribution of π , it follows that $p_i > p_w$. Then, given that only those pupils for whom p exceeds one-half prefer to stay at school, the

proportions in each class preferring this outcome are given by the area under the unspecified distribution function above the point

$$z_s = \frac{\frac{1}{2} - p_s}{\sigma_{pS}}$$

for the service class and analogously for the working class.

Differences in Ability and Expectations of Success

Thus far we have been assuming that the option of continuing in education is open to all pupils. But, of course, this is often not the case and successive levels of education may only be open to those who meet some criterion, such as a given level of performance in a previous examination. Let us assume, for the sake of simplicity, that this criterion can be expressed directly in terms of ability, so that, for example, a pupil may only continue in education if his or her ability level exceeds some threshold, k : i.e. we impose the condition that a_i must be greater than k . Recalling our assumption regarding primary effects that the mean level of ability is higher in the service class than in the working class but that both have the same variance in ability, it follows that the proportion of service-class children who meet this condition exceeds the proportion of working-class children.

However, we might also suppose that pupils' own knowledge of their ability helps shape the subjective probability they attach to being successful in the next stage of education, which we labelled π_i . So we can write $\pi_i = g(a_i)$, where g indicates that π is a function of a . If we then denote by π_s^* and π_w^* the required minimum subjective probabilities compatible with continuing in education (these are the smallest values of π_i for which $p_i > 0.5$) we can write the probability of continuing in education as

$$\begin{aligned}
 & pr(a_i > k)pr(\pi_i > \pi^* \mid a_i > k) \\
 & = pr(a_i > k, \pi_i > \pi^*) \\
 & = pr(a_i > k, g(a_i) > \pi^*) \quad (4)
 \end{aligned}$$

If

$$pr(g(a_i) > \pi^*) \leq pr(a_i > k)$$

then (4) reduces to

$$pr(\pi_i > \pi^*).$$

If pupils' expectations about how well they will perform at the next level of education are upwardly bounded by how well they have performed in their most recent examination—for example, if there are no pupils who, although they have failed to exceed the threshold k are nevertheless sufficiently optimistic about their future examination performance to wish to continue in education—then ability differences will be wholly captured in differences in the subjective parameter π . This will cause the average value of π to be lower among working- than service-class pupils because of the class difference in average ability levels.

Differences in Resources

Thus far we have assumed education to be costless. If we relax this assumption we need to take account of class differences in the resources that families in different locations in the class structure can devote to their children's education. Assume, therefore, that pupils can continue in education if and only if $r_i > c$ where r_i is the level of resources available for children's education in the i^{th} family. Given that service-class families have, on average, greater resources than working-class families ($r_s > r_w$) and that resources have the same dispersion within each class, it follows that the proportion of service-class pupils for whom this resource requirement is met will exceed the proportion of working-class pupils.

We have now suggested three mechanisms which, taken together, give rise to class differentials in the proportions of children who choose to stay on in education. Our first mechanism shows how, solely because of the relative risk aversion that is seen as being common across classes, there will be a stronger preference among service- than working-class pupils for remaining in education given that no costs attach to doing so. Our second mechanism then allows for class differences in average ability levels and in turn in expectations of success. The effect of this is to introduce class differences in the values of π (the subjective probability of future educational success), which further widen class differences in the value of p and thus in the strength of the preference for staying on in education. Finally, our third mechanism takes account of the costs of continuing in education and allows for a further source of class differentials, the average resource levels available to meet these costs. The effect of this is to promote class differences via the proportion of families in each class whose resources exceed the costs of their children continuing in education or, more simply, who can afford to allow their children to continue. . . .

Explaining Empirical Generalizations

We now seek to apply our model to the explanation of the empirical generalizations that were set out in the Introduction, beginning with that of the widely observed persistence of class differentials in educational attainment in the context of an overall increase in educational participation rates. To account for the latter trend is fairly easy: the relative costs of education have declined over time in all economically advanced countries. As the period of compulsory schooling has been extended, the costs of successively higher levels of education have been reduced through the abolition of fees, the introduction

of maintenance grants, soft loans, etc. In our model this change is treated via mechanism (iii)—class differences in resources—and is captured in a decline in the size of the parameter c . This will lead to an increase in the proportions of children from both service- and working-class origins continuing in education, providing, of course, that the preference for continuing (given by our p parameter) does not decline. However, far from p_i declining over time it is more plausible to believe that there is a widespread increase in the desire to remain in full-time education as educational credentials come to take on increasing importance in the labour market and in securing a relatively advantaged class position. Indeed, in so far as education is regarded as a 'positional' good, p_i could be expected to rise steadily simply as a consequence of educational expansion itself.

At the same time our model can provide an explanation of how, within a context of educational expansion, class differentials may none the less persist. To see this, recall that class differences in educational attainment are usually measured by odds ratios which compare the odds of continuing in education versus leaving for pairs of origin classes. Under our model, the odds ratio between the service and the working class is equal to

$$\frac{\phi_S / (1 - \phi_S)}{\phi_W / (1 - \phi_W)}$$

where we use ϕ_S to mean the proportion of service-class pupils who remain in education and similarly for ϕ_W . It is then possible to show (see Breen and Goldthorpe 1997, 300–2) that, given a decline over time in c , together with an increase in the proportion of both service- and working-class pupils who consider it in their best interests to remain in education, the odds of continuing in education increase by a roughly constant amount for each class, and so preserve a similar constancy in the odds ratio. This tells us

that, under these circumstances, a uniform decline in the costs of education, i.e. uniform across classes, will result in the odds for children of all classes choosing to continue being multiplied by something like a common factor. So if, for example, some level of education is made free of charge (in the sense that fees are no longer levied) class differences in participation (as measured by odds ratios) at this level will remain more or less unchanged even though the overall participation rate will increase.

Our model also sheds some new light on the concept of 'maximally maintained inequality' in education (Raftery and Hout 1990; Hout, Raftery and Bell 1993). These authors argue that class differences in educational attainment will only begin to decline when participation in a given level of education of children of more advantaged backgrounds reaches saturation. In our model, such a reduction will occur once c declines to the value at which all members of the service class have resources that exceed it. At this point, all service-class families will possess resources that exceed the costs of remaining in education and thus the proportion in this class who choose to continue in education will be equal to the proportion who perceive it to be in their interests (i.e. for whom $p_i > 0.5$). Further reductions in c will then have no influence on the numbers of service-class children who choose to continue but will still increase the proportion of working-class children who do so. Under these conditions, the relevant odds ratio could be expected to move towards unity.³ However, it should be recognized that, as understood in terms of our model, maximally maintained inequality does not imply that a decline in class differentials can only commence at the point at which all children of more advantaged class origins continue in education. Rather, this effect occurs once all such children whose p_i is greater than one-half continue, in other words, once all those who perceive it to be in

their best interests to continue are able to act accordingly. It is true that in some instances the achievement of this latter condition will, in fact, give rise to 100 percent continuation among children of more advantaged classes.⁴ But further declines in c , even if they lead to $r_{iW} > c$ for all members of the working class, will not lead to equality in the proportions continuing in education in each class so long as there still remains a class difference in the proportion who prefer to continue.

It further follows from our model that class differentials in educational attainment will also respond to changes in the costs of education which, rather than being uniform, have a variable impact across classes. Such changes could be brought about directly through the selective subsidization of young people according to their class of origin, as occurred, for example, in some post-war Communist societies. However, essentially the same effect could follow from a general reduction in inequality of condition between classes. Specifically, if class differences in resources, r , become smaller, our model would predict that differentials in educational attainment, as measured by odds-ratios, would in turn decline.

It is then in this way that the model may be seen as applying to the national case that most obviously deviates from the typical pattern of persisting class inequalities in education, i.e. that of Sweden, in which, as earlier noted, a narrowing of such inequalities over the post-war decades is well attested. There is indeed further extensive evidence (for reviews, see Erikson and Jonsson 1996a; Goldthorpe 1996) that in this same period the average income levels of different classes in Sweden became more equal, while the degree of economic insecurity experienced by members of the working class was steadily reduced. And through time-series analysis, correlations can in fact be established between these latter tendencies and the growing equality in educational outcomes that are at

all events consistent with the hypothesis of a causal influence (Erikson 1996).

As against the constancy in class differentials in educational attainment, to which exceptions are few, the decline in gender differentials that has occurred in virtually all advanced societies since the 1970s must appear as rather dramatic. Because gender differentials arise within, rather than between, families, neither changes in the costs of education nor in inequalities in resources among families are appropriate to explaining their reduction. In the light of our model, this may rather be seen as resulting from shifts in the perception of educational returns that have been prompted by changes in women's labour market participation. It would be fair comment to say that the pattern of returns to different educational decisions that we have thus far envisaged would, for most of the 20th century, be more applicable to young men than to young women. Until quite recently, it is likely that educational decisions in the case of girls were shaped in the main by the expectation that their primary social roles would be those of wife and mother, and that their class positions would therefore be determined more by whom they married than by how they themselves fared in the labour market. In so far as this were the case, then the relative returns to education for women would be somewhat different to those we have supposed in the exposition of our model: at all events, the returns associated with any particular educational decision would be less highly differentiated than for men. So, for example, young women of service-class origins could be thought best able to retain their position in this class through marriage; but to meet young service-class men did not necessitate that they themselves should acquire the educational qualifications that led to a service-class occupation. Rather, their qualifications had to be such as to provide them with employment that would bring them into contact with

potential service-class husbands, and this requirement might be met through only relatively modest levels of educational attainment, leading to a job as, say, a secretary or nurse. And within both the home and the educational system alike, as much emphasis was indeed placed on the acquisition of social and domestic skills as on skills that would have value in the labour market.

Such a flatter 'gradient' in the returns to different educational pathways would, if incorporated into our model, have two consequences. First, the proportion of women choosing to remain in education at each decision point would be smaller than the proportion of men; and, second, class differentials would tend to be less among women than among men. The former result follows from the lesser incentive to continue in successively higher levels of education that would be held out to women of all class origins alike; the latter comes about because the magnitude of the class differences among those choosing to remain in education (for given values of ability and resources) is directly proportional to the differences in returns associated with the various possible educational outcomes. If we consider equations (1) and (2) shown earlier, then as the difference between, say, α_1 , β_1 and γ_1 diminishes, so the difference between p_{iS} and p_{iW} will also diminish.

Over the past 20 years, we would suggest, the pattern of returns to education for women has drawn closer to that for men, as rates of women's labour market participation and, especially rates for married women, have increased and as a woman's own employment has taken on greater significance in determining the standard of living enjoyed by her family and further, perhaps, her own class position. In other words, our model as expressed in Figure 1 has come increasingly to apply to women: the 'gradient' in their returns to education has steepened. According to our model, then, such a change should

have two effects: gender differentials in educational attainment should decline, as indeed they have, and at the same time the magnitude of class differences among women should increase.

Conclusions: Theoretical Implications of the Model

As regards the theoretical implications of our model, we would see these as being of main significance in their bearing on explanatory strategy. The model represents children and their families as acting in a (subjectively) rational way, i.e. as choosing among the different educational options available to them on the basis of evaluations of their costs and benefits and of the perceived probabilities of more or less successful outcomes. It then accounts for stability, or change, in the educational differentials that ensue by reference to a quite limited range of situational features. For example, in the case of persisting class differentials, the explanatory emphasis falls on similarly persisting inequalities in the resources that members of different classes can command in the face of the constraints and opportunities that their class positions typically entail. Class differences in demonstrated academic ability are also recognized, but not—as we have emphasized—class differences of a (sub)cultural character.

To the extent, then, that our model holds good, i.e. that it can provide an adequate account of the regularities we have considered and that its further empirical implications are not rejected—the relatively parsimonious strategy of the rational action approach is supported; and, we might add, in an area in which 'culturalist' theories of one kind or another have hitherto enjoyed great popularity—even if not great explanatory success (see Goldthorpe 1996). In turn, the case for attempting to pursue this strategy in other areas of sociological enquiry is strengthened.

Finally, though, we would wish to allude to certain theoretical implications that might be regarded as following from our model but that do not in fact do so. To begin with, we are not required to suppose that, in making educational choices, children and their parents in fact go through the processes of ratiocination that the model might appear to attribute to them. We do take it to be the case that the actors in question have some knowledge of how their society works, have some concern for their own, or for family interests, and seek to use the former to promote the latter. But we can at the same time accept that the decisions they make may only rarely result from any entirely explicit procedures, rather than, say, 'emerging' over a period of time and, in all probability, reflecting also various non-rational influences. What underlies our approach is the idea that it is rational considerations that are, not the only, but the main common factor at work across individual instances, and that will therefore shape patterns of educational choices in aggregate and, in turn, the regularities that constitute our *explananda*. Our model then aims to represent these considerations in an 'idealized' way, so as to capture the key generative processes involved, rather than to represent decision-making as it actually occurs at the level of particular families.

Further, while we do not in explaining class differentials in education invoke systematic variation in values or derived norms, this does not mean that we have to deny their very existence. Thus, in so far as class-specific norms may be identified—which is an empirical issue—we could recognize them as serving as *guides* to rational action that have evolved over time out of distinctive class experience and that may substitute for detailed calculation when educational choices arise. Understood in this way, such norms could conceivably be of some explanatory significance as inertial forces in

cases where the structure of constraints and opportunities or the distribution of resources is changing. But what we would in fact expect, and the decline in gender differentials would, at least by analogy, lend support, is that norms, in being essentially epiphenomenal, would rather quickly come into line with patterns of action that display a rational adaptation to the new circumstances that have come into being.

In sum, our model implies an explanatory strategy that is undoubtedly 'reductionist' so far as the relation of norms to rational action is concerned (see Elster 1991). However, we do not in this regard seek what Popper (1972: Ch. 8) has criticized as reduction by fiat, but only reduction in so far as it is warranted by the empirical support that our theoretical arguments can obtain in the particular area in which they have been applied.⁵

NOTES

1. Strictly speaking the mathematics of our model require a slightly weaker condition, namely that $\alpha \geq \gamma_1 / (\gamma_1 + \gamma_2)$. This imposes a condition on the magnitude of the difference in the chances of access to the service class as between remaining at school and passing the examination and leaving immediately. The conditional probability of access to the service class for those who leave immediately should not be greater than $\gamma_1 + \gamma_2$ times the conditional probability of access to the service class for those who remain at school and pass the examination. However, because of condition (iii), condition (iv) will always be met if $\alpha > 0.5$.

2. Note that, whereas p_{iW} can take any value between zero and one depending on the value of π , if $\beta_1 \geq \gamma_1$ then p_{iS} will exceed one-half for all values of π .

3. Though empirically this will be observed only if the proportion of service-class children who consider it in their best interests to remain in education does not change for other reasons. For example, given an increase over time in the importance of educational qualifications in obtaining jobs we might see changes in the relative values of the α , β and γ parameters causing the proportion for whom $p_i > 0.5$ to increase in both classes. Under these conditions a narrowing of the odds-ratio will not necessarily follow.

4. In our model this will be the case for the service class if (in addition to conditions (i) to (iv)) $\beta_1 \geq \gamma_1$ but it need not be so if this inequality does not hold.

5. Elster (1991) criticizes several different versions of the argument that action taken in conformity with social norms is reducible to rational action. However, his efforts to show that no version entails that such a reduction is always possible are of greater philosophical than sociological interest. One could entirely agree with Elster, yet still wish to maintain that, in a particular instance of sociological explanation, a reductionist view could in fact be upheld; or, more generally, that it is good strategy to start from a reductionist position and to modify it only in so far as the evidence requires.

REFERENCES

- Arum, R. and M. Hout. 1995. 'The Early Returns: The Transition from School to Work in the United States.' European University Institute, School-to-Work Conference, Florence.
- Boudon, R. 1974. *Education, Opportunity and Social Inequality*. New York: Wiley.
- Breen, R. and J.H. Goldthorpe. 1997. 'Explaining Educational Differentials. Towards a Formal Rational Action Theory.' *Rationality and Society* 9: 275–305.
- De Graaf, P.M. and H.G.B. Ganzeboom. 1993. 'Family Background and Educational Attainment in the Netherlands for the 1891–1960 Birth Cohorts.' In *Persistent Inequality: Changing Educational Attainment in Thirteen Countries*, eds Y. Shavit and H.-P. Blossfeld. Boulder: Westview Press.
- Elster, J. 1991. 'Rationality and Social Norms.' *Archives Européennes de Sociologie* 32: 109–29.
- Erikson, R. 1996. 'Can We Account for the Change in Inequality of Educational Opportunity?' In *Can Education be Equalised? The Swedish Case in Comparative Perspective*, eds R. Erikson and J.O. Jonsson. Boulder: Westview Press.
- Erikson, R. and J.O. Jonsson. 1993. *Ursprung och Utbildning*. Stockholm: Statens Offentliga Utredningar.
- Erikson, R. and J.O. Jonsson. 1996a. 'Explaining Class Inequality in Education: the Swedish Test Case.' In *Can Education be Equalised? The Swedish Case in Comparative Perspective*, eds R. Erikson and J.O. Jonsson. Boulder: Westview Press.
- Erikson, R. and J.O. Jonsson. (eds). 1996b. *Can Education be Equalised? The Swedish Case in Comparative Perspective*. Boulder: Westview Press.
- Goldthorpe, J.H. 1996. 'Class Analysis and the Reorientation of Class Theory: The Case of Persisting Differentials in Educational Attainment.' *British Journal of Sociology* 45(3): 481–506.
- Hout, M., A.E. Raftery and E.O. Bell. 1993. 'Making the Grade: Educational Stratification in the United States, 1925–1989.' In *Persistent Inequality: Changing Educational Attainment in Thirteen Countries*, eds Y. Shavit and H.-P. Blossfeld. Boulder: Westview Press.
- Jonsson, J.O., C. Mills and W. Müller. 1996. 'A Half Century of Increasing Educational Openness? Social Class, Gender and Educational Attainment in Sweden, Germany and Britain.' In *Can Education be Equalised? The Swedish Case in Comparative Perspective*, eds R. Erikson and J.O. Jonsson. Boulder: Westview Press.
- Müller, W. and D. Haun. 1994. 'Bildungungleichheit im Sozialen Wandel.' *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 46: 1–42.
- Popper, K.R. 1972. *Objective Knowledge*. Oxford: Clarendon Press.
- Raftery, A.E. and M. Hout. 1990. 'Maximally Maintained Inequality: Expansion, Reform and Opportunity in Irish Education, 1921–1975.' ISA Research Committee on Social Stratification and Mobility, Madrid.
- Shavit, Y. and H.-P. Blossfeld. (eds). 1993. *Persistent Inequality: Changing Educational Attainment in Thirteen Countries*. Boulder, Westview Press.

PART VI

Race and Ethnicity

Constructing Racial Categories

65 ■ Reynolds Farley

Racial Identities in 2000

The Response to the Multiple-Race Response Option

The greatest change in the measurement of race in the history of the United States occurred in the census of 2000. For more than two centuries, the federal statistical system had classified each respondent into a single race. That is no longer the case. According to the new rules, anyone may now identify with as many races as he or she desires. . . .

The efforts to change the one-race-only classification system came from individuals who believed that many Americans shared multiple racial origins, but information about the size or growth of this population was seldom presented. Indeed, no reliable data existed before the Census Bureau pretested queries for the census of 2000. Although a small number of studies attempted to estimate changes over time in the frequency of interracial marriages (Kalmijn 1993, 1998), few data sets lend themselves to the measurement of rare events; until 1967, some states prohibited marriages across racial lines, and so there are few data on the mixed-race offspring of interracial unions. However, public-use files from census data may be used to ascertain the percentage of married couples in which the husband and wife reported different races.

As a result, studies of interracial marriage have become numerous (Fu 2001; Gilbertson, Fitzpatrick, and Yang 1996; Heer 1974; Hwang, Saenz, and Aguirre 1995; Model and Fisher 2001; Qian 1997; Rosenfield 2001; Sandefur and McKinnell 1986; Schoen and Thomas 1989). Presumably, as the proportion of interracial marriages increases, so too will the percentage of children who think of themselves as multiple by race.

Table 1 reports data about interracial marriage compiled from the 1990 census and the March 2000 Current Population Survey. Spouses were classified into one of five mutually exclusive and exhaustive categories: Hispanic, non-Hispanic white, non-Hispanic black or African American, non-Hispanic Asian or Pacific Islander, and non-Hispanic American Indian. For these tabulations, Hispanic origin is treated as if it were equivalent to a race.

According to the table, the percentage of married couples in which the races of the spouses differ rose from 4.5 percent at the start of the 1990s to 5.4 percent at the end—an increase from 2.4 million interracial couples to about 3.0 million. This finding is consistent with previous studies that reported a

Table 1. Estimates of Racial Intermarriage, 1990 and 2000, by Race and Sex (Thousands)

| <i>Racial Group</i> | Men | | Women | |
|------------------------|-------------|-------------|-------------|-------------|
| | <i>1990</i> | <i>2000</i> | <i>1990</i> | <i>2000</i> |
| White | | | | |
| Number | 43,710 | 43,809 | 43,546 | 43,590 |
| Percentage married out | 2.7 | 3.4 | 2.2 | 2.9 |
| Black | | | | |
| Number (000) | 3,566 | 4,089 | 3,383 | 3,872 |
| Percentage married out | 6.3 | 8.4 | 2.4 | 3.3 |
| Hispanic | | | | |
| Number (000) | 3,489 | 5,181 | 3,562 | 5,318 |
| Percentage married out | 18.7 | 15.0 | 19.9 | 17.2 |
| Asian | | | | |
| Number (000) | 1,299 | 1,922 | 1,543 | 2,201 |
| Percentage married out | 11.3 | 9.9 | 24.1 | 21.3 |
| American Indian | | | | |
| Number (000) | 300 | 351 | 330 | 371 |
| Percentage married out | 57.8 | 52.3 | 61.5 | 53.9 |
| Total | | | | |
| Number (000) | 52,364 | 55,352 | 52,364 | 55,352 |
| Percentage married out | 4.5 | 5.4 | 4.5 | 5.4 |

Source: U.S. Bureau of the Census, Public Use Microdata Files from the 5 Percent Sample of the Census of 1990 and from the March 2000 Current Population Survey.

Note: These data tabulate the race of spouses for persons married and living with a spouse. Five mutually exclusive races were used: Hispanic, non-Hispanic white, non-Hispanic black, non-Hispanic Asian (including Pacific Islanders), and non-Hispanic American Indian.

secular trend toward increasing interracial marriage, especially among blacks and whites (Farley 1999, figures 5.7 and 5.8). Presumably, a growing but still small minority of children have parents of different races.

The primary reason for change through the 1990s was that white and black men and women increasingly married persons from races other than their own. The percentage of black men with nonblack wives rose from 6.3 to 8.4 percent, while that of white men with nonwhite wives went from 2.7 to 3.4 percent. Interestingly, the percentage of Hispanic and Asian spouses marrying outside their race decreased. Migration from abroad has produced rapid growth for these groups. Immigrant enclaves are getting much larger, and this presumably promotes greater racial homogeneity for Hispanics and Asians who marry, especially among immigrants. . . .

The Official Governmental Decision About Measuring Race

Because the new 1997 OMB (Office of Management and Budget) guidelines for measuring race will have consequences for decades to come, it is important to summarize them carefully. They include the following:

- When self-identification is used to identify race in the federal statistical system, respondents should be given the opportunity to identify with more than one race, but the term “multiracial” is not to be used. Rather, the names of specific races are to be presented as choices for the respondent.
- When self-identification is the method of data collection, separate questions should be asked to determine race and Hispanic ethnicity.

- The Hispanic-ethnicity question should precede the race question (an important change from the 1990 census procedure).
- “Arab” and “Cape Verdean” should not be used as races. (Spokespersons for these groups had advocated the addition of these terms as races.)
- The term “American Indian” should not be replaced by “Native American,” but the term “Hawaiian” should be replaced by “Native Hawaiian.”
- The term “Alaskan Native” should replace “Aleut” or “Eskimo.”
- Asian races should be distinguished from Pacific Islander races. A broad racial category called “Native Hawaiian or Other Pacific Islander” (NHOPI or NHPI) should be used.
- The five major racial categories to be used are white, black or African American, American Indian, Asian, and Native Hawaiian or other Pacific Islander.

The radical shift away from the assumption that each person had only one racial identity did not occur because powerful civil rights organization or professional associations advocated a change nor because federal enforcement agencies demanded improved data to carry out the responsibilities mandated by Congress and the courts. Rather, a small number of persistent advocates who were greatly upset by the status quo effectively challenged the traditional system. They found a sympathetic hearing in Congress and in the Clinton administration: elected and appointed officials did not want to be in the untenable position of insisting that a person had only one racial identity when he or she wished to identify with two or more races. The idea that every individual can be classified as a member of a single race may gradually disappear in the United States, just as the idea that people were either Angles or Saxons disappeared in England.

How the OMB Guidelines Were Put into Operation for the 2000 Enumeration

After data were gathered in 2000, the Census Bureau developed coding rules to process the results. These rules increased the number of major races by adding a sixth option, “some other race.” Persons who identified with a single race—those who either checked only one box or checked no box but wrote a term directly denoting race, such as white, black, or African American, in the “some other race” space or anywhere near the race question on the census questionnaire—were classified into one major race.

Other respondents presented challenges, and the Census Bureau devised new procedures for their classification. A person who marked two Asian races, such as Chinese and Vietnamese, or two Pacific Islander races, such as Guamanian and Samoan, was classified as identifying with a single major race—Asian or NHOPI, respectively.

Although not consistent with OMB recommendations, the Census 2000 enumeration form included a “some other race” box where a respondent could write a word or phrase. The size of the multiple race population depended greatly upon the coding procedures the Census Bureau developed for coding entries written in the “some other race” box.

If a person wrote any term indicating a Hispanic origin, he or she was automatically considered to be “other” by race. For example, a person who wrote Castilian, Cuban, or Argentinean was classified into the “other” race category. If the respondent had already checked a box for his or her and then wrote a Hispanic term, he or she was multiple by race. Thus persons who marked white or black for their race and then went on to write Dominican or Puerto Rican in the “some other race” box were classified as multiracial. This explains why two of the most frequently reported combinations

Table 2. Population by Number of Races Reported, 2000 Census

| <i>Number of Races Reported</i> | <i>Number</i> | <i>Percentage of Total Population</i> |
|---------------------------------|---------------|---------------------------------------|
| One only | 274,595,678 | 97.4 |
| Two | 6,368,075 | 2.3 |
| Three | 410,285 | 0.1 |
| Four | 38,408 | <0.1 |
| Five | 8,637 | <0.1 |
| Six | 823 | <0.1 |
| Total | 281,421,906 | 100.0 |

Source: U.S. Bureau of the Census 2001, table 1.

were white and “some other race” and black and “some other race.”

If a person wrote a term in the “some other race” box indicating an origin other than Hispanic, the Census Bureau may or may not have classified that individual into the “other” race category. The agency considered the word or phrase written on the 2000 Census form and then examined data for persons who wrote that same word or phrase for their ancestry in the enumeration of 1990. If 70 percent or more of the persons reporting that ancestry in 1990 identified with a specific race, that race was assigned to the 2000 respondent. If less than 70 percent of those reporting that ancestry in 1990 had identified with a race, then the person was considered “other” by race in 2000. Several examples illustrate this procedure. Those who wrote French or Irish in the “some other race” box on the Census 2000 form were classified as white since more than 70 percent of those reporting French or Irish as their ancestry in 1990 had marked white for their race. Those who wrote South African in the “some other race” box in Census 2000 were considered to be “other” by race since, in the 1990 census, those writing South African for their ancestry reported a variety of races: white, black, and Asian Indian. These coding rules mean that if a person marked white in 2000 and then went on to write “Spanish” in the “some other race” box, he or she was multiracial just as was a person who checked black and then wrote “Dominican” in the “some other race” box. However, a person who checked white on the Census 2000 and then

wrote “French” or “Italian” in the “some other race” box was not multiple by race.

Identifying with Multiple Races: How Many Used This Option?

There were no instructions on the census form telling a respondent what to do to be tabulated as multiracial. Thus, the reported size of the multiracial population largely results from Census Bureau coding rules.

As table 2 indicates, 6.8 million of the 281.4 million Americans counted in 2000 marked two or more boxes or wrote a Hispanic-origin term for their second race, such that they were classified as identifying with two or more of the six major races used in the census. Approximately one American in forty identified with two races, whereas about one in one thousand marked boxes to identify with three or more races. Fewer than one thousand persons checked sufficient numbers of boxes to identify with all six major races.

Table 3 lists the racial combinations reported by at least one hundred thousand in 2000; the pie chart in figure 1 illustrates the distribution of the multirace population. The most popular combination—reported by one-third of those who identified with two or more races—was white and “some other race.” Typically, these were people who marked both the box for white and the box for “some other race” and then wrote in a term denoting Hispanic origin. White and American Indian was the next most popular combination, followed by white and Asian. The most commonly

Table 3. Frequency of Report of Multiple-Race Ancestry, 2000 Census

| <i>Racial Combination</i> | <i>Number Reported</i> | <i>Share of the Total Multiple-Race Population (Percentage)</i> |
|--|------------------------|---|
| Most Frequently Reported | | |
| White and some other race | 2,206,251 | 32.3 |
| White and American Indian | 1,082,683 | 15.9 |
| White and Asian | 868,395 | 12.7 |
| White and black | 784,764 | 11.5 |
| Black and some other race | 417,249 | 6.1 |
| Asian and some other race | 249,108 | 3.6 |
| Black and American Indian | 182,494 | 2.7 |
| Asian and NHOPI | 138,802 | 2.0 |
| White and NHOPI | 112,964 | 1.7 |
| White, black, and American Indian | 112,207 | 1.6 |
| Black and Asian | 106,782 | 1.6 |
| All other combinations | 564,529 | 8.3 |
| Least Frequently Reported | | |
| White–Black–Indian–NHOPI—some other race | 68 | <0.1 |
| Black–Indian–NHOPI—some other race | 111 | <0.1 |
| Indian–Asian–NHOPI—some other race | 207 | <0.1 |
| Black–Indian–Asian–NHOPI—some other race | 216 | <0.1 |
| White–Indian–NHOPI—some other race | 309 | <0.1 |
| White–black–NHOPI—some other race | 325 | <0.1 |
| Black–Indian–Asian—some other race | 334 | <0.1 |
| White–black–Asian–NHOPI—some other race | 379 | <0.1 |
| Indian–NHOPI—some other race | 586 | <0.1 |
| White–Indian–Asian–NHOPI—some other race | 639 | <0.1 |
| Total | 6,826,228 | 100.0 |

Source: U.S. Bureau of the Census 2001.

Note: NHOPI refers to the major racial group of Native Hawaiians and other Pacific Islanders. Ninety-seven percent of those who marked “Some other race” as their only race went on to identify with a Spanish-Hispanic-Latino origin on the separate Hispanic-origin inquiry. Fifty-eight percent of those who marked “Some other race” in combination with one or more other major races went on to identify with a Spanish-Hispanic-Latino origin.

reported second race for those who marked the box for black was also “some other race,” just as it was for whites. Subsequent data from the census will allow us to confirm the hypothesis that many of the individuals who identified themselves as black and “some other race” were Caribbeans or Central Americans who identified with their African origin and then wrote a Spanish term, such as Dominican. The only three-race combination marked by one hundred thousand or more respondents was white, black, and American Indian—reflecting, undoubtedly, the high rates at which American Indians have married outside their own race since the arrival of the first Europeans.

The Census of 2000 provides counts for 6 major race designations (including “some other”) and for all possible combinations of multiple reporting. There are 57 such possible combinations of 2 to 6 races—implying that this enumeration gives us counts for 63 distinct races or combinations of races down to the lowest level of census geography, the city block. When these 63 races are tabulated according to reported Hispanic ethnicity, data are available for 126 different groups.

Table 3 also lists the ten least frequently reported races. Some people identified with every one of the combinations, but fewer than ten thousand persons identified with 35 of the

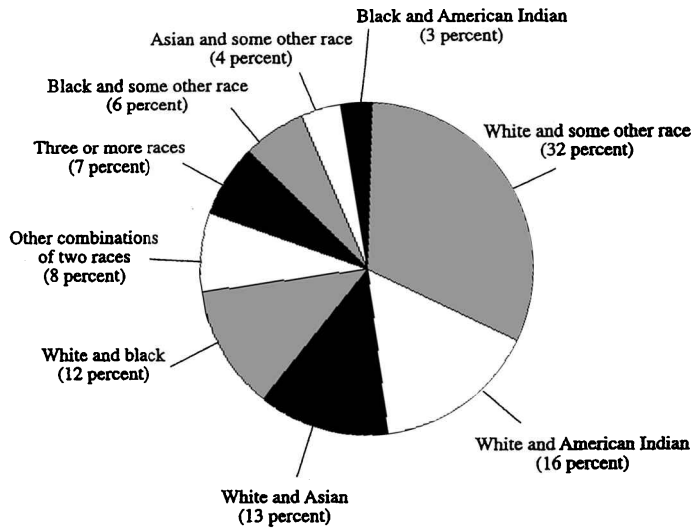


Figure 1. Distribution of the Multiple-Race Population, 2000

Source: U.S. Bureau of the Census 2001.

57 possible multiple race groups, and fewer than one thousand persons checked boxes to identify with 14 of the combinations. As this table shows, the combination white–black–American Indian–NHOPI–some other race was chosen least frequently, with a national count of just sixty-eight persons. Most of the 57 multiple-race combinations were reported so infrequently that the groups will not be subject to scholarly analysis.

Maximum and Minimum Counts of the Race in 2000

The 2000 census data generate thirty-two different estimates of the size of each of the major races; that is, it provides a count of those who identified with each race only, five different counts for the two-race combination involving each race, ten different counts for both three-race and four-race combinations, five different counts for five-race combinations, and, finally, the count that includes persons who claimed all six major races. No longer is there an unambiguous answer to the question, how many whites, or how many American Indians, did the census count.

Figure 2 presents a comparison of the minimum and maximum counts of each of the six major race categories. There are substantial differences between these counts. Among those who marked a box to report they were white by race, for example, few went on to identify with a second race, and so the minimum count of whites is close to the maximum count. Among Native Hawaiians and other Pacific Islanders, however, there were many more who identified with a second or third race than who identified as NHOPI only.

Whites and African Americans differed from the other major races in that those who identified as white or black were relatively unlikely to identify with a second race—thus the minimum and maximum counts are fairly similar. Although 5.2 million Americans marked both white and a second race, 97.4 percent of those who marked white claimed white as their only race. Among those who marked black, 95.2 percent marked only black–African American–Negro.

The “some other race” group presented in figure 2 is largely—but not entirely—made up of Latinos; 90.4 percent of those who checked the “some other race” box went on

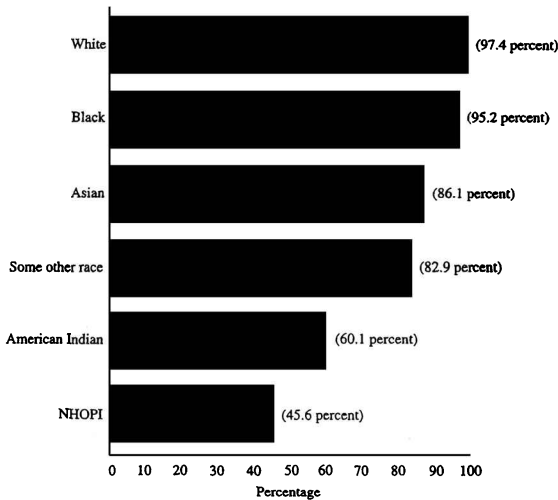


Figure 2. Minimum Counts of Census Race Groups—Races as Percentage of Maximum Counts, 2000

Source: U.S. Bureau of the Census 2001.

to claim a Spanish-Hispanic-Latino identity on the distinct Hispanic-ethnicity question. Of those who checked only the “some other race” box, 97 percent marked themselves as Hispanic in origin.

American Indians and the NHOPI races were distinguished by the frequency with which they were reported in combination with other races. The estimates of the American Indian population are spread across a broad range in the census of 2000: from 2.5 million who identified uniquely with this race to a maximum of 4.1 million, if those who marked Indian along with a second or third race are included in the count. . . .

Conclusion

Putting this new conceptualization of race into practice on the self-administered census form presented a formidable challenge to the Census Bureau. Their decision was to classify all persons into one or more of six race categories: white, black, American Indian, Asian, Native Hawaiian and other Pacific Islander, and “some other race.”

Current regulations of the Office of Management and Budget strongly recommend that

the Hispanic-origin question be asked, in addition to the mandatory race question. This may confuse respondents and certainly adds a burden for users of data: approximately 40 percent of those who identified with a Hispanic origin in both the 1990 and 2000 censuses did not mark a box for any one of the major races. Instead, they wrote a term denoting a Hispanic origin for their race and were thereby classified by the Census Bureau as “some other race.”

In the census of 2000, approximately 2.4 percent of the total population—one person in forty—identified with two or more races. This number is misleading, however, because it includes those who wrote in a term denoting Hispanic origin for their second race. If those who wrote in a Hispanic-origin term are excluded, approximately 1.6 percent of the population identified with two or more of the five major races—that is, about one person in sixty-five.

Multiple-race reporting was more common among the young, reflecting the demographic consequences of increasing interracial marriage. Among those under the age of eighteen, 4.4 percent identified with two or more races; among those eighteen years and older, that figure fell to 1.9 percent. If those who were counted as multiple in race because they wrote in a Hispanic-origin term for their second race are excluded, then the multiple-race population accounted for 3.2 percent of the population under the age of eighteen and 1.4 percent for their elders. Stated differently, about one American child in thirty-one was reported to be multiracial, compared with one in seventy-one for the population eighteen and older.

The reporting of multiple races displays strong geographic patterns. It is most common in Hawaii, Alaska, and California. Given the large black and white populations, the history of miscegenation, and a significant literature describing generations of mulattoes, one might expect multiple-race reporting to be high in the southern states. Instead, such reporting was lowest in the Deep South. . . .

The social movement that led to the change in racial identification has not been active in the years since census data were collected. It has faded. Because of apparent increases in marriages across racial lines, however, there will inevitably be a growing population of children whose parents differ by race. At this point, we do not know whether the current OMB racial classification system will be readily used by these individuals or whether many of them may identify with the race of one parent in some circumstances and that of the other parent in other circumstances. The traditional system of categorizing persons according to membership in a unique race is no longer accurate nor acceptable, but it is far from certain that the approach used in Census 2000 is practical, ideal, or useful (Skerry 2000).

We have yet to determine whether the concept “multiracial” will enter popular discourse. I believe that it will not, and that into the foreseeable future, most people will continue to assume that every American has a “basic” or “essential” race.

The multiple-race option is a dramatic change in the federal statistical system, but there is a telling demographic analogy. From 1870 through 1970 the census asked the birthplaces of respondents’ parents. After the restrictive immigration laws of the early twentieth century went into effect, the second-generation population declined, and in 1980, an open-ended question about ancestry supplanted the query on parents’ birthplaces. Respondents were encouraged to write in a term or terms to identify their ethnicity, ancestry, or national origin. The Census Bureau coded the first two terms unless the individual had entered a religious identity. At first glance, one might assume that analysts would exploit these new ancestry data to describe patterns of ethnic intermarriage, the nation’s ethnic heterogeneity, and the way ancestry or national origin was

linked to achievement. Few scholars, however, have analyzed ancestry and multiple-ancestry information from the last three censuses, and as far as I know, these data have not been used by governmental agencies.

REFERENCES

- Farley, Reynolds. 1999. “Racial Issues: Recent Trends in Residential Patterns and Intermarriage.” In *Diversity and Its Discontents: Cultural Conflict and Common Ground in Contemporary American Society*, edited by Neil J. Smelser and Jeffrey C. Alexander. Princeton, N.J.: Princeton University Press.
- Fu, Vincent Kang. 2001. “Racial Intermarriage Pairings.” *Demography* 38(2): 147–59.
- Gilberston, Greta, Joseph Fitzpatrick, and Lijun Yang. 1996. “Hispanic Intermarriage in New York City: New Evidence from 1991.” *International Migration Review* 30(summer): 445–59.
- Heer, David. 1974. “The Prevalence of Black-White Intermarriage in the United States: 1960 and 1970.” *Journal of Marriage and the Family* 36(2): 246–58.
- Hwang, Sean-Smong, Rogelio Saenz, and Benigno Aguirre. 1995. “The SES Selectivity of Interracially Married Asians.” *International Migration Review* 29(summer): 469–91.
- Kalmijn, Matthijs. 1993. “Trends in Black-White Intermarriage.” *Social Forces* 72(1): 19–146.
- . 1998. “Intermarriage and Homogamy: Causes, Patterns, and Trends.” *Annual Review of Sociology* 24: 395–421.
- Model, Suzanne, and Gene Fisher. 2001. “Black-White Unions: West Indians and African Americans Compared.” *Demography* 38(2): 177–85.
- Qian, Zhenchao. 1997. “Breaking the Racial Barriers: Variations in Interracial Marriage Between 1980 and 1990.” *Demography* 34(3): 263–76.
- Rosenfield, Michael J. 2001. “The Salience of Pan-National Hispanic and Asian Identities in U.S. Marriage Markets.” *Demography* 38(2): 161–75.
- Sandefur, Gary, and Trudy McKinnell. 1986. “American Indian Intermarriage.” *Social Science Research* 15(4): 347–71.
- Schoen, Robert, and B. Thomas. 1989. “Intergroup Marriage in Hawaii: 1969–1971 and 1979–1981.” *Sociological Perspectives* 32: 365–82.
- Skerry, Peter. 2000. *Counting on the Census? Race, Group Identity, and the Evasion of Politics*. Washington, D.C.: Brookings Institution.

Modes of Incorporation

■ CLASSIC MODES ■

66 ■ Edna Bonacich

A Theory of Ethnic Antagonism

The Split Labor Market

Societies vary considerably in their degree of ethnic and racial antagonism. Such territories as Brazil, Mexico, and Hawaii are generally acknowledged to be relatively low on this dimension; while South Africa, Australia, and the United States are considered especially high. Literally hundreds of variables have been adduced to account for these differences, ranging from religions of dominant groups, to whether the groups who migrate are dominant or subordinate, to degrees of difference in skin color, to an irreducible “tradition” of ethnocentrism. While some writers have attempted to synthesize or systematize some subset of these (e.g., Lieberman, 1961; Mason, 1970; Noel, 1968; Schermerhorn, 1970; van den Berghe, 1966), one is generally struck by the absence of a developed theory accounting for variations in ethnic antagonism.

One approach to this problem is to consider an apparent anomaly, namely that ethnic antagonism has taken two major, seemingly antithetical forms: exclusion movements, and so-called caste systems.¹ An example of the former is the “white Australia” policy; while

South Africa’s color bar illustrates the latter. The United States has shown both forms, with a racial caste system in the South and exclusion of Asian and “new” immigrants² from the Pacific and eastern seaboard respectively. Apart from manifesting antagonism between ethnic elements, exclusion and caste seem to have little in common. In the one, an effort is made to prevent an ethnically different group from being part of the society. In the other, an ethnically different group is essential to the society: it is an exploited class supporting the entire edifice. The Deep South felt it could not survive without its black people; the Pacific coast could not survive with its Japanese. This puzzle may be used as a touchstone for solving the general problem of ethnic antagonism, for to be adequate a theory must be able to explain it.

The theory presented here is, in part, a synthesis of some of the ideas used by Oliver Cox to explain the Japanese-white conflict on the U.S. Pacific coast (Cox, 1948:408–22), and by Marvin Harris to analyze the difference between Brazil and the Deep South in rigidity of the “color line” (Harris, 1964:79–94). It

Edna Bonacich. “A Theory of Ethnic Antagonism: The Split Labor Market,” in *American Sociological Review* (October 1972), pp. 547–559. Copyright © 1972 by the American Sociological Association. Used by permission of the American Sociological Association and the author.

stresses the role of a certain kind of economic competition in the development of ethnic antagonism. Economic factors have, of course, not gone unnoticed, though until recently sociological literature has tended to point them out briefly, then move on to more “irrational” factors (even such works as *The Economics of Discrimination*, Becker, 1957). A resurgence of Marxian analysis (e.g., Blauner, 1969; Reich, 1971) has thrust economic considerations to the fore, but I shall argue that even this approach cannot adequately deal with the problem posed by exclusion movements and caste systems. In addition, both Marxist and non-Marxist writers assume that racial and cultural differences in themselves prompt the development of ethnic antagonism. This theory challenges that assumption, suggesting that economic processes are more fundamental.

No effort is made to prove the accuracy of the following model. Such proof depends on a lengthier exposition. Historical illustrations are presented to support it.

Ethnic Antagonism

“Ethnic” rather than “racial” antagonism was selected as the dependent variable because the former is seen to subsume the latter. Both terms refer to groups defined socially as sharing a common ancestry in which membership is therefore inherited or ascribed, whether or not members are currently physically or culturally distinctive.³ The difference between race and ethnicity lies in the size of the locale from which a group stems, races generally coming from continents, and ethnicities from national sub-sections of continents. In the past the term “race” has been used to refer to both levels, but general usage today has reversed this practice (e.g., Schermerhorn, 1970; Shibutani and Kwan, 1965). Ethnicity has become the generic term.

Another reason for choosing this term is that exclusion attempts and caste-like arrangements are found among national groupings within a racial category. For exam-

ple, in 1924 whites (Europeans) attempted to exclude whites of different national backgrounds from the United States by setting up stringent immigration quotas.

The term “antagonism” is intended to encompass all levels of intergroup conflict, including ideologies and beliefs (such as racism and prejudice), behaviors (such as discrimination, lynchings, riots), and institutions (such as laws perpetuating segregation). Exclusion movements and caste systems may be seen as the culmination of many pronouncements, actions, and enactments, and are continuously supported by more of the same. “Antagonism” was chosen over terms like prejudice and discrimination because it carries fewer moralistic and theoretical assumptions (see Schermerhorn, 1970:6–9). For example, both of these terms see conflict as emanating primarily from one side: the dominant group. Antagonism allows for the possibility that conflict is mutual; i.e. a product of interaction.

The Split Labor Market

The central hypothesis is that ethnic antagonism first germinates in a labor market split along ethnic lines. To be split, a labor market must contain at least two groups of workers whose price of labor differs for the same work, or would differ if they did the same work. The concept “price of labor” refers to labor’s total cost to the employer, including not only wages, but the cost of recruitment, transportation, room and board, education, health care (if the employer must bear these), and the costs of labor unrest. The degree of worker “freedom” does not interfere with this calculus; the cost of a slave can be estimated in the same monetary units as that of a wage earner, from his purchase price, living expenses, policing requirements, and so on.

The price of a group of workers can be roughly calculated in advance and comparisons made even though two groups are not engaged in the same activity at the same

Table 1. Estimated Cost of Three Types of Labor to Be Shepherds in New South Wales, 1841

| | <i>Free Man (White)</i> | | | <i>Prisoner (White)</i> | | | <i>Coolie (Indian)</i> | | |
|--------------------|-----------------------------|----|----|-----------------------------|----|----|----------------------------|----|----|
| | £ | s. | d. | £ | s. | d. | £ | s. | d. |
| Rations | 16 | 18 | 0 | 13 | 14 | 4 | 9 | 6 | 4 |
| Clothing | — | — | — | 3 | 3 | 0 | 1 | 1 | 8 |
| Wages | 25 | 0 | 0 | — | — | — | 6 | 0 | 0 |
| Passage from India | — | — | — | — | — | — | 2 | 0 | 0 |
| Total per Annum | 41 | 18 | 0 | 16 | 17 | 4 | 18 | 8 | 0 |

From Yarwood (1968:13)

time. Thus in 1841 in the colony of New South Wales, the Legislative Council's Committee on Immigration estimated the relative costs of recruiting three groups of laborers to become shepherds. Table 1 shows their findings. The estimate of free white labor, for example, was based on what it would take to attract these men from competing activities.

Factors Affecting the Initial Price of Labor

Labor markets that are split by the entrance of a new group develop a dynamic which may in turn affect the price of labor. One must therefore distinguish initial from later price determinants. The initial factors can be divided into two broad categories: resources and motives.

1. Resources

Three types of resources are important price determinants. These are:

a. Level of Living, or Economic Resources.

The ethnic groups forming the labor market in a contact situation derive from different economic systems, either abroad or within a conquered territory. For members of an ethnic group to be drawn into moving, they must at least raise their wage level. In general, the poorer the economy of the recruits, the less the inducement needed for them to

enter the new labor market. Crushing poverty may drive them to sell their labor relatively cheaply. For example, Lind (1968:199) describes the effect of the living level on the wage scale received by immigrant workers to Hawaii:

In every case [of labor importations] the superior opportunities for gaining a livelihood have been broadcast in regions of surplus manpower, transportation facilities have been provided, and finally a monetary return larger than that already received has been offered to the prospective laborer. The monetary inducement has varied considerably, chiefly according to the plane of living of the population being recruited, and the cheapest available labor markets have, of course, been most extensively drawn upon.

Workers need not accept the original wage agreement for long after they have immigrated, since other opportunities may exist; for instance, there may be ample, cheap land available for individual farming. One capitalist device for keeping wages low at least for a time is to bind immigrants to contracts before they leave the old economy. The Indian indenture system, for example, rested on such an arrangement (Gillion, 1962:19–38).

b. Information. Immigrants may be pushed into signing contracts out of igno-

rance. They may agree to a specific wage in their homeland not knowing the prevailing wage in the new country, or having been beguiled by a false account of life and opportunity there. Williams (1944:11), for example, describes some of the false promises made to draw British and Germans as workers to West Indian sugar plantations before the advent of African slavery. Chinese labor to Australia was similarly “obtained under ‘false and specious pretences’” (Willard, 1967:9).

The possibilities for defrauding a population lacking access to the truth are obvious. In general, the more people know about conditions obtaining in the labor market to which they are moving, the better can they protect themselves against disadvantageous wage agreements.

c. Political Resources. By political resources I mean the benefits to a group of organizing. Organization can exist at the level of labor, or it can occur at higher levels, for example, in a government that protects them. These levels are generally related in that a strong government can help organize its emigrants. There are exceptions, however: strong emigrant governments tend not to extend protection to their deported convicts or political exiles; and some highly organized groups, like the Jews in the United States, have not received protection from the old country.

Governments vary in the degree to which they protect their emigrants. Japan kept close watch over the fate of her nationals who migrated to Hawaii and the Pacific coast; and the British colonial government in India tried to guard against abuses of the indenture system (for example, by refusing to permit Natal to import Indian workers for their sugar plantations until satisfactory terms had been agreed to; cf. Ferguson-Davie, 1952:4–10). In contrast Mexican migrant workers to the United States have received little protection from their government, and African states were unable to intervene on behalf of slaves brought to America. Often the indigenous populations

of colonized territories have been politically weak following conquest. Thus African nations in South Africa have been unable to protect their migrant workers in the cities.

In general, the weaker a group politically, the more vulnerable it is to the use of force, hence to an unfavorable wage bargain (or to no wage bargain at all, as with slavery). The price of a labor group varies inversely with the amount of force that can be used against it, which in turn depends on its political resources.

2. Motives

Two motives affect the price of labor, both related to the worker’s intention of not remaining permanently in the labor force. Temporary workers tend to cost less than permanent workers for two reasons. First, they are more willing to put up with undesirable work conditions since these need not be endured forever. If they are migrants, this tolerance may extend to the general standard of living. Often migrant temporary workers are males who have left the comforts of home behind and whose employers need not bear the cost of housing and educating their families. Even when families accompany them, such workers tend to be willing to accept a lower standard of living since it is only short term.

Second, temporary workers avoid involvement in lengthy labor disputes. Since they will be in the labor market a short while, their main concern is immediate employment. They may be willing to undercut wage standards if need be to get a job and are therefore ripe candidates for strikebreaking. Permanent workers also stand to lose from lengthy conflict, but they hope for benefits to their progeny. If temporary workers are from elsewhere, they have no such interest in future business-labor relations. Altogether, temporary workers have little reason to join the organizations and unions of a permanent work force, and tend not to do so.

a. Fixed or Supplementary Income Goal. Some temporary workers enter the market

either to supplement family income, or to work toward a specific purchase. The worker's standard of living does not, therefore, depend on his earnings on the job in question, since his central source of employment or income lies elsewhere. Examples of this phenomenon are to be found throughout Africa:

. . . the characteristic feature of the labor market in most of Africa has always been the massive circulation of Africans between their villages and paid employment outside. In some places villagers engage in wage-earning seasonally. More commonly today they work for continuous though short-term periods of roughly one to three years, after which they return to the villages. . . . the African villager, the potential migrant into paid employment, has a relatively low, clearly-defined and rigid income goal; he wants money to pay head and hut taxes, to make marriage payments required of prospective bridegrooms, or to purchase some specific consumer durable (a bicycle, a rifle, a sewing machine, a given quantity of clothing or textiles, etc.) (Berg, 1966:116-8).

Such a motive produces the “backward-sloping labor supply function” characteristic of many native peoples in colonized territories. In addition to the general depressing effects on wages of being temporary, this motive leads to a fairly rapid turnover in personnel, making organization more difficult and hindering the development of valuable skills which could be used for bargaining. If wages were to rise, workers would reach their desired income and withdraw more quickly from the market, thereby lessening their chances of developing the political resources necessary to raise their wages further.

b. Fortune Seeking. Many groups, commonly called sojourners (see Siu, 1952), migrate long distances to seek their fortune, with the ultimate intention of improving their position in their homeland. Such was the case

with Japanese immigrants on the west coast and Italian immigrants in the east. Such workers stay longer in the labor market, and can develop political resources. However, since they are temporary they have little incentive to join the organizations of the settled population. Instead they tend to create competing organizations composed of people who will play a part in their future in the homeland, i.e. members of the same ethnic group.

Sojourner laborers have at least three features which affect the price of labor: lower wages, longer hours, and convenience to the employer. The Japanese show all three. Millis (1915:45) cites the U.S. Immigration Commission on the question of relative wages:

The Japanese have usually worked for a lower wage than the members of any other race save the Chinese and the Mexican. In the salmon canneries the Chinese have been paid higher wages than the Japanese engaged in the same occupations. In the lumber industry, all races, including the East Indian, have been paid higher wages than the Japanese doing the same kind of work. As section hands and laborers in railway shops they have been paid as much or more than the Mexicans, but as a rule less than the white men of many races.

And so on. The lower wage level of Japanese workers reflects both a lower standard of living, and a desire to get a foothold in the labor market. As Iwata (1962:27) puts it: “Their willingness to accept even lower wages than laborers of other races enabled the Japanese to secure employment readily.”

Millis (1915:155) describes a basket factory in Florin, California, where Japanese workers had displaced white female workers because the latter were unwilling to work more than ten hours a day or on weekends. The Japanese, anxious to return to Japan as quickly as possible, were willing to work twelve to fourteen hours per day and on weekends, thereby saving their employers the costs of a special overtime work force.

The Japanese immigrants developed political resources through a high degree of community organization. This could be used for the convenience of the employer, by solving his recruitment problems, seeing that work got done, and providing workers with board and lodging. In the case of seasonal labor, the Japanese community could provide for members during the off-season by various boarding arrangements and clubs, and by transporting labor to areas of demand (Ichihashi, 1932:172–6; Millis, 1915:44–5). These conveniences saved the employer money.⁴

As the reader may have noted, I have omitted a factor usually considered vital in determining the price of labor, i.e. differences in skills. I would contend, however, that this does not in itself lead to that difference in price for the same work which distinguishes a split labor market. While a skilled worker may be able to get a higher paying job, an unskilled laborer of another ethnicity may be trained to fill that job for the same wage. Skills are only indirectly important in that they can be used to develop political resources, which in turn may lead to a difference in wage level for the same work.

Price of Labor and Ethnicity

Ethnic differences need not always produce a price differential. Thus, if several ethnic groups who are approximately equal in resources and/or goals enter the same economic system, a split labor market will not develop. Alternatively, in a two-group contact situation, if one ethnic group occupies the position of a business elite and has no members in the labor force (or in a class that could easily be pushed into the labor force, e.g. low-capital farmers) then regardless of the other group's price, the labor market will not be split. This statement is a generalization of the point made by Harris (1964) that the critical difference in race relations between the deep south and Brazil was that the former had a white yeomanry in direct com-

petition with ex-slaves, while the Portuguese only occupied the role of a business elite (plantation owners).

Conversely, a split labor force does not only stem from ethnic differences. For example, prison and female labor have often been cheaper than free male labor in Western societies. Prison labor has been cheap because prisoners lack political resources, while women often labor for supplementary incomes (cf. Hutchinson, 1968:59–61; Hene-man and Yoder, 1965:543–4).

That initial price discrepancies in labor should ever fall along ethnic lines is a function of two forces. First, the original wage agreement arrived at between business and new labor often takes place in the labor group's point of origin. This is more obviously a feature of immigrant labor, but also occurs within a territory when conquered peoples enter their conquerors' economy. In other words, the wage agreement is often concluded within a national context, these nationalities coming to comprise the ethnic elements of the new labor market. One would thus expect the initial wages of co-nationals to be similar.

Second, nations or peoples that have lived relatively separately from one another are likely to have developed different employment motives and levels of resources (wealth, organization, communication channels). In other words, the factors that affect the price of labor are likely to differ grossly between nations, even though there may be considerable variation within each nation, and overlap between nations. Color differences in the initial price of labor only seem to be a factor because resources have historically been roughly correlated with color around the world.⁵ When color and resources are not correlated in the "expected" way, then I would predict that price follows resources and motives rather than color.

In sum, the prejudices of business do not determine the price of labor, darker skinned or culturally different persons being paid less

because of them. Rather, business tries to pay as little as possible for labor, regardless of ethnicity, and is held in check by the resources and motives of labor groups. Since these often vary by ethnicity, it is common to find ethnically split labor markets.

The Dynamics of Split Labor Markets

In split labor markets, conflict develops between three key classes: business, higher paid labor, and cheaper labor. The chief interests of these classes are as follows:

1. Business or Employers

This class aims at having as cheap and docile a labor force as possible to compete effectively with other businesses. If labor costs are too high (owing to such price determinants as unions), employers may turn to cheaper sources, importing overseas groups or using indigenous conquered populations. In the colony of Queensland in Australia, for example, it was believed that cotton farming would be the most suitable economic enterprise:

However, such plantations (being too large) could not be worked, much less cleared, by their owners; neither could the work be done by European laborers because sufficient numbers of these were not available—while even had there been an adequate supply, the high rates of wages would have been prohibitive. This was a consideration which assumed vast importance when it was realized that cotton would have to be cultivated in Queensland at a considerably lower cost than in the United States in order to compensate for the heavier freights from Queensland—the more distant country from England. It seemed then that there was no possibility of successful competition with America unless the importation of some form of cheap labor was permitted. (Moles, 1968:41)

Cheaper labor may be used to create a new industry having substantially lower labor

costs than the rest of the labor market, as in Queensland. Or they may be used as strike-breakers or replacements to undercut a labor force trying to improve its bargaining position with business. If cheap labor is unavailable, business may turn to mechanization, or try to relocate firms in areas of the world where the price of labor is lower.

2. Higher Paid Labor

This class is very threatened by the introduction of cheaper labor into the market, fearing that it will either force them to leave the territory or reduce them to its level. If the labor market is split ethnically, the class antagonism takes the form of ethnic antagonism. It is my contention (following Cox, 1948:411n) that, while much rhetoric of ethnic antagonism concentrates on ethnicity and race, it really in large measure (though probably not entirely) expresses this class conflict.

The group comprising higher paid labor may have two components. First, it may include current employees demanding a greater share of the profits or trying to maintain their position in the face of possible cuts. A second element is the small, independent, entrepreneur, like the subsistence farmer or individual miner. The introduction of cheaper labor into these peoples' line can undermine their position, since the employer of cheaper labor can produce at lower cost. The independent operator is then driven into the labor market. The following sequence occurs in many colonies: settlement by farmers who work their own land, the introduction of intensive farming using cheaper labor, a rise in land value and a consequent displacement of independent farmers. The displaced class may move on (as occurred in many of the West Indies when African slave labor was introduced to raise sugar), but if it remains, it comes to play the role of higher paid labor.

The presence of cheaper labor in areas of the economy where higher paid labor is not currently employed is also threatening to the latter, since the former attract older indus-

tries. The importance of potential competition cannot be over-stressed. Oftentimes writers assert the irrationality of ethnic antagonism when direct economic competition is not yet in evidence owing to few competitors having entered the labor market, or to competitors having concentrated in a few industries. Thus Daniels (1966:29) belittles the role of trade unions in the Asiatic Exclusion League by describing one of the major contributors as “an organization whose members, like most trade unionists in California, were never faced with job competition from Japanese.” It does not take direct competition for members of a higher priced labor group to see the possible threat to their well-being, and to try to prevent its materializing. If they have reason to believe many more low-priced workers are likely to follow an initial “insignificant trickle” (as Daniels, 1966:1, describes the Japanese immigration, failing to mention that it was insignificant precisely because a larger anticipated flow had been thwarted, and diverted to Brazil), or if they see a large concentration of cheaper labor in a few industries which could easily be used to undercut them in their own, they will attempt to forestall undercutting.

Lest you think this fear misguided, take note that, when business could override the interests of more expensive labor, the latter have indeed been displaced or undercut. In British Guiana the local labor force, composed mainly of African ex-slaves, called a series of strikes in 1842 and 1847 against planters’ attempts to reduce their wages. Plantation owners responded by using public funds to import over 50,000 cheaper East Indian indentured workers (Despres, 1969). A similar situation obtained in Mississippi, where Chinese were brought in to undercut freed blacks. Loewen (1971:23) describes the thinking of white land-owners: “the ‘China-man’ would not only himself supply a cheaper and less troublesome work force but in addition his presence as a threatening alternative would intimidate the Negro into

resuming his former docile behavior.” Such displacement has occurred not only to non-white more expensive labor, but, as the effects of slavery in the West Indies show, to whites by white capitalists.

3. Cheaper Labor

The employer uses this class partly to undermine the position of more expensive labor, through strikebreaking and undercutting. The forces that make the cheaper group cost less permit this to occur. In other words, either they lack the resources to resist an offer or use of force by business, or they seek a quick return to another economic and social base.

With the possible exception of sojourners, cheaper labor does not intentionally undermine more expensive labor; it is paradoxically its weakness that makes it so threatening, for business can more thoroughly control it. Cox makes this point (1948:417–8) in analyzing why Pacific coast white and Asian workers could not unite in a coalition against business:

. . . the first generation of Asiatic workers is ordinarily very much under the control of labor contractors and employers, hence it is easier for the employer to frustrate any plans for their organization. Clearly this cultural bar helped antagonize white workers against the Asiatics. The latter were conceived of as being in alliance with the employer. It would probably have taken two or three generations before, say, the East Indian low-caste worker on the Coast became sufficiently Americanized to adjust easily to the policies and aims of organized labor.

Ethnic antagonism is specifically produced by the competition that arises from a price differential. An oversupply of equal-priced labor does not produce such antagonism, though it too threatens people with the loss of their job. However, hiring practices will not necessarily fall along ethnic lines, there being no advantage to the employer in hiring workers of one or another ethnicity. All workingmen are on

the same footing, competing for scarce jobs (cf. Blalock, 1967:84–92, who uses this model of labor competition). When one ethnic group is decidedly cheaper than another (i.e. when the labor market is split) the higher paid worker faces more than the loss of his job; he faces the possibility that the wage standard in all jobs will be undermined by cheaper labor.

Victory for More Expensive Labor

If an expensive labor group is strong enough (strength generally depending on the same factors that influence price), they may be able to resist being displaced. Both exclusion and caste systems represent such victories for higher paid labor.

1. Exclusion

Exclusion movements generally occur when the majority of a cheaper labor group resides outside a given territory but desires to enter it (often at the request of business groups). The exclusion movement tries to prevent the physical presence of cheaper labor in the employment area, thereby preserving a non-split, higher priced labor market.

There are many examples of exclusion attempts around the world. In Australia, for instance, a group of white workers was able to prevent capitalists from importing cheaper labor from India, China, Japan, and the Pacific Islands. Attempts at importation were met with strikes, boycotts, petitions, and deputations (Willard, 1967:51–7). Ultimately, organized white labor pressed for strong exclusion measures, and vigilantly ensured their enforcement. As Yarwood (1964:151–2) puts it: “A comparison of the records of various governments during our period [1896–1923] leaves no doubt as to the special role of the Labour Party as the guardian of the ports.” In other words, a white Australia policy (i.e. the exclusion of Asian and Polynesian immigrants) appears to have sprung from a conflict of interests between employers who wanted to import

cheap labor, and a labor force sufficiently organized to ward off such a move.

California’s treatment of Chinese and Japanese labor is another example of exclusion. A socialist, Cameron H. King, Jr., articulates the threatened labor group’s position:

Unskilled labor has felt this competition [from the Japanese] for some time being compelled to relinquish job after job to the low standard of living it could not endure. The unskilled laborers are largely unorganized and voiceless. But as the tide rises it is reaching the skilled laborers and the small merchants. These are neither unorganized nor voiceless, and viewing the menace to their livelihood they loudly demand protection of their material interests. We of the Pacific Coast certainly know that exclusion is an effective solution. In the seventh decade of the nineteenth century the problem arose of the immigration of Chinese laborers. The Republican and Democratic parties failed to give heed to the necessities of the situation and the Workingman’s party arose and swept the state with the campaign cry of “The Chinese must go.” Then the two old parties woke up and have since realized that to hold the labor vote they must stand for Asiatic exclusion. (King, 1908:665–6)

King wrote this around the time of the Gentlemen’s Agreement, an arrangement of the U.S. and Japanese governments to prevent further immigration of Japanese labor to the Pacific Coast (Bailey, 1934). The Agreement was aimed specifically at labor and not other Japanese immigrants, suggesting that economic and not racial factors were at issue.

Exclusion movements clearly serve the interests of higher paid labor. Its standards are protected, while the capitalist class is deprived of cheaper labor.

2. Caste

If cheaper labor is present in the market, and cannot be excluded, then higher paid labor

will resort to a caste arrangement, which depends on exclusiveness rather than exclusion. Caste is essentially an aristocracy of labor (a term borrowed from Lenin, e.g. 1964), in which higher paid labor deals with the undercutting potential of cheaper labor by excluding them from certain types of work. The higher paid group controls certain jobs exclusively and gets paid at one scale of wages, while the cheaper group is restricted to another set of jobs and is paid at a lower scale. The labor market split is submerged because the differentially priced workers ideally never occupy the same position.

Ethnically distinct cheaper groups (as opposed to women, for example, who face a caste arrangement in many Western societies) may reside in a territory for two reasons: either they were indigenous or they were imported early in capitalist-labor relations, when the higher paid group could not prevent the move. Two outstanding examples of labor aristocracies based on ethnicity are South Africa, where cheaper labor was primarily indigenous, and the U.S. South, where they were imported as slaves.

Unlike exclusion movements, caste systems retain the underlying reality of a price differential, for if a member of the subordinate group were to occupy the same position as a member of the stronger labor group he would be paid less. Hence, caste systems tend to become rigid and vigilant, developing an elaborate battery of laws, customs, and beliefs aimed to prevent undercutting. The victory has three facets. First, the higher paid group tries to ensure its power in relation to business by monopolizing the acquisition of certain essential skills, thereby ensuring the effectiveness of strike action, or by controlling such important resources as purchasing power. Second, it tries to prevent the immediate use of cheaper labor as undercutters and strikebreakers by denying them access to general education thereby making their training as quick replacements more difficult, or by ensuring through such devices as “influx control” that

the cheaper group will retain a base in their traditional economies. The latter move ensures a backward-sloping labor supply function (cf. Berg, 1966) undesirable to business. Third, it tries to weaken the cheaper group politically, to prevent their pushing for those resources that would make them useful as undercutters. In other words, the solution to the devastating potential of weak, cheap labor is, paradoxically, to weaken them further, until it is no longer in business' immediate interest to use them as replacements.

South Africa is perhaps the most extreme modern example of an ethnic caste system. A split labor market first appeared there in the mining industry. With the discovery of diamonds in 1869, a white working class emerged.⁶ At first individual whites did the searching, but, as with the displacement of small farms by plantations, they were displaced by consolidated, high-capital operations, and became employees of the latter (Doxey, 1961:18). It was this class together with imported skilled miners from Cornwall (lured to Africa by high wages) which fought the capitalists over the use of African labor. Africans were cheaper because they came to the mines with a fixed income goal (e.g. the price of a rifle) and did not view the mines as their main source of livelihood. By contrast, European workers remained in the mines and developed organizations to further their interests.

Clearly, it would have been to the advantage of businessmen, once they knew the skills involved, to train Africans to replace the white miners at a fraction of the cost; but this did not happen. The mining companies accepted a labor aristocracy, not out of ethnic solidarity with the white workers but:

(as was to be the case throughout the later history of mining) they had little or no choice because of the collective strength of the white miners. . . . The pattern which was to emerge was that of the Europeans showing every sign of preparedness to use their collective strength to ensure their exclusive supremacy in the

labour market. Gradually the concept of trade unionism, and, for that matter, of socialism, became accepted in the minds of the European artisans as the means of maintaining their own position against non-white inroads. (Doxey, 1961:23–4)

The final showdown between mine owners and white workers occurred in the 1920s when the owners tried to substitute cheaper non-white labor for white labor in certain semi-skilled occupations. This move precipitated the “Rand Revolt,” a general strike of white workers on the Witwatersrand, countered by the calling in of troops and the declaration of martial law. The result was a coalition between Afrikaner nationalists (predominantly workers and small-scale farmers being pushed off the land by larger, British-owned farms) and the English-speaking Labor Party (Van der Horst, 1965:117–8). The Revolt “showed the lengths to which white labour was prepared to go to defend its privileged position. From that time on, mine managements have never directly challenged the colour-bar in the mining industry” (Van der Horst, 1965:118).

The legislative history of much of South Africa (and of the post-bellum Deep South) consists in attempts by higher priced white labor to ward off undercutting by cheaper groups, and to entrench its exclusive control of certain jobs.⁷

This interpretation of caste contrasts with the Marxist argument that the capitalist class purposefully plays off one segment of the working class against the other (e.g. Reich, 1971). Business, I would contend, rather than desiring to protect a segment of the working class supports a liberal or laissez-faire ideology that would permit all workers to compete freely in an open market. Such open competition would displace higher paid labor. Only under duress does business yield to labor aristocracy, a point made in *Deep South*, a book written when the depression had caused the displacement of white tenant farmers and industrial workers by blacks:

The economic interests of these groups [employers] would also demand that cheaper colored labor should be employed in the “white collar” jobs in business offices, governmental offices, stores, and banks. In this field, however, the interests of the employer group conflict not only with those of the lower economic group of whites but also with those of the more literate and aggressive middle group of whites. A white store which employed colored clerks, for example, would be boycotted by both these groups. The taboo upon the employment of colored workers in such fields is the result of the political and purchasing power of the white middle and lower groups. (Davis, *et al.*, 1941:480)

In sum, exclusion and caste are similar reactions to a split labor market. They represent victories for higher paid labor. The victory of exclusion is more complete in that cheaper labor is less available to business. For this reason I would hypothesize that a higher paid group prefers exclusion to caste, even though exclusion means they have to do the dirty work. Evidence for this comes from Australia where, in early attempts to import Asian labor, business tried to buy off white labor’s opposition by offering to form them into a class of “mechanics” and foremen over the “coolies” (Yarwood, 1968:16, 42). The offer was heartily rejected in favor of exclusion. Apartheid in South Africa can be seen as an attempt to move from caste to the exclusion of the African work force.

Most of our examples have contained a white capitalist class, a higher paid white labor group, and a cheaper, non-white labor group. Conditions in Europe and around the world, and not skin color, yield such models. White capitalists would gladly dispense with and undercut their white working-class brethren if they could, and have done so whenever they had the opportunity. In the words of one agitator for excluding Chinese from the U.S. Pacific coast: “I have seen men . . . American born, who certainly would, if I

Table 2. Numerical Proportion of Dominant to Subordinate Ethnic Groups

| <i>Category</i> | | |
|------------------------|---------------------|----------------------|
| <i>Domination</i> | <i>Paternalism</i> | <i>Competition</i> |
| <i>Situations</i> | | |
| South Africa (1960) | Nigeria (1952) | Britain (1968) |
| 1-4 | 1-2000 | 50-1 |
| U. S. South (1960) | Nyasaland (1966) | U.S. North (1960) |
| 4-1 | 1-570 | 15-1 |
| Rhodesia (1960) | Tanganyika | New Zealand |
| 1-16 | 1-450 | 13-1 |
| | Uganda 1-650 | |

Adapted from Mason (1970:64)

may use a strong expression, employ devils from Hell if the devils would work for 25 cents less than a white man" (cited in Daniels and Kitano, 1970:43).

In addition, cases have occurred of white workers playing the role of cheap labor, and facing the same kind of ethnic antagonism as non-white workers. Consider the riots against Italian strikebreakers in the coal fields of Pennsylvania in 1874 (Higham, 1965:47-8). In the words of one writer: "Unions resented the apparently inexhaustible cheap and relatively docile labor supply which was streaming from Europe obviously for the benefit of their employers" (Wittke, 1953:10).

Even when no ethnic differences exist, split labor markets may produce ethnic-like antagonism. Carey McWilliams (1945:82-3) describes an instance:

During the depression years, "Old Stock"—that is, white, Protestant, Anglo-Saxon Americans, from Oklahoma, Arkansas, and Texas—were roundly denounced in California as "interlopers." The same charges were made against them that were made against the Japanese: they were "dirty"; they had "enormous families"; they engaged in unfair competition; they threatened to "invade" the state and to "undermine" its institutions. During

these turgid years (1930-1938) California attempted *to exclude*, by various extra-legal devices, those yeoman farmers just as it had excluded the Chinese and Japanese. "Okies" were "inferior" and "immoral." There was much family discord when Okie girl met California boy, and vice versa. . . . The prejudice against the Okies was obviously not "race" prejudice; yet it functioned in much the same manner.

Conclusion

Obviously, this type of three-way conflict is not the only important factor in ethnic relations. But it does help explain some puzzles, including, of course, the exclusion-caste anomaly. For example, Philip Mason (1970:64) develops a typology of race relations and finds that it relates to numerical proportions without being able to explain the dynamic behind this correlation. Table 2 presents a modified version of his chart. My theory can explain these relationships. Paternalism arises in situations where the cleavage between business and labor corresponds to an ethnic difference. A small business elite rules a large group of workers who entered the labor market at approximately the same price or strength. No split labor market

existed, hence no ethnic caste system arises. The higher proportion of the dominant ethnicity under “Domination” means that part of the dominant group must be working class. A labor element that shares ethnicity with people who have sufficient resources to become the business elite is generally likely to come from a fairly wealthy country and have resources of its own. Such systems are likely to develop split labor markets. Finally, competition has under it societies whose cheaper labor groups have not been a major threat because the indigenous population available as cheap labor has been small and/or exclusion has effectively kept business groups from importing cheap labor in large numbers.

This theory helps elucidate other observations. One is the underlying similarity in the situation of blacks and women. Another is the history of political sympathy between California and the South. And, a third is the conservatism of the American white working class, or what Daniels and Kitano (1970:45) consider to be an “essential paradox of American life: [that] movements for economic democracy have usually been violently opposed to a thorough-going ethnic democracy.” Without having to resort to psychological constructs like “authoritarianism,” this theory is able to explain the apparent paradox.

In sum, in comparing those countries with the most ethnic antagonism with those having the least, it is evident that the difference does not lie in the fact that the former are Protestant and the latter Catholic: Protestants are found in all three of Mason’s types, and Hawaii is a Protestant-dominated territory. It does not lie in whether the dominant or subordinate group moves: South Africa and the Deep South show opposite patterns of movement. It is evident that some of the most antagonistic territories have been British colonies, but not all British colonies have had this attribute. The characteristic

that those British colonies and other societies high on ethnic antagonism share is that they all have a powerful white, or more generally higher paid, working class.

NOTES

1. I do not wish to enter the debate over the applicability of the term “caste” to race relations (cf. Cox, 1948; Davis, *et al.*, 1941). It is used here only for convenience and implies no particular theoretical bent.

2. The term “exclusion” has not usually been applied to immigrant quotas imposed on eastern and southern European immigrants; but such restrictions were, in effect, indistinguishable from the restrictions placed on Japanese immigration.

3. This usage contrasts with that of van den Berghe (1967a:9–10) who reserves the term “ethnic” for groups socially defined by cultural differences. In his definition, ethnicity is not necessarily inherited. I would contend that, while persons of mixed ancestry may be problematic and are often assigned arbitrarily by the societies in which they reside, inheritance is implied in the common application of the word.

4. Sojourners often use their political resources and low price of labor to enter business for themselves (a process which will be fully analyzed in another paper). This does not remove the split in the labor market, though it makes the conflict more complex.

5. It is, of course, no accident that color and resources have been historically related. Poverty among non-white nations has in part resulted from European imperialism. Nevertheless, I would argue that the critical factor in the development of ethnic segmentation in a country is the meeting that occurs in the labor market of that country. The larger economic forces help determine the resources of entering parties, but it is not such forces to which workers respond. Rather they react to the immediate conflicts and threats in their daily lives.

6. Such a split was not found in the early Cape Colony, where business was one ethnicity—white—and labor another—non-white. Actually in neither case was the ethnic composition simple or homogeneous; but the important fact is that, among the laborers, who included so-called Hottentots, and slaves from Madagascar, Mozambique, and the East Indies (cf. van den Berghe, 1967b:14), no element was significantly more expensive. The early Cape is thus structurally similar, in terms of the variables I

consider important, to countries like Brazil and Mexico. And it is also noted for its “softened” tone of race relations as reflected in such practices as intermarriage.

7. Ethnically based labor aristocracies are much less sensitive about cheap labor in any form than are systems that do not arrive at this resolution because they are protected from it. Thus, Sutherland and Cressey (1970:561–2) report that both the Deep South and South Africa continue to use various forms of prison contract labor, in contrast to the northern U.S. where the contract system was attacked by rising labor organizations as early as 1880.

REFERENCES

- Bailey, Thomas A. 1934. *Theodore Roosevelt and the Japanese-American Crises*. Stanford: Stanford University Press.
- Becker, Gary. 1957. *The Economics of Discrimination*. Chicago: University of Chicago Press.
- Berg, E. J. 1966. “Backward-sloping labor supply functions in dual economies—the Africa case.” Pp. 114–136 in Immanuel Wallerstein (ed.), *Social Change: The Colonial Situation*. New York: Wiley.
- Blalock, H. M., Jr. 1967. *Toward a Theory of Minority-Group Relations*. New York: Wiley.
- Blauner, Robert. 1969. “Internal colonialism and ghetto revolt.” *Social Problems* 16 (Spring): 393–408.
- Cox, Oliver C. 1948. *Caste, Class and Race*. New York: Modern Reader.
- Daniels, Roger. 1966. *The Politics of Prejudice*. Gloucester, Massachusetts: Peter Smith.
- Daniels, Roger, and Harry H. L. Kitano. 1970. *American Racism*. Englewood Cliffs: Prentice-Hall.
- Davis, Allison W., B. B. Gardner, and M. R. Gardner. 1941. *Deep South*. Chicago: University of Chicago Press.
- Despres, Leo A. 1969. “Differential adaptations and micro-cultural evolution in Guyana.” *Southwestern Journal of Anthropology* 25 (Spring): 14–44.
- Doxey, G. V. 1961. *The Industrial Colour Bar in South Africa*. Cape Town: Oxford University Press.
- Ferguson-Davie, C. J. 1952. *The Early History of Indians in Natal*. Johannesburg: South African Institute of Race Relations.
- Gillion, K. L. 1962. *Fiji’s Indian Migrants*. Melbourne: Oxford University Press.
- Harris, Marvin. 1964. *Patterns of Race in the Americas*. New York: Walker.
- Heneman, H. G., and Dale Yoder. 1965. *Labor Economics*. Cincinnati: Southwestern.
- Higham, John. 1965. *Strangers in the Land*. New York: Athenium.
- Hutchison, Emilie J. 1968. *Women’s Wages*. New York: Ams Press.
- Ichihashi, Yamato. 1932. *Japanese in the United States*. Stanford: Stanford University Press.
- Iwata, Masakazu. 1962. “The Japanese immigrants in California agriculture.” *Agricultural History* 36 (January):25–37.
- King, Cameron H., Jr. 1908. “Asiatic exclusion.” *International Socialist Review* 8 (May):661–669.
- Lenin, V. I. 1964. “Imperialism and the split in socialism.” Pp. 105–120 in *Collected Works, Volume 23, August 1916–March 1917*. Moscow: Progress.
- Lieberson, Stanley. 1961. “A societal theory of race and ethnic relations.” *American Sociological Review* 26 (December):902–910.
- Lind, Andrew W. 1968. *An Island Community*. New York: Greenwood.
- Loewen, James W. 1971. *The Mississippi Chinese*. Cambridge: Harvard University Press.
- Mason, Philip. 1970. *Patterns of Dominance*. London: Oxford University Press.
- McWilliams, Carey. 1945. *Prejudice: Japanese-Americans*. Boston: Little, Brown.
- Millis, H. A. 1915. *The Japanese Problem in the United States*. New York: Macmillan.
- Moles, I. N. 1968. “The Indian coolie labour issue.” Pp. 40–48 in A. T. Yarwood (ed.), *Attitudes to Non-European Immigration*. Melbourne: Cassell Australia.
- Noel, Donald L. 1968. “A theory of the origin of ethnic stratification.” *Social Problems* 16 (Fall):157–172.
- Reich, Michael. 1971. “The economics of racism.” Pp. 107–113 in David M. Gordon (ed.), *Problems in Political Economy*. Lexington, Massachusetts: Heath.
- Schermerhorn, R. A. 1970. *Comparative Ethnic Relations*. New York: Random House.
- Shibutani, Tamotsu, and Kian M. Kwan. 1965. *Ethnic Stratification*. New York: Macmillan.
- Siu, Paul C. P. 1952. “The sojourner.” *American Journal of Sociology* 58 (July):34–44.
- Sutherland, Edwin H., and Donald R. Cressey. 1970. *Criminology*. Philadelphia: Lippincott.
- van den Berghe, Pierre L. 1966. “Paternalistic versus competitive race relations: an ideal-type approach.”

- Pp. 53–69 in Bernard E. Segal (ed.), *Racial and Ethnic Relations*. New York: Crowell.
- _____. 1967a. *Race and Racism*. New York: Wiley.
- _____. 1967b. *South Africa: A Study in Conflict*. Berkeley: University of California Press.
- Van der Horst, Sheila T. 1965. “The effects of industrialization on race relations in South Africa.” Pp. 97–140 in Guy Hunter (ed.), *Industrialization and Race Relations*. London: Oxford University Press.
- Willard, Myra. 1967. *History of the White Australia Policy to 1920*. London: Melbourne University Press.
- Williams, Eric. 1944. *Capitalism and Slavery*. Chapel Hill: University of North Carolina Press.
- Wittke, Carl. 1953. “Immigration policy prior to World War I.” Pp. 1–10 in Benjamin M. Ziegler (ed.), *Immigration: An American Dilemma*. Boston: Heath.
- Yarwood, A. T. 1964. *Asian Immigration to Australia*. London: Cambridge University Press.
- Yarwood, A. T. (ed.). 1968. *Attitudes to Non-European Immigration*. Melbourne: Cassell Australia.

67 ■ Alejandro Portes and Robert D. Manning

The Immigrant Enclave

Theory and Empirical Examples

I. Introduction

The purpose of this chapter is to review existing theories about the process of immigrant adaptation to a new society and to recapitulate the empirical findings that have led to an emerging perspective on the topic. This emerging view revolves around the concepts of different modes of structural incorporation and of the immigrant enclave as one of them. These concepts are set in explicit opposition to two previous viewpoints on the adaptation process, generally identified as assimilation theory and the segmented labor markets approach.

The study of immigrant groups in the United States has produced a copious historical and sociological literature, written mostly from the assimilation perspective. Although the experiences of particular groups varied, the common theme of these writings is the unrelenting efforts of immigrant mi-

norities to surmount obstacles impeding their entry into the “mainstream” of American society (Handlin, 1941, 1951; Wittke, 1952; Child, 1943; Vecoli, 1977).

From this perspective, the adaptation process of particular immigrant groups followed a sequential path from initial economic hardship and discrimination to eventual socioeconomic mobility arising from increasing knowledge of American culture and acceptance by the host society (Warner and Srole, 1945; Gordon, 1964; Sowell, 1981). The focus on a “core” culture, the emphasis on consensus-building, and the assumption of a basic patterned sequence of adaptation represent central elements of assimilation theory. From this perspective, the failure of individual immigrants or entire ethnic groups to move up through the social hierarchies is linked either to their reluctance to shed traditional values or to the resistance of the native majority to ac-

cept them because of racial, religious, or other shortcomings. Hence, successful adaptation depends, first of all, on the willingness of immigrants to relinquish a “backward” way of life and, second, on their acquisition of characteristics making them acceptable to the host society (Eisenstadt, 1970). Throughout, the emphasis is placed on the social psychological processes of motivation, learning, and interaction and on the cultural values and perceptions of the immigrants themselves and those who surround them.

The second general perspective takes issue with this psychosocial and culturalist orientation as well as with the assumption of a single basic assimilation path. This alternative view begins by noting that immigrants and their descendants do not necessarily “melt” into the mainstream and that many groups seem not to want to do so, preferring instead to preserve their distinct ethnic identities (Greeley, 1971; Glazer and Moynihan, 1970). A number of writers have focused on the resilience of these communities and described their functions as sources of mutual support and collective political power (Suttles, 1968; Alba and Chamlin, 1983; Parenti, 1967). Others have gone beyond descriptive accounts and attempted to establish the causes of the persistence of ethnicity. Without exception, these writers have identified the roots of the phenomenon in the economic sphere and, more specifically, in the labor-market roles that immigrants have been called on to play.

Within this general perspective, several specific theoretical approaches exist. The first focuses on the situation of the so-called unmeltable ethnics—blacks, Chicanos, and American Indians—and finds the source of their plight in a history of internal colonialism during which these groups have been confined to specific areas and made to work under uniquely unfavorable conditions. In a sense, the role of colonized minorities has been to bypass the free labor market, yielding in the process distinct benefits both to direct employers of their labor and, indirectly, to

other members of the dominant racial group (Blauner, 1972; Geschwender, 1978). The continuation of colonialist practices to our day explains, according to this view, the spatial isolation and occupational disadvantages of these minorities (Barrera, 1980).

A second approach attempts to explain the persistence of ethnic politics and ethnic mobilization on the basis of the organization of subordinate groups to combat a “cultural division of labor.” The latter confined members of specific minorities to a quasi-permanent situation of exploitation and social inferiority. Unlike the first view, this second approach does not envision the persistence of ethnicity as a consequence of continuing exploitation, but rather as a “reactive formation” on the part of the minority to reaffirm its identity and its interests (Hechter, 1977; Despres, 1975). For this reason, ethnic mobilizations are often most common among groups who have already abandoned the bottom of the social ladder and started to compete for positions of advantage with members of the majority (Nagel and Olzak, 1982).

A final variant focuses on the situation of contemporary immigrants to the United States. Drawing on the dual labor market literature, this approach views recent immigrants as the latest entrants into the lower tier of a segmented labor market where women and other minorities already predominate. Relative to the latter, immigrants possess the advantages of their lack of experience in the new country, their legal vulnerability, and their greater initial motivation. All of these traits translate into higher productivity and lower labor costs for the firms that employ them (Sassen-Koob, 1980). Jobs in the secondary labor market are poorly paid, require few skills, and offer limited mobility opportunities. Hence, confinement of immigrants to this sector insures that those who do not return home are relegated to a quasi-permanent status as disadvantaged and discriminated minorities (Piore, 1975, 1979).

What these various structural theories have in common is the view of resilient ethnic

communities formed as the result of a consistently disadvantageous economic position and the consequent absence of a smooth path of assimilation. These situations, ranging from slave labor to permanent confinement to the secondary labor market, are not altered easily. They have given rise, in time, either to hopeless communities of “unmeltable” ethnics or to militant minorities, conscious of a common identity and willing to support a collective strategy of self-defense rather than rely on individual assimilation.

These structural theories have provided an effective critique of the excessively benign image of the adaptation process presented by earlier writings. However, while undermining the former, the new structural perspective may have erred in the opposite direction. The basic hypothesis advanced in this chapter is that several identifiable modes of labor-market incorporation exist and that not all of them relegate newcomers to a permanent situation of exploitation and inferiority. Thus, while agreeing with the basic thrust of structural theories, we propose several modifications that are necessary for an adequate understanding of the different types of immigrant flows and their distinct processes of adaptation.

II. Modes of Incorporation

In the four decades since the end of World War II, immigration to the United States has experienced a vigorous surge reaching levels comparable only to those at the beginning of the century (National Research Council, 1985, chapter 2). Even if one restricts attention to this movement, disregarding multiple other migrations elsewhere in the world, it is not the case that the inflow has been of a homogeneous character. Low-wage labor immigration itself has taken different forms, including temporary contract flows, undocumented entries, and legal immigration. More importantly, it is not the case that all immigrants have been directed to the secondary labor market. For example, since the promulgation

of the Immigration Act of 1965, thousands of professionals, technicians, and craftsmen have come to the United States, availing themselves of the occupational preference categories of the law. This type of inflow, dubbed “brain drain” in the sending nations, encompasses today sizable contingents of immigrants from such countries as India, South Korea, the Philippines, and Taiwan, each an important contributor to U.S. annual immigration.

The characteristics of this type of migration have been described in detail elsewhere (Portes, 1976, 1981). Two such traits deserve mention, however. First, occupationally skilled immigrants—including doctors, nurses, engineers, technicians, and craftsmen—generally enter the “primary” labor market; they contribute to alleviate domestic shortages in specific occupations and gain access, after a period of time, to the mobility ladders available to native workers. Second, immigration of this type does not generally give rise to spatially concentrated communities; instead, immigrants are dispersed throughout many cities and regions, following different career paths.

Another sizable contingent of entrants whose occupational future is not easily characterized *a priori* are political refugees. Large groups of refugees, primarily from Communist-controlled countries, have come to the United States, first after the occupation of Eastern Europe by the Soviet Army, then after the advent of Fidel Castro to power in Cuba, and finally in the aftermath of the Vietnam War. Unlike purely “economic” immigrants, refugees have often received resettlement assistance from various governmental agencies (Zolberg, 1983; Keely, 1981). All the available evidence runs contrary to the notion of a uniform entry of political refugees into low-wage secondary occupations; on the contrary, there are indications of their employment in many different lines of work.

A third mode of incorporation has gained the attention of a number of scholars in recent years. It consists of small groups of immigrants who are inserted or insert themselves as

commercial intermediaries in a particular country or region. These “middleman minorities” are distinct in nationality, culture, and sometimes race from both the superordinate and subordinate groups to which they relate (Bonacich, 1973; Light, 1972). They can be used by dominant elites as a buffer to deflect mass frustration and also as an instrument to conduct commercial activities in impoverished areas. Middlemen accept these risks in exchange for the opportunity to share in the commercial and financial benefits gained through such instruments as taxation, higher retail prices, and usury. Jews in feudal and early modern Europe represent the classic instance of a middleman minority. Other examples include Indian merchants in East Africa, and Chinese entrepreneurs in Southeast Asia and throughout the Pacific Basin (Bonacich and Modell, 1980, chapter 1). Contemporary examples in the United States include Jewish, Korean, and other Oriental merchants in inner-city ghetto areas and Cubans in Puerto Rico (Kim, 1981; Cobas, 1984).

Primary labor immigration and middleman entrepreneurship represent two modes of incorporation that differ from the image of an homogeneous flow into low-wage employment. Political refugees, in turn, have followed a variety of paths, including both of the above as well as insertion into an ethnic enclave economy. The latter represents a fourth distinct mode. Although frequently confused with middleman minorities, the emergence and structure of an immigrant enclave possess distinct characteristics. The latter have significant theoretical and practical implications, for they set apart groups adopting this entry mode from those following alternative paths. We turn now to several examples of immigrant enclaves to clarify their internal dynamics and causes of their emergence.

III. Immigrant Enclaves

Immigration to the United States before World War I was, overwhelmingly, an un-

skilled labor movement. Impoverished peasants from southern Italy, Poland, and the eastern reaches of the Austro-Hungarian Empire settled in dilapidated and crowded areas, often immediately adjacent to their points of debarkation, and took any menial jobs available. From these harsh beginnings, immigrants commenced a slow and often painful process of acculturation and economic mobility. Theirs was the saga captured by innumerable subsequent volumes written from both the assimilation and the structural perspectives.

Two sizable immigrant groups did not follow this pattern, however. Their most apparent characteristic was the economic success of the first generation, even in the absence of extensive acculturation. On the contrary, both groups struggled fiercely to preserve their cultural identity and internal solidarity. Their approach to adaptation thus directly contradicted subsequent assimilation predictions concerning the causal priority of acculturation to economic mobility. Economic success and “clannishness” also earned for each minority the hostility of the surrounding population. These two immigrant groups did not have a language, religion, or even race in common and they never overlapped in significant numbers in any part of the United States. Yet, arriving at opposite ends of the continent, Jews and Japanese pursued patterns of economic adaptation that were quite similar both in content and in their eventual consequences.

A. Jews in Manhattan

The first major wave of Jewish immigration to the United States consisted of approximately 50,000 newcomers of German origin, arriving between 1840 and 1870. These immigrants went primarily into commerce and achieved, in the course of a few decades, remarkable success. By 1900, the average income of German-Jewish immigrants surpassed that of the American population (Rischin, 1962). Many individuals who started as street peddlers and small merchants

had become, by that time, heads of major industrial, retail, and financial enterprises.

The second wave of Jewish immigration exhibited quite different characteristics. Between 1870 and 1914, over two million Jews left the Pale of Settlement and other Russian-dominated regions, escaping Czarist persecution. Major pogroms occurred before and during this exodus (Dinnerstein, 1977). Thus, unlike most immigrants of the period, the migration of Russian and Eastern Europe Jews was politically motivated and their move was much more permanent. In contrast to German Jews, who were relatively well educated, the Yiddish-speaking newcomers came, for the most part, from modest origins and had only a rudimentary education. Although they viewed the new Russian wave with great apprehension, German Jews promptly realized that their future as an ethnic minority depended on the successful integration of the newcomers (Rischin, 1962). Charitable societies were established to provide food, shelter, and other necessities, and private schools were set up to teach immigrants English, civics, and the customs of the new country (Howe and Libo, 1979).

Aside from its size and rapidity of arrival, turn-of-the-century Jewish immigration had two other distinct characteristics. First was its strong propensity toward commerce and self-employment in general in preference to wage labor: as German Jews before them, many Russian immigrants moved directly into street peddling and other commercial activities of the most modest sort. Second was its concentration into a single, densely populated urban area—the lower East Side of Manhattan. Within this area, those who did not become storekeepers and peddlers from the start found employment in factories owned by German Jews, learning the necessary rudiments for future self-employment (Sowell, 1981, chapter 4).

The economic activities of this population created, in the course of two decades, a dense network of industrial, commercial, and finan-

cial enterprises. Close physical proximity facilitated exchanges of information and access to credit and raw materials. Characteristic of this emerging Jewish enclave is that production and marketing of goods was not restricted to the ethnic community, but went well beyond it into the general economy. Jews entered the printing, metal, and building trades; they became increasingly prominent in jewelry and cigar-making; above all, the garment industry became the primary domain of Jewish entrepreneurship, with hundreds of firms of all sizes engaged in the trade (Rischin, 1962; Howe and Libo, 1979).

The economic success of many of these ventures did not require and did not entail rapid acculturation. Immigrants learned English and those instrumental aspects of the new culture required for economic advancement. For the rest, they preferred to remain with their own and maintained, for the most part, close adherence to their original religion, language, and values (Wirth, 1956; Howe, 1976). Jewish enclave capitalism depended, for its emergence and advancement, precisely on those resources made available by a solidaristic ethnic community: protected access to labor and markets, informal sources of credit, and business information. It was through these resources that upstart immigrant enterprises could survive and eventually compete effectively with better-established firms in the general economy.

The emergence of a Jewish enclave in East Manhattan helped this group bypass the conventional assimilation path and achieve significant economic mobility in the course of the first generation, well ahead of complete acculturation. Subsequent generations also pursued this path, but the resources accumulated through early immigrant entrepreneurship were dedicated primarily to further the education of children and their entry into the professions. It was at this point that outside hostility became most patent, as one university after another established quotas to prevent the onrush of Jewish students. The last

of these quotas did not come to an end until after World War II (Dinnerstein, 1977).

Despite these and other obstacles, the movement of Jews into higher education continued. Building on the economic success of the first generation, subsequent ones achieved levels of education, occupation, and income that significantly exceed the national average (Featherman, 1971; Sowell, 1981, chapter 4). The original enclave is now only a memory, but it provided in its time the necessary platform for furthering the rapid social and economic mobility of the minority. Jews did enter the mainstream of American society, but they did not do so starting uniformly at the bottom, as most immigrant groups had done; instead, they translated resources made available by early ethnic entrepreneurship into rapid access to positions of social prestige and economic advantage.

B. Japanese on the West Coast

The specific features of Japanese immigration differ significantly from the movement of European Jews, but their subsequent adaptation and mobility patterns are similar. Beginning in 1890 and ending with the enactment of the Gentlemen's Agreement of 1908, approximately 150,000 Japanese men immigrated to the West Coast. They were followed primarily by their spouses until the Immigration Act of 1924 banned any further Asiatic immigration. Although nearly 300,000 Japanese immigrants are documented in this period (Daniels, 1977), less than half of this total remained in the United States (Petersen, 1971). This is due, in contrast to the case of the Jews, to the sojourner character of Japanese immigrants: the intention of many was to accumulate sufficient capital for purchasing farm land or settling debts in Japan. Hence this population movement included commercial and other members of the Japanese middle class who, not incidentally, were explicitly sponsored by their national government.

The residential patterns of Japanese immigrants were not as concentrated as those of

Jews in Manhattan, but they were geographically clustered. Almost two-thirds of the 111,010 Japanese reported in the U.S. Census of 1920 lived in California. Further, one-third of California's Japanese residents lived in Los Angeles County in 1940, while another one-third lived in six nearby counties (Daniels, 1977). However, it was not the residential segregation of Japanese immigrants but rather their occupational patterns that eventually mobilized the hostility of the local population.

Japanese immigrants were initially welcomed and recruited as a form of cheap agricultural labor. Their reputation as thrifty and diligent workers made them preferable to other labor sources. Nativist hostilities crystallized, however, when Japanese immigrants shifted from wage labor to independent ownership and small-scale farming. This action not only reduced the supply of laborers but it also increased competition for domestic growers in the fresh-produce market. In 1900, only about 40 Japanese farmers in the entire United States leased or owned a total of 5000 acres of farmland. By 1909, the number of Japanese farmers had risen to 6000 and their collective holdings exceeded 210,000 acres (Petersen, 1971). Faced with such "unfair" competition, California growers turned to the political means at their disposal. In 1913, the state legislature passed the first Alien Land Law, which restricted land ownership by foreigners. This legislation did not prove sufficient, however, and, in subsequent years, the ever-accommodating legislature passed a series of acts closing other legal loopholes to Japanese farming (Petersen, 1971).

These proscriptions, which barred most of the Japanese from the lands, accelerated their entry into urban enterprise. In 1909, Japanese entrepreneurs owned almost 3000 small shops in several Western cities. Forty percent of Japanese men in Los Angeles were self-employed. They operated businesses such as dry-cleaning establishments, fisheries, lunch

counters, and produce stands that marketed the production of Japanese farms (Light, 1972).

The ability of the first-generation *Issei* to escape the status of stoop labor in agriculture was based on the social cohesion of their community. Rotating credit associations offered scarce venture capital, while mutual-aid organizations provided assistance in operating farms and urban businesses. Light (1972) reports that capitalizations as high as \$100,000 were financed through ethnic credit networks. Economic success was again accompanied by limited instrumental acculturation and by careful preservation of national identity and values. It was the availability of investment capital, cooperative business associations, and marketing practices (forward and backward economic linkages) within the ethnic enclave that enabled Japanese entrepreneurs to expand beyond its boundaries and compete effectively in the general economy. This is illustrated by the production and marketing of fresh produce. In 1920, the value of Japanese crops was about 10% of the total for California, when the Japanese comprised less than 1% of the state's population; many retail outlets traded exclusively with a non-Japanese clientele (Light, 1972; Petersen, 1971).

During the early 1940s, the Japanese ethnic economy was seriously disrupted but not eliminated by the property confiscations and camp internments accompanying World War II. After the war, economic prosperity and other factors combined to reduce local hostility toward the Japanese. Older *Issei* and many of their children returned to small business, while other second-generation *Nisei*, like their Jewish predecessors, pursued higher education and entered the white-collar occupations *en masse*. This mobility path was completed by the third or *Sansei* generation, with 88% of their members attending college. Other third-generation Japanese have continued, however, the entrepreneurial tradition of their parents (Bonacich and

Modell, 1980). Like Jews before them, Japanese Americans have made use of the resources made available by early immigrant entrepreneurship to enter the mainstream of society in positions of relative advantage. The mean educational and occupational attainment of the group's 600,000 members surpasses at present all other ethnic and native groups, while its average family income is exceeded among American ethnic groups only by the Jews (Sowell, 1981).¹ . . .

IV. Conclusion: A Typology of the Process of Incorporation

We can now attempt a summary description of the characteristics of immigrant enclaves and how they differ from other paths. The emergence of an ethnic enclave economy has three prerequisites: first, the presence of a substantial number of immigrants, with business experience acquired in the sending country; second, the availability of sources of capital; and third, the availability of sources of labor. The latter two conditions are not too difficult to meet. The requisite labor can usually be drawn from family members and, more commonly, from recent arrivals. Surprisingly perhaps, capital is not a major impediment either since the sums initially required are usually small. When immigrants did not bring them from abroad, they could be accumulated through individual savings or pooled resources in the community. It is the first condition that appears critical. The presence of a number of immigrants skilled in what Franklin Frazier (1949) called the art of "buying and selling" is common [not only to the Jewish and Japanese cases reviewed above but also to contemporary enclave economies among Koreans and Cubans]. Such an entrepreneurial-commercial class among early immigrant cohorts can usually overcome other obstacles; conversely, its absence within an immigrant community will confine the community to wage employment even if sufficient resources of capital and labor are available.

Enclave businesses typically start small and cater exclusively to an ethnic clientele. Their expansion and entry into the broader market requires, as seen above, an effective mobilization of community resources. The social mechanism at work here seems to be a strong sense of reciprocity supported by collective solidarity that transcends the purely contractual character of business transactions. For example, receipt of a loan from a rotating credit association entails the duty of continuing to make contributions so that others can have access to the same source of capital. Although, in principle, it would make sense for the individual to withdraw once his loan is received, such action would cut him off from the very sources of community support on which his future business success depends (Light, 1972).

Similarly, relations between enclave employers and employees generally transcend a contractual wage bond. It is understood by both parties that the wage paid is inferior to the value of labor contributed. This is willingly accepted by many immigrant workers because the wage is only *one* form of compensation. Use of their labor represents often the key advantage making poorly capitalized enclave firms competitive. In reciprocity, employers are expected to respond to emergency needs of their workers and to promote their advancement through such means as on-the-job training, advancement to supervisory positions, and aid when they move into self-employment. These opportunities represent the other part of the “wage” received by enclave workers. The informal mobility ladders thus created are, of course, absent in the secondary labor market where there is no primary bond between owners and workers or no common ethnic community to enforce the norm of reciprocity.

Paternalistic labor relations and strong community solidarity are also characteristic of middleman minorities. Although both modes of incorporation are similar and are thus frequently confused, there are three major struc-

tural differences between them. First, immigrant enclaves are not exclusively commercial. Unlike middleman minorities, whose economic role is to mediate commercial and financial transactions between elites and masses, enclave firms include in addition a sizable productive sector. The latter may comprise agriculture, light manufacturing, and construction enterprises; their production, marketed often by coethnic intermediaries, is directed toward the general economy and not exclusively to the immigrant community.

Second, relationships between enclave businesses and established native ones are problematic. Middleman groups tend to occupy positions complementary and subordinate to the local owning class; they fill economic niches either disdained or feared by the latter. Unlike them, enclave enterprises often enter in direct competition with existing domestic firms. There is no evidence, for example, that domestic elites deliberately established or supported the emergence of the Jewish, Japanese, Korean, or Cuban business communities as means to further their own economic interests. There is every indication, on the other hand, that this mode of incorporation was largely self-created by the immigrants, often in opposition to powerful domestic interests. Although it is true that enclave entrepreneurs have been frequently employed as subcontractors by outside firms in such activities as garment and construction (Bonacich, 1978), it is incorrect to characterize this role as the exclusive or dominant one among these enterprises.

Third, the enclave is concentrated and spatially identifiable. By the very nature of their activities, middleman minorities must often be dispersed among the mass of the population. Although the immigrants may live in certain limited areas, their businesses require proximity to their mass clientele and a measure of physical dispersion within it. It is true that middleman activities such as money-lending have been associated in several historical instances with certain streets and neighborhoods,

but this is not a necessary or typical pattern. Street peddling and other forms of petty commerce require merchants to go into the areas where demand exists and avoid excessive concentration of the goods and services they offer. This is the typical pattern found today among middleman minorities in American cities (Cobas, 1984; Kim, 1981).

Enclave businesses, on the other hand, are spatially concentrated, especially in their early stages. This is so for three reasons: first, the need for proximity to the ethnic market which they initially serve; second, proximity to each other which facilitates exchange of information, access to credit, and other supportive activities; third, proximity to ethnic labor supplies on which they crucially depend. Unlike the Jewish, Korean, or Cuban cases, the Japanese enclave economy does partially depart from the pattern of high physical concentration. This can be attributed to the political persecution to which this group was subjected. Originally, Japanese concentration was a rural phenomenon based on small farms linked together by informal bonds and cooperative associations. Forced removal of this minority from the land compelled their entry into urban businesses and their partial dispersal into multiple activities.

Physical concentration of enclaves underlies their final characteristic. Once an enclave economy has fully developed, it is possible for a newcomer to live his life entirely within the confines of the community. Work, education, and access to health care, recreation, and a variety of other services can be found without leaving the bounds of the ethnic economy. This institutional completeness is what enables new immigrants to move ahead economically, despite very limited knowledge of the host culture and language. Supporting empirical evidence comes from studies showing low levels of English knowledge among enclave minorities and the absence of a net effect of knowledge of English on their average income levels (Light, 1980; Portes and Bach, 1985).

Table 1 summarizes this discussion by presenting the different modes of incorporation and their principal characteristics. Two caveats are necessary. First, this typology is not exhaustive, since other forms of adaptation have existed and will undoubtedly emerge in the future. Second, political refugees are not included, since this entry label does not necessarily entail a unique adaptation path. Instead, refugees can select or be channelled in many different directions, including self-employment, access to primary labor markets, or confinement to secondary sector occupations.

Having discussed the characteristics of enclaves and middleman minorities, a final word must be said about the third alternative to employment in the lower tier of a dual labor market. As a mode of incorporation, primary sector immigration also has distinct advantages, although they are of a different order from those pursued by “entrepreneurial” minorities. Dispersal throughout the receiving country and career mobility based on standard promotion criteria make it imperative for immigrants in this mode to become fluent in the new language and culture (Stevens, Goodman, and Mick, 1978). Without a supporting ethnic community, the second generation also becomes thoroughly steeped in the ways of the host society. Primary sector immigration thus tends to lead to very rapid social and cultural integration. It represents the path that approximates most closely the predictions of assimilation theory with regard to (1) the necessity of acculturation for social and economic progress and (2) the subsequent rewards received by immigrants and their descendants for shedding their ethnic identities.

Clearly, however, this mode of incorporation is open only to a minority of immigrant groups. In addition, acculturation of professionals and other primary sector immigrants is qualitatively different from that undergone by others. Regardless of their differences, immigrants in other modes tend to learn the new

Table 1. Typology of Modes of Incorporation

| Variable | Primary sector immigration | Secondary sector immigration | Immigrant enclaves | Middleman minorities |
|---|--|-------------------------------|--|---|
| Size of immigrant population | Small | Large | Large | Small |
| Spatial concentration, national | Dispersed | Dispersed | Concentrated | Concentrated |
| Spatial concentration, local | Dispersed | Concentrated | Concentrated | Dispersed |
| Original class composition | Homogeneous: skilled workers and professionals | Homogeneous: manual laborers | Heterogeneous: entrepreneurs, professionals, and workers | Homogeneous: merchants and some professionals |
| Present occupational status distribution | High mean status/low variance | Low mean status/low variance | Mean status/high variance | Mean status/low variance |
| Mobility opportunities | High: formal promotion ladders | Low | High: informal ethnic ladders | Average: informal ethnic ladders |
| Institutional diversification of ethnic community | None | Low: weak social institutions | High: institutional completeness | Medium: strong social and economic institutions |
| Participation in ethnic organizations | Little or none | Low | High | High |
| Resilience of ethnic culture | Low | Average | High | High |
| Knowledge of host country language | High | Low | Low | High |
| Knowledge of host country institutions | High | Low | Average | High |
| Modal reaction of host community | Acceptance | Discrimination | Hostility | Mixed: elite acceptance/mass hostility |

language and culture with a heavy “local” content. Although acculturation may be slow, especially in the case of enclave groups, it carries with it elements unique to the surrounding community—its language inflections, particular traditions, and loyalties (Greeley, 1971; Suttles, 1968). On the contrary, acculturation of primary sector immigrants is of a more cosmopolitan sort. Because career requirements often entail physical mobility, the new language and culture are learned more rapidly and more generally, without strong attachments to a particular community. Thus, while minorities entering menial labor, enclave, or middleman enterprise in the United States have eventually become identified with a certain city or region, the same is not true for immigrant professionals, who tend to “disappear,” in a cultural sense, soon after their arrival (Stevens *et al.*, 1978; Cardona and Cruz, 1980).

Awareness of patterned differences among immigrant groups in their forms of entry and labor market incorporation represents a significant advance, in our view, from earlier undifferentiated descriptions of the adaptation process. This typology is, however, a provisional effort. Just as detailed research on the condition of particular minorities modified or replaced earlier broad generalizations, the propositions advanced here will require revision. New groups arriving in the United States at present and a revived interest in immigration should provide the required incentive for empirical studies and theoretical advances in the future.

NOTES

1. The original article from which this chapter was drawn includes a further discussion of contemporary Korean and Cuban enclaves.—ED.

REFERENCES

Alba, Richard D., and Chamlin, Mitchell B. (1983). Ethnic identification among whites. *American Sociological Review* 48:240–247.

- Barrera, Mario (1980). *Race and class in the Southwest: A theory of racial inequality*. Notre Dame, Indiana: Notre Dame University Press.
- Blauner, Robert (1972). *Racial oppression in America*. New York: Harper and Row.
- Bonacich, Edna (1973). A theory of middleman minorities. *American Sociological Review* 38 (October):583–594.
- Bonacich, Edna (1978). U.S. capitalism and Korean immigrant small business. Riverside, California: Department of Sociology, University of California—Riverside, mimeographed.
- Bonacich, Edna, and Modell, John (1980). *The economic basis of ethnic solidarity: Small business in the Japanese-American community*. Berkeley, California: University of California Press.
- Cardona, Ramiro C., and Cruz, Carmen I. (1980). *El exodo de Colombianos*. Bogota: Ediciones Tercer Mundo.
- Child, Irving L. (1943). *Italian or American? The second generation in conflict*. New Haven, Connecticut: Yale University Press.
- Cobas, Jose (1984). Participation in the ethnic economy, ethnic solidarity and ambivalence toward the host society: The case of Cuban emigres in Puerto Rico. Presented at the American Sociological Association Meeting, San Antonio, Texas: August.
- Daniels, Roger (1977). The Japanese-American experience: 1890–1940. In *Uncertain Americans* (L. Dinnerstein and F. C. Jaher, eds.), pp. 250–267. New York: Oxford University Press.
- Despres, Leo (1975). Toward a theory of ethnic phenomena. In *Ethnicity and resource competition* (Leo Despres, ed.), pp. 209–212. The Hague: Mouton.
- Dinnerstein, Leonard (1977). The East European Jewish migration. In *Uncertain Americans* (L. Dinnerstein and F. C. Jaher, eds.), pp. 216–231. New York: Oxford University Press.
- Eisenstadt, S. N. (1970). The process of absorbing new immigrants in Israel. In *Integration and development in Israel* (S. N. Eisenstadt, RivKah Bar Yosef, and Chaim Adler, eds.), pp. 341–367. Jerusalem: Israel University Press.
- Featherman, David L. (1971). The socioeconomic achievement of white religio-ethnic sub-groups: Social and psychological explanations. *American Sociological Review* 36 (April): 207–222.
- Frazier, E. Franklin (1949). *The Negro in the United States*. New York: Macmillan.
- Geschwender, James A. (1978). *Racial stratification in America*. Dubuque, Iowa: William C. Brown.
- Glazer, Nathan, and Moynihan, Daniel P. (1970). *Beyond the melting pot: The Negroes, Puerto Ricans,*

- Jews, Italians and Irish of New York City*. Cambridge, Massachusetts: M.I.T. Press.
- Gordon, Milton M. (1964). *Assimilation in American life: The role of race, religion, and national origins*. New York: Oxford University Press.
- Greeley, Andrew (1971). *Why can't they be like us? America's white ethnic groups*. New York: Dutton.
- Handlin, Oscar (1941). *Boston's immigrants: A study of acculturation*. Cambridge: Harvard University Press.
- Handlin, Oscar (1951). *The uprooted: The epic story of the great migrations that made the American people*. Boston: Little, Brown.
- Hechter, Michael (1977). *Internal colonialism, the Celtic fringe in British national development, 1536–1966*. Berkeley, California: University of California Press.
- Howe, Irving (1976). *World of our fathers*. New York: Harcourt, Brace, Jovanovich.
- Howe, Irving, and Libo, Kenneth (1979). *How we lived, a documentary history of immigrant Jews in America*. New York: Richard March.
- Keely, Charles B. (1981). *Global refugee policy: The case for a development-oriented strategy*. New York: The Population Council.
- Kim, Illsoo (1981). *New urban immigrants, the Korean community in New York*. Princeton, New Jersey: Princeton University Press.
- Light, H. Ivan (1972). *Ethnic enterprise in America: Business and welfare among Chinese, Japanese, and Blacks*. Berkeley, California: University of California Press.
- Light, H. Ivan (1980). Asian enterprise in America: Chinese, Japanese, and Koreans in small business. In *Self-help in urban America* (Scott Cummings, ed.), pp. 33–57. New York: Kennikat Press.
- Nagel, Joane, and Olzak, Susan (1982). Ethnic mobilization in new and old states: An extension of the competition model. *Social Problems* 30:127–143.
- National Research Council (1985). Immigration statistics: A story of neglect. Report of the Panel on Immigration Statistics. Washington, D.C.: National Academy of Sciences.
- Parenti, Michael (1967). Ethnic politics and the persistence of ethnic identification. *American Political Science Review* 61:717–726.
- Petersen, William (1971). *Japanese Americans, oppression and success*. New York: Random House.
- Piore, Michael J. (1975). Notes for a theory of labor market stratification. In *Labor market segmentation* (Richard C. Edwards, Michael Reich, and David M. Gordon, eds.), pp. 125–171. Lexington, Massachusetts: Heath.
- Piore, Michael J. (1979). *Birds of passage, migrant labor and industrial societies*. New York: Cambridge University Press.
- Portes, Alejandro (1976). Determinants of the brain drain. *International Migration Review* 10 (Winter):489–508.
- Portes, Alejandro (1981). Modes of structural incorporation and theories of labor immigration. In *Global Trends in Migration, Theory and Research on International Population Movements* (Mary M. Kritz, Charles B. Keely, and Silvano M. Tomasi, eds.), pp. 279–297. New York: Center for Migration Studies.
- Portes, Alejandro, and Bach, Robert L. (1985). *Latin journey, Cuban and Mexican immigrants in the United States*. Berkeley, California: University of California Press.
- Rischin, Moses (1962). *The promised city, New York Jews 1870–1914*. Cambridge, Massachusetts: Harvard University Press.
- Sassen-Koob, Saskia (1980). Immigrant and minority workers in the organization of the labor process. *Journal of Ethnic Studies* (1/Spring):1–34.
- Sowell, Thomas (1981). *Ethnic America: A history*. New York: Basic Books.
- Stevens, Rosemary, Goodman, Louis W., and Mick, Stephen (1978). *The alien doctors, foreign medical graduates in American hospitals*. New York: Wiley.
- Suttles, Gerald D. (1968). *The social order of the slum: ethnicity and territory in the inner city*. Chicago: University of Chicago Press.
- Vecoli, Rudolph (1977). The Italian Americans. In *Uncertain Americans* (L. Dinnerstein and F. C. Jaher, eds.), pp. 201–215. New York: Oxford University Press.
- Warner, W. Lloyd, and Srole, Leo (1945). *The social systems of American ethnic groups*. New Haven: Yale University Press.
- Wirth, Louis (1956). *The ghetto*. Chicago: University of Chicago Press.
- Witke, Carl (1952). *Refugees of revolution: The German Forty-eighters in America*. Philadelphia: University of Pennsylvania Press.
- Zolberg, Aristide (1983). Contemporary transnational migrations in historical perspective: Patterns and dilemmas. In *U.S. immigration and refugee policy* (Mary M. Kritz, ed.), pp. 15–51. Lexington, Massachusetts: Heath.

■ NEW MODES ■

68 ■ Alejandro Portes and Min Zhou

The New Second Generation

Segmented Assimilation and Its Variants

*My name is Herb
and I'm not poor;
I'm the Herbie that you're looking for,
like Pepsi,
a new generation
of Haitian determination—
I'm the Herbie that you're looking for.*

A beat tapped with bare hands, a few dance steps, and the Haitian kid was rapping. His song, titled “Straight Out of Haiti,” was being performed at Edison High, a school that sits astride Little Haiti and Liberty City, the largest black area of Miami. The lyrics captured well the distinct outlook of his immigrant community. The panorama of Little Haiti contrasts sharply with the bleak inner city. In Miami’s Little Haiti, the storefronts leap out at the passersby. Bright blues, reds, and oranges vibrate to Haitian merengue blaring from sidewalk speakers.¹ Yet, behind the gay Caribbean exteriors, a struggle goes on that will define the future of this community. As we will see later on, it involves the second generation—children like Herbie—subject to conflicting pressure from parents and peers and to pervasive outside discrimination.

Growing up in an immigrant family has always been difficult, as individuals are torn by conflicting social and cultural demands while they face the challenge of entry into an unfamiliar and frequently hostile world. And yet the difficulties are not always the same. The process of growing up American oscillates between smooth acceptance and traumatic confrontation depending on the characteristics that immigrants and their children bring along and the social context that receives them. In this article, we explore some of these factors and their bearing on the process of social adaptation of the immigrant second generation. We propose a conceptual framework for understanding this process and illustrate it with selected ethnographic material and survey data from a recent survey of children of immigrants.

Research on the new immigration—that which arose after the passage of the 1965 Im-

migration Act—has been focused almost exclusively on the first generation, that is, on adult men and women coming to the United States in search of work or to escape political persecution. Little noticed until recently is the fact that the foreign-born inflow has been rapidly evolving from single adult individuals to entire family groups, including infant children and those born to immigrants in the United States. By 1980, 10 percent of dependent children in households counted by the census were second-generation immigrants.² In the late 1980s, another study put the number of students in kindergarten through twelfth grade in American schools who spoke a language other than English at home at 3 to 5 million.³

The great deal of research and theorizing on post-1965 immigration offers only tentative guidance on the prospects and paths of adaptation of the second generation because the outlook of this group can be very different from that of their immigrant parents. For example, it is generally accepted among immigration theorists that entry-level menial jobs are performed without hesitation by newly arrived immigrants but are commonly shunned by their U.S.-reared offspring. This disjuncture gives rise to a race between the social and economic progress of first-generation immigrants and the material conditions and career prospects that their American children grow to expect.⁴

Nor does the existing literature on second-generation adaptation, based as it is on the experience of descendants of pre-World War I immigrants, offer much guidance for the understanding of contemporary events. The last sociological study of children of immigrants was Irving Child's *Italian or American? The Second Generation in Conflict*, published fifty years ago.⁵ Conditions at the time were quite different from those confronting settled immigrant groups today. Two such differences deserve special mention. First, descendants of European immigrants who confronted the dilemmas of conflicting cul-

tures were uniformly white. Even if of a somewhat darker hue than the natives, their skin color reduced a major barrier to entry into the American mainstream. For this reason, the process of assimilation depended largely on individual decisions to leave the immigrant culture behind and embrace American ways. Such an advantage obviously does not exist for the black, Asian, and mestizo children of today's immigrants.

Second, the structure of economic opportunities has also changed. Fifty years ago, the United States was the premier industrial power in the world, and its diversified industrial labor requirements offered to the second generation the opportunity to move up gradually through better-paid occupations while remaining part of the working class. Such opportunities have increasingly disappeared in recent years following a rapid process of national deindustrialization and global industrial restructuring. This process has left entrants to the American labor force confronting a widening gap between the minimally paid menial jobs that immigrants commonly accept and the high-tech and professional occupations requiring college degrees that native elites occupy.⁶ The gradual disappearance of intermediate opportunities also bears directly on the race between first-generation economic progress and second-generation expectations, noted previously. . . .

Assimilation as a Problem

The Haitian immigrant community of Miami is composed of some 75,000 legal and clandestine immigrants, many of whom sold everything they owned in order to buy passage to America. First-generation Haitians are strongly oriented toward preserving a strong national identity, which they associate both with community solidarity and with social networks promoting individual success.⁷ In trying to instill national pride and an achievement orientation in their children, they clash, however, with the youngsters' everyday experiences in

school. Little Haiti is adjacent to Liberty City, the main black inner-city area of Miami, and Haitian adolescents attend predominantly inner-city schools. Native-born youths stereotype Haitians as too docile and too subservient to whites and they make fun of French and Creole and of the Haitians' accent. As a result, second-generation Haitian children find themselves torn between conflicting ideas and values: to remain Haitian they would have to face social ostracism and continuing attacks in school; to become American—black American in this case—they would have to forgo their parents' dreams of making it in America on the basis of ethnic solidarity and preservation of traditional values.⁸

An adversarial stance toward the white mainstream is common among inner-city minority youths who, while attacking the newcomers' ways, instill in them a consciousness of American-style discrimination. A common message is the devaluation of education as a vehicle for advancement of all black youths, a message that directly contradicts the immigrant parents' expectations. Academically outstanding Haitian American students, "Herbie" among them, have consciously attempted to retain their ethnic identity by cloaking it in black American cultural forms, such as rap music. Many others, however, have followed the path of least effort and become thoroughly assimilated. Assimilation in this instance is not into mainstream culture but into the values and norms of the inner city. In the process, the resources of solidarity and mutual support within the immigrant community are dissipated.

An emerging paradox in the study of today's second generation is the peculiar forms that assimilation has adopted for its members. As the Haitian example illustrates, adopting the outlooks and cultural ways of the native-born does not represent, as in the past, the first step toward social and economic mobility but may lead to the exact opposite. At the other end, immigrant youths who remain firmly ensconced in their respec-

tive ethnic communities may, by virtue of this fact, have a better chance for educational and economic mobility through use of the material and social capital that their communities make available.⁹

This situation stands the cultural blueprint for advancement of immigrant groups in American society on its head. As presented in innumerable academic and journalistic writings, the expectation is that the foreign-born and their offspring will first acculturate and then seek entry and acceptance among the native-born as a prerequisite for their social and economic advancement. Otherwise, they remain confined to the ranks of the ethnic lower and lower-middle classes.¹⁰ This portrayal of the requirements for mobility, so deeply embedded in the national consciousness, stands contradicted today by a growing number of empirical experiences.

A closer look at these experiences indicates, however, that the expected consequences of assimilation have not entirely reversed signs, but that the process has become segmented. In other words, the question is into what sector of American society a particular immigrant group assimilates. Instead of a relatively uniform mainstream whose mores and prejudices dictate a common path of integration, we observe today several distinct forms of adaptation. One of them replicates the time-honored portrayal of growing acculturation and parallel integration into the white middle-class; a second leads straight in the opposite direction to permanent poverty and assimilation into the underclass; still a third associates rapid economic advancement with deliberate preservation of the immigrant community's values and tight solidarity. This pattern of segmented assimilation immediately raises the question of what makes some immigrant groups become susceptible to the downward route and what resources allow others to avoid this course. In the ultimate analysis, the same general process helps explain both outcomes. We advance next our hypotheses as to how this process takes place and how the contrast-

ing outcomes of assimilation can be explained. This explanation is then illustrated with recent empirical material in the final section.

Vulnerability and Resources

Along with individual and family variables, the context that immigrants find upon arrival in their new country plays a decisive role in the course that their offspring's lives will follow. This context includes such broad variables as political relations between sending and receiving countries and the state of the economy in the latter and such specific ones as the size and structure of preexisting coethnic communities. The concept of modes of incorporation provides a useful theoretical tool to understand this diversity. As developed in prior publications, modes of incorporation consist of the complex formed by the policies of the host government; the values and prejudices of the receiving society; and the characteristics of the coethnic community. These factors can be arranged in a tree of contextual situations, illustrated by Figure 1. This figure provides a first approximation to our problem.¹¹

To explain second-generation outcomes and their segmented character, however, we need to go into greater detail into the meaning of these various modes of incorporation from the standpoint of immigrant youths. There are three features of the social contexts encountered by today's newcomers that create vulnerability to downward assimilation. The first is color, the second is location, and the third is the absence of mobility ladders. As noted previously, the majority of contemporary immigrants are nonwhite. Although this feature may appear at first glance as an individual characteristic, in reality it is a trait belonging to the host society. Prejudice is not intrinsic to a particular skin color or racial type, and, indeed, many immigrants never experienced it in their native lands. It is by virtue of moving into a new social environment, marked by different values and preju-

ices, that physical features become redefined as a handicap.

The concentration of immigrant households in cities and particularly in central cities, as documented previously, gives rise to a second source of vulnerability because it puts new arrivals in close contact with concentrations of native-born minorities. This leads to the identification of the condition of both groups—immigrants and the native poor—as the same in the eyes of the majority. More important, it exposes second-generation children to the adversarial subculture developed by marginalized native youths to cope with their own difficult situation.¹² This process of socialization may take place even when first-generation parents are moving ahead economically and, hence, their children have no objective reasons for embracing a counter-cultural message. If successful, the process can effectively block parental plans for intergenerational mobility.

The third contextual source of vulnerability has to do with changes in the host economy that have led to the evaporation of occupational ladders for intergenerational mobility. As noted previously, new immigrants may form the backbone of what remains of labor-intensive manufacturing in the cities as well as in their growing personal services sector, but these are niches that seldom offer channels for upward mobility. The new hourglass economy, created by economic restructuring, means that children of immigrants must cross a narrow bottleneck to occupations requiring advanced training if their careers are to keep pace with their U.S.-acquired aspirations. This race against a narrowing middle demands that immigrant parents accumulate sufficient resources to allow their children to effect the passage and to simultaneously prove to them the viability of aspirations for upward mobility. Otherwise, assimilation may not be into mainstream values and expectations but into the adversarial stance of impoverished groups confined to the bottom of the new economic hourglass.

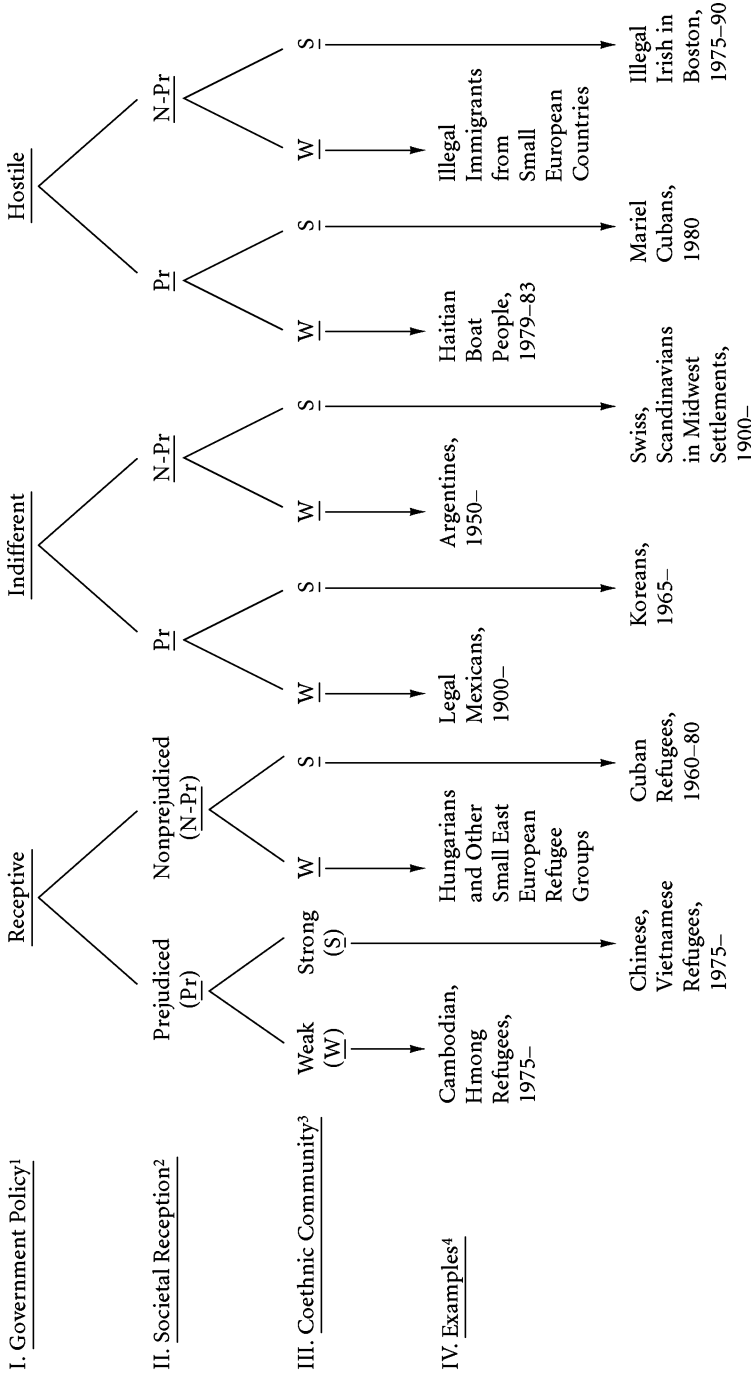


Figure 1. Models of Incorporation: A Typology

Source: Adapted from Alejandro Portes and Rubén G. Rumbaut, *Immigrant America: A Portrait* (Berkeley: University of California Press, 1990), p. 91. Copyright © 1990 by the Regents of the University of California.

1. Receptive policy is defined as legal entry with resettlement assistance, indifferent as legal entry without resettlement assistance, hostile as active opposition to a group's entry or permanence in the country; 2. Prejudiced reception is defined as that accorded to nonphenotypically white groups; nonprejudiced is that accorded to European and European-origin whites; 3. Weak coethnic communities are either small in numbers or composed primarily of manual workers; strong communities feature sizable numerical concentrations and a diversified occupational structure including entrepreneurs and professionals; 4. Examples include immigrant groups arriving from the start of the century to the present. Dates of migration are approximate. Groups reflect broadly but not perfectly the characteristics of each ideal type.

The picture is painted in such stark terms here for the sake of clarity, although in reality things have not yet become so polarized. Middle-level occupations requiring relatively modest educational achievements have not completely vanished. By 1980, skilled blue-collar jobs—classified by the U.S. census as “precision production, craft, and repair occupations”—had declined by 1.1 percent relative to a decade earlier but still represented 13 percent of the experienced civilian labor force, or 13.6 million workers. Mostly clerical administrative support occupations added another 16.9 percent, or 17.5 million jobs. In 1980, occupations requiring a college degree had increased by 6 percent in comparison with 1970, but they still employed less than a fifth—18.2 percent—of the American labor force.¹³ Even in the largest cities, occupations requiring only a high school diploma were common by the late 1980s. In New York City, for example, persons with 12 years or less of schooling held just over one half of the jobs in 1987. Clerical, service, and skilled blue-collar jobs not requiring a college degree represented 46 percent.¹⁴ Despite these figures, there is little doubt that the trend toward occupational segmentation has increasingly reduced opportunities for incremental upward mobility through well-paid blue-collar positions. The trend forces immigrants today to bridge in only one generation the gap between entry-level jobs and professional positions that earlier groups took two or three generations to travel.

Different modes of incorporation also make available, however, three types of resources to confront the challenges of contemporary assimilation. First, certain groups, notably political refugees, are eligible for a variety of government programs including educational loans for their children. The Cuban Loan Program, implemented by the Kennedy administration in connection with its plan to resettle Cuban refugees away from South Florida, gave many impoverished first- and second-generation Cuban youths a chance to attend college. The high proportion of professionals

and executives among Cuban American workers today, a figure on a par with that for native white workers, can be traced, at least in part, to the success of that program.¹⁵ Passage of the 1980 Refugee Act gave to subsequent groups of refugees, in particular Southeast Asians and Eastern Europeans, access to a similarly generous benefits package.¹⁶

Second, certain foreign groups have been exempted from the traditional prejudice endured by most immigrants, thereby facilitating a smoother process of adaptation. Some political refugees, such as the early waves of exiles from Castro’s Cuba, Hungarians and Czechs escaping the invasions of their respective countries, and Soviet Jews escaping religious persecution, provide examples. In other cases, it is the cultural and phenotypical affinity of newcomers to ample segments of the host population that ensures a welcome reception. The Irish coming to Boston during the 1980s are a case in point. Although many were illegal aliens, they came into an environment where generations of Irish Americans had established a secure foothold. Public sympathy effectively neutralized governmental hostility in this case, culminating in a change of the immigration law directly benefiting the newcomers.¹⁷

Third, and most important, are the resources made available through networks in the coethnic community. Immigrants who join well-established and diversified ethnic groups have access from the start to a range of moral and material resources well beyond those available through official assistance programs. Educational help for second-generation youths may include not only access to college grants and loans but also the existence of a private school system geared to the immigrant community’s values. Attendance at these private ethnic schools insulates children from contact with native minority youths, while reinforcing the authority of parental views and plans.

In addition, the economic diversification of several immigrant communities creates niches

of opportunity that members of the second generation can occupy, often without a need for an advanced education. Small-business apprenticeships, access to skilled building trades, and well-paid jobs in local government bureaucracies are some of the ethnic niches documented in the recent literature.¹⁸ In 1987, average sales per firm of the smaller Chinese, East Indian, Korean, and Cuban enterprises exceeded \$100,000 per year and they jointly employed over 200,000 workers. These figures omit medium-sized and large ethnic firms, whose sales and work forces are much larger.¹⁹ Fieldwork in these communities indicates that up to half of recently arrived immigrants are employed by coethnic firms and that self-employment offers a prime avenue for mobility to second-generation youths.²⁰ Such community-mediated opportunities provide a solution to the race between material resources and second-generation aspirations not available through competition in the open labor market. Through creation of a capitalism of their own, some immigrant groups have thus been able to circumvent outside discrimination and the threat of vanishing mobility ladders.

In contrast to these favorable conditions are those foreign minorities who either lack a community already in place or whose coethnics are too poor to render assistance. The condition of Haitians in South Florida, cited earlier, provides an illustration of one of the most handicapped modes of incorporation encountered by contemporary immigrants, combining official hostility and widespread social prejudice with the absence of a strong receiving community.²¹ From the standpoint of second-generation outcomes, the existence of a large but downtrodden coethnic community may be even less desirable than no community at all. This is because newly arrived youths enter into ready contact with the reactive subculture developed by earlier generations. Its influence is all the more powerful because it comes from individuals of the same national origin, “people like us” who can

more effectively define the proper stance and attitudes of the newcomers. To the extent that they do so, the first-generation model of upward mobility through school achievement and attainment of professional occupations will be blocked.

Three Examples

Mexicans and Mexican Americans

Field High School (the name is fictitious) is located in a small coastal community of central California whose economy has long been tied to agricultural production and immigrant farm labor. About 57 percent of the student population is of Mexican descent. An intensive ethnographic study of the class of 1985 at Field High began with school records that showed that the majority of U.S.-born Spanish-surname students who had entered the school in 1981 had dropped out by their senior year. However, only 35 percent of the Spanish-surname students who had been originally classified by the school as limited English proficient (LEP) had dropped out. The figure was even lower than the corresponding one for native white students, 40 percent. LEP status is commonly assigned to recently arrived Mexican immigrants.²²

Intensive ethnographic fieldwork at the school identified several distinct categories in which the Mexican-origin population could be classified. Recent Mexican immigrants were at one extreme. They dressed differently and unstylishly. They claimed an identity as Mexican and considered Mexico their permanent home. The most academically successful of this group were those most proficient in Spanish, reflecting their prior levels of education in Mexico. Almost all were described by teachers and staff as courteous, serious about their schoolwork, respectful, and eager to please as well as naive and unsophisticated. They were commonly classified as LEP.

The next category comprised Mexican-oriented students. They spoke Spanish at home and were generally classified as fluent English

proficient (FEP). They had strong bicultural ties with both Mexico and the United States, reflecting the fact that most were born in Mexico but had lived in the United States for more than five years. They were proud of their Mexican heritage but saw themselves as different from the first group, the *recién llegados* (recently arrived), as well as from the native-born Chicanos and Cholos, who were derided as people who had lost their Mexican roots. Students from this group were active in soccer and the Sociedad Bilingue and in celebrations of May 5th, the anniversary of the Mexican defeat of French occupying forces. Virtually all of the Mexican-descent students who graduated in the top 10 percent of their class in 1981 were identified as members of this group.

Chicanos were by far the largest Mexican-descent group at Field High. They were mostly U.S.-born second- and third-generation students whose primary loyalty was to their in-group, seen as locked in conflict with white society. Chicanos referred derisively to successful Mexican students as “schoolboys” and “schoolgirls” or as “wannabes.” According to M. G. Matute-Bianchi,

To be a Chicano meant in practice to hang out by the science wing . . . *not* eating lunch in the quad where all the “gringos” and “schoolboys” hang out . . . cutting classes by faking a call slip so you can be with your friends at the 7–11 . . . sitting in the back of classes and not participating . . . *not* carrying your books to class . . . *not* taking the difficult classes . . . doing the minimum to get by.²³

Chicanos merge imperceptibly into the last category, the Cholos, who were commonly seen as “low riders” and gang members. They were also native-born Mexican Americans, easily identifiable by their deliberate manner of dress, walk, speech, and other cultural symbols. Chicanos and Cholos were generally regarded by teachers as “irresponsible,” “disrespectful,” “mistrusting,”

“sullen,” “apathetic,” and “less motivated,” and their poor school performance was attributed to these traits.²⁴ According to Matute-Bianchi, Chicanos and Cholos were faced with what they saw as a forced-choice dilemma between doing well in school or being a Chicano. To act white was regarded as disloyalty to one’s group.

The situation of these last two groups exemplifies losing the race between first-generation achievements and later generations’ expectations. Seeing their parents and grandparents confined to humble menial jobs and increasingly aware of discrimination against them by the white mainstream, U.S.-born children of earlier Mexican immigrants readily join a reactive subculture as a means of protecting their sense of self-worth. Participation in this subculture then leads to serious barriers to their chances of upward mobility because school achievement is defined as antithetical to ethnic solidarity. Like Haitian students at Edison High, newly arrived Mexican students are at risk of being socialized into the same reactive stance, with the aggravating factor that it is other Mexicans, not native-born strangers, who convey the message. The principal protection of *mexicanos* against this type of assimilation lies in their strong identification with home-country language and values, which brings them closer to their parents’ cultural stance.

Punjabi Sikhs in California

Valleyside (a fictitious name) is a northern California community where the primary economic activity is orchard farming. Farm laborers in this area come often from India; they are mainly rural Sikhs from the Punjab. By the early 1980s, second-generation Punjabi students already accounted for 11 percent of the student body at Valleyside High. Their parents were no longer only farm laborers, since about a third had become orchard owners themselves and another third worked in factories in the nearby San Francisco area. An ethnographic study of Valleyside High School

in 1980–82 revealed a very difficult process of assimilation for Punjabi Sikh students. According to its author, M. A. Gibson, Valley-side is “redneck country,” and white residents are extremely hostile to immigrants who look different and speak a different language: “Punjabi teenagers are told they stink . . . told to go back to India . . . physically abused by majority students who spit at them, refuse to sit by them in class or in buses, throw food at them or worse.”²⁵

Despite these attacks and some evidence of discrimination by school staff, Punjabi students performed better academically than majority Anglo students. About 90 percent of the immigrant youths completed high school, compared to 70–75 percent of native whites. Punjabi boys surpassed the average grade point average, were more likely to take advanced science and math classes, and expressed aspirations for careers in science and engineering. Girls, on the other hand, tended to enroll in business classes, but they paid less attention to immediate career plans, reflecting parental wishes that they should marry first. This gender difference is indicative of the continuing strong influence exercised by the immigrant community over its second generation. According to Gibson, Punjabi parents pressured their children against too much contact with white peers who may “dishonor” the immigrants’ families, and defined “becoming Americanized” as forgetting one’s roots and adopting the most disparaged traits of the majority, such as leaving home at age 18, making decisions without parental consent, dating, and dancing. At the same time, parents urged children to abide by school rules, ignore racist remarks and avoid fights, and learn useful skills, including full proficiency in English.²⁶

The overall success of this strategy of selective assimilation to American society is remarkable because Punjabi immigrants were generally poor on their arrival in the United States and confronted widespread discrimination from whites without the benefit of ei-

ther governmental assistance or a well-established coethnic community. In terms of our typology of vulnerability and resources, the Punjabi Sikh second generation was very much at risk except for two crucial factors. First, immigrant parents did not settle in the inner city or in close proximity to any native-born minority whose offspring could provide an alternative model of adaptation to white-majority discrimination. In particular, the absence of a downtrodden Indian American community composed of children of previous immigrants allowed first-generation parents to influence decisively the outlook of their offspring, including their ways of fighting white prejudice. There was no equivalent of a Cholo-like reactive subculture to offer an alternative blueprint of the stance that “people like us” should take.

Second, Punjabi immigrants managed to make considerable economic progress, as attested by the number who had become farm owners, while maintaining a tightly knit ethnic community. The material and social capital created by this first-generation community compensated for the absence of an older coethnic group and had decisive effects on second-generation outlooks. Punjabi teenagers were shown that their parents’ ways paid off economically, and this fact, plus their community’s cohesiveness, endowed them with a source of pride to counteract outside discrimination. Through this strategy of selective assimilation, Punjabi Sikhs appeared to be winning the race against the inevitable acculturation of their children to American-style aspirations.

Caribbean Youths in South Florida

Miami is arguably the American city that has been most thoroughly transformed by post-1960 immigration. The Cuban Revolution had much to do with this transformation, as it sent the entire Cuban upper class out of the country, followed by thousands of refugees of more modest backgrounds. Over time, Cubans created a highly diversified and pros-

perous ethnic community that provided resources for the adaptation process of its second generation. Reflecting this situation are average Cuban family incomes that, by 1989, approximated those of the native-born population; the existence in 1987 of more than 30,000 Cuban-owned small businesses that formed the core of the Miami ethnic enclave; and the parallel rise of a private school system oriented toward the values and political outlook of this community.²⁷ In terms of the typology of vulnerability and resources, well-sheltered Cuban American teenagers lack any extensive exposure to outside discrimination, they have little contact with youths from disadvantaged minorities, and the development of an enclave creates economic opportunities beyond the narrowing industrial and tourist sectors on which most other immigrant groups in the area depend. Across town, Haitian American teenagers face exactly the opposite set of conditions, as has been shown.

Among the other immigrant groups that form Miami's ethnic mosaic, two deserve mention because they represent intermediate situations between those of the Cubans and Haitians. One comprises Nicaraguans escaping the Sandinista regime during the 1980s. They were not as welcomed in the United States as were the Cuban exiles, nor were they able to develop a large and diversified community. Yet they shared with Cubans their language and culture, as well as a militant anti-Communist discourse. This common political outlook led the Cuban American community to extend its resources in support of their Nicaraguan brethren, smoothing their process of adaptation.²⁸ For second-generation Nicaraguans, this means that the preexisting ethnic community that provides a model for their own assimilation is not a downtrodden group but rather one that has managed to establish a firm and positive presence in the city's economy and politics.

The second group comprises West Indians coming from Jamaica, Trinidad, and other English-speaking Caribbean republics. They

generally arrive in Miami as legal immigrants, and many bring along professional and business credentials as well as the advantage of fluency in English. These individual advantages are discounted, however, by a context of reception in which these mostly black immigrants are put in the same category as native-born blacks and discriminated against accordingly. The recency of West Indian migration and its small size have prevented the development of a diversified ethnic community in South Florida. Hence new arrivals experience the full force of white discrimination without the protection of a large coethnic group and with constant exposure to the situation and attitudes of the inner-city population. Despite considerable individual resources, these disadvantages put the West Indian second generation at risk of bypassing white or even native black middle-class models to assimilate into the culture of the underclass. . . .

Conclusion

Fifty years ago, the dilemma of Italian American youngsters studied by Irving Child consisted of assimilating into the American mainstream, sacrificing in the process their parents' cultural heritage in contrast to taking refuge in the ethnic community from the challenges of the outside world. In the contemporary context of segmented assimilation, the options have become less clear. Children of nonwhite immigrants may not even have the opportunity of gaining access to middle-class white society, no matter how acculturated they become. Joining those native circles to which they do have access may prove a ticket to permanent subordination and disadvantage. Remaining securely ensconced in their coethnic community, under these circumstances, may be not a symptom of escapism but the best strategy for capitalizing on otherwise unavailable material and moral resources. As the experiences of Punjabi Sikh and Cuban American students suggest, a strategy of paced, selective assimilation may prove the best

course for immigrant minorities. But the extent to which this strategy is possible also depends on the history of each group and its specific profile of vulnerabilities and resources. The present analysis represents a preliminary step toward understanding these realities.

NOTES

1. Alejandro Portes and Alex Stepick, *City on the Edge: The Transformation of Miami* (Berkeley: University of California Press, 1993), chap. 8.

2. Defined as native-born children with at least one foreign-born parent or children born abroad who came to the United States before age 12.

3. Joan N. First and John W. Carrera, *New Voices: Immigrant Students in U.S. Public Schools* (Boston: National Coalition of Advocates for Students, 1988).

4. Michael Piore, *Birds of Passage* (New York: Cambridge University Press, 1979); Herbert Gans, "Second-Generation Decline: Scenarios for the Economic and Ethnic Futures of the Post-1965 American Immigrants," *Ethnic and Racial Studies* 15:173-92 (Apr. 1992).

5. Irving L. Child, *Italian or American? The Second Generation in Conflict* (New Haven, CT: Yale University Press, 1943).

6. See, for example, Saskia Sassen, "Changing Composition and Labor Market Location of Hispanic Immigrants in New York City, 1960-1980," in *Hispanics in the U.S. Economy*, ed. George J. Borjas and Marta Tienda (New York: Academic Press, 1985), pp. 299-322.

7. See Alex Stepick, "Haitian Refugees in the U.S." (Report no. 52, Minority Rights Group, London, 1982); Alex Stepick and Alejandro Portes, "Flight into Despair: A Profile of Recent Haitian Refugees in South Florida," *International Migration Review*, 20:329-50 (Summer 1986).

8. This account is based on fieldwork in Miami conducted in preparation for a survey of immigrant youths in public schools.

9. On the issue of social capital, see James S. Coleman, "Social Capital in the Creation of Human Capital," *American Journal of Sociology*, supplement, 94:S95-121 (1988); Alejandro Portes and Min Zhou, "Gaining the Upper Hand: Economic Mobility among Immigrant and Domestic Minorities," *Ethnic and Racial Studies*, 15:491-522 (Oct. 1992). On ethnic entrepreneurship, see Ivan H. Light, *Ethnic Enterprise in America: Business and Welfare among Chinese, Japanese, and Blacks* (Berkeley: University of

California Press, 1972); Kenneth Wilson and W. Allen Martin, "Ethnic Enclaves: A Comparison of the Cuban and Black Economies in Miami," *American Journal of Sociology*, 88:135-60 (1982).

10. See W. Lloyd Warner and Leo Srole, *The Social Systems of American Ethnic Groups* (New Haven, CT: Yale University Press, 1945); Thomas Sowell, *Ethnic America: A History* (New York: Basic Books, 1981).

11. See Alejandro Portes and Rubén G. Rumbaut, *Immigrant America: A Portrait* (Berkeley: University of California Press, 1990), chap. 3.

12. See Mercer L. Sullivan, "Getting Paid": *Youth, Crime, and Work in the Inner City* (Ithaca, NY: Cornell University Press, 1989), chaps. 1, 5.

13. U.S. Department of Commerce, Bureau of the Census, *Census of Population and Housing, 1980: Public Use Microdata Samples A (MRDF)* (Washington, DC: Department of Commerce, 1983).

14. Thomas Bailey and Roger Waldinger, "Primary, Secondary, and Enclave Labor Markets: A Training System Approach," *American Sociological Review*, 56:432-45 (1991).

15. Professionals and executives represented 25.9 percent of Cuban-origin males aged 16 years and over in 1989; the figure for the total adult male population was 26 percent. See Jesus M. García and Patricia A. Montgomery, *The Hispanic Population of the United States: March 1990*, Current Population Reports, ser. P-20, no. 449 (Washington, DC: Department of Commerce, 1991).

16. Portes and Rumbaut, *Immigrant America*, pp. 23-25; Robert L. Bach et al., "The Economic Adjustment of Southeast Asian Refugees in the United States," in *World Refugee Survey, 1983* (Geneva: United Nations High Commission for Refugees, 1984), pp. 51-55.

17. The 1990 Immigration Act contains tailor-made provisions to facilitate the legalization of Irish immigrants. Those taking advantage of the provisions are popularly dubbed "Kennedy Irish" in honor of the Massachusetts Senator who coauthored the act. On the 1990 act, see Michael Fix and Jeffrey S. Passel, "The Door Remains Open: Recent Immigration to the United States and a Preliminary Analysis of the Immigration Act of 1990" (Working paper, Urban Institute and RAND Corporation, 1991). On the Irish in Boston, see Karen Tumulty, "When Irish Eyes Are Hiding. . .," *Los Angeles Times*, 29 Jan. 1989.

18. Bailey and Waldinger, "Primary, Secondary, and Enclave Labor Markets"; Min Zhou, *New York's Chinatown: The Socioeconomic Potential of an Urban Enclave* (Philadelphia: Temple University Press, 1992); Wilson and Martin, "Ethnic Enclaves";

Suzanne Model, "The Ethnic Economy: Cubans and Chinese Reconsidered" (Manuscript, University of Massachusetts at Amherst, 1990).

19. U.S. Department of Commerce, Bureau of the Census, *Survey of Minority-Owned Business Enterprises, 1987*, MB-2 and MB-3 (Washington, DC: Department of Commerce, 1991).

20. Alejandro Portes and Alex Stepick, "Unwelcome Immigrants: The Labor Market Experiences of 1980 (Mariel) Cuban and Haitian Refugees in South Florida," *American Sociological Review*, 50:493–514 (Aug. 1985); Zhou, *New York's Chinatown*; Luis E. Guarnizo, "One Country in Two: Dominican-Owned Firms in New York and the Dominican Republic" (Ph.D. diss. Johns Hopkins University, 1992); Bailey and Waldinger, "Primary, Secondary, and Enclave Labor Markets."

21. Stepick, "Haitian Refugees in the U.S.;" Jake C. Miller, *The Plight of Haitian Refugees* (New York: Praeger, 1984).

22. M. G. Matute-Bianchi, "Ethnic Identities and Patterns of School Success and Failure among Mexi-

can-Descent and Japanese-American Students in a California High School," *American Journal of Education*, 95:233–55 (Nov. 1986). This study is summarized in Rubén G. Rumbaut, "Immigrant Students in California Public Schools: A Summary of Current Knowledge" (Report no. 11, Center for Research on Effective Schooling for Disadvantaged Children, Johns Hopkins University, Aug. 1990).

23. Matute-Bianchi, "Ethnic Identities and Patterns," p. 253.

24. Rumbaut, "Immigrant Students," p. 25.

25. M. A. Gibson, *Accommodation without Assimilation: Sikh Immigrants in an American High School* (Ithaca, NY: Cornell University Press, 1989), p. 268.

26. Gibson, *Accommodation without Assimilation*. The study is summarized in Rumbaut, "Immigrant Students," pp. 22–23.

27. García and Montgomery, Hispanic Population; U.S. Department of Commerce, Bureau of the Census, *Survey of Minority-Owned Business Enterprises*, MB-2.

28. Portes and Stepick, *City on the Edge*, chap. 7.

69 ■ Mary C. Waters

Black Identities

West Indian Immigrant Dreams and American Realities

A key proposition of the new models of immigrant incorporation studies how the social capital immigrants bring with them, and the racial and ethnic definitions of nonwhite immigrants as minorities, combine to create a situation where becoming American in terms of culture and identity and achieving economic success are decoupled. Some immigrants and their children do better economically by maintaining a strong ethnic identity and culture and by

resisting American cultural and identity influences. In fact, many authors now suggest that remaining immigrant or ethnic-identified eases economic and social incorporation into the United States. These new assumptions turn models of identity change on their head—now those who resist becoming American do well and those who lose their immigrant ethnic distinctiveness become downwardly mobile. West Indians, it turns out, fit this model very well because

when West Indians lose their distinctiveness as immigrants or ethnics they become not just Americans, but black Americans. Given the ongoing prejudice and discrimination in American society, this represents downward mobility for the immigrants and their children. . . .

The main argument of this chapter is that black immigrants from the Caribbean come to the United States with a particular identity/culture/worldview that reflects their unique history and experiences. This culture and identity are different from the immigrant identity and culture of previous waves of European immigrants because of the unique history of the origin countries and because of the changed contexts of reception the immigrants face in the United States. This culture and identity are also different from the culture and identity of African Americans.

At first, two main aspects of the culture of West Indians help them to be successful in America. First, because they are immigrants they have a different attitude toward employment, work, and American society than native-born Americans. Employers value this highly. Their background characteristics, including human capital and social network ties, ease their entry into the U.S. labor force. Middle-class immigrants come with qualifications and training that are needed in the U.S. economy (nurses, for example). Because English is their native language, they are able to transfer their foreign qualifications (teaching credentials, nursing degrees) into American credentials. In addition, working-class immigrants have extensive networks of contacts that facilitate their entry into low-level jobs.

Second, the immigrants' unique understanding and expectations of race relations allow them to interact with American racial structures in a successful way. Specifically, their low anticipation of sour race relations allows them to have better interpersonal in-

teractions with white Americans than many native African Americans. Because they come from a society with a majority of blacks and with many blacks in high positions, the immigrants have high ambitions and expectations. Yet their experience with blocked economic mobility due to race and their strong racial identities lead them to challenge blocked mobility in a very militant fashion when they encounter it. This combination of high ambitions, friendly relations with whites on an interpersonal level, and strong militance in encountering any perceived discrimination leads to some better outcomes in the labor market for West Indians than for black Americans.

Ultimately, however, the structural realities of American race relations begin to swamp the culture of the West Indians. Persistent and obvious racial discrimination undermines the openness toward whites the immigrants have when they first arrive. Low wages and poor working conditions are no longer attractive to the children of the immigrants, who use American, not Caribbean, yardsticks to measure how good a job is. Racial discrimination in housing channels the immigrants into neighborhoods with inadequate city services and high crime rates. Inadequate public schools undermine their hopes for their children's future. Over time the distinct elements of West Indian culture the immigrants are most proud of—a willingness to work hard, a lack of attention to racialism, a high value on education, and strong interests in saving for the future—are undermined by the realities of life in the United States.

These changes are particularly concentrated among the working-class and poor immigrants. Middle-class immigrants are able to pass along aspects of their culture and worldview to their children, but the majority of the working-class immigrants are not. Race as a master status in the United States soon overwhelms the identities of the immi-

grants and their children, and they are seen as black Americans. Many of the children of the immigrants develop “oppositional identities” to deal with that status. These identities stress that doing well in school is “acting white.” The cultural behaviors associated with these oppositional identities further erode the life chances of the children of the West Indian immigrants.

While many white conservatives blame the culture of African Americans for their failures in the economy, the experiences of the West Indians show that even “good culture” is no match for racial discrimination. Over the course of one generation the structural realities of American race relations and the American economy undermine the cultures of the West Indian immigrants and create responses among the immigrants, and especially their children, that resemble the cultural responses of African Americans to long histories of exclusion and discrimination. . . .

The key factor brought to fresh light by the West Indian immigrants’ experiences is the role of continuing racial inequality—the institutional failures in our inner cities to provide jobs, education, and public safety—in sustaining a cultural response of disinvestment in the face of discrimination rather than increased striving. A lifetime of interpersonal attacks based on race can lead to bitterness and anger on the part of an individual. A community of people coping with economic marginality and a lack of any avenues of institutional support for individual mobility leads to a culture of opposition. That culture might serve individuals well for those times when it protects them from the sting of racism and discrimination, but ultimately as a long-term political response to discrimination and exclusion it serves to prevent people from taking advantage of the new opportunities that do arise. Those opportunities are reserved by whites in power for immigrants who make them feel less un-

comfortable about relations between the races and especially about taking orders from white supervisors or customers.

One of the African-American teachers [interviewed in this study] eloquently describes how even one act of cruelty or disdain by a white person can have long-standing effects on a young black person and on the whole cycle of black-white race relations:

I have had this happen to me so let me relate this incident. I have been going or coming from a building and held the door for some old white person and had them walk right past me, as if I am supposed to hold the door for them. Not one word, a thank-you, or an acknowledgment of your presence was made. A seventeen-year-old, when he has something like that happen to him, the next time he is going to slam that door in the old lady’s face, because I had that tendency myself. I had the hostility build up in me. The next time it happens I won’t do that because I am older, but when I was seventeen it might not have made any difference to me that this was a different old lady. I would have flashed back to that previous incident and said I am not holding that door for you. Now that old lady who may have been a perfectly fine individual, who got this door slammed in her face by this young black person, her attitude is “boy, all those people are really vicious people.” Her not understanding how it all came to pass. That on a large scale is what is happening in our country today. That’s why our young people are very aggressive and very, very hostile when they are put in a situation of black-white confrontation. They say to themselves, “I am not going to let you treat me the way you treated my grandparents, or the way I have read or seen in books or movies that they were treated.” I would rather for you to hate me than to disrespect me, is the attitude I think is coming out from our black youth today. (Black American male teacher, age 41)

The policy implications of this study lie in the ways in which the economic and cultural disinvestment in American cities erodes the social capital of immigrant families. The families need recognition of their inherent strengths and the supports necessary to maintain their ambitions. The erosion of the optimism and ambition of the first generation that I saw in their children could be stopped if job opportunities were more plentiful, inner-city schools were nurturing and safe environments that provided good educations, and neighborhoods were safer. Decent jobs, effective schools, and safe streets are not immigrant- or race-based policies. They are universal policies that would benefit all urban residents. Indeed, the strengths these immigrant families have may be in part due to their immigrant status, but the problems they face are much more likely to be due to their class status and their urban residence. Policies that benefit immigrants would equally benefit Americans.

But, in addition, the experience of these immigrants tells us that we must recognize the continuing significance of interpersonal

racism in creating psychological tensions and cultural adaptations in the black community. The cycle of attack and disrespect from whites, anger and withdrawal from blacks, and disengagement and blaming behaviors by whites must be broken by changing whites' behaviors. This involves policies that specifically address racial discrimination. The immigrants' tales of blatant housing and job discrimination directly point to needed vigilance in protecting all blacks in the United States from unequal treatment in the private sector. The more difficult problem is dealing with the everyday subtle forms of prejudice and discrimination that also plague foreign-born and American-born blacks. We cannot pass laws forbidding white women from clutching their handbags when black teenagers walk past them. We cannot require old white women to thank young black men who show them courteous behavior. Those kinds of behaviors can only change when whites no longer automatically fear blacks and when whites begin to perceive the humanity and diversity of the black people they encounter.

Discrimination

70 ■ Marianne Bertrand and Sendhil Mullainathan

Are Emily and Greg More Employable Than Lakisha and Jamal?

A Field Experiment on Labor Market Discrimination

Every measure of economic success reveals significant racial inequality in the U.S. labor market. Compared to Whites, African-Americans are twice as likely to be unemployed and earn nearly 25 percent less when they are employed (Council of Economic Advisers, 1998). This inequality has sparked a debate as to whether employers treat members of different races differentially. When faced with observably similar African-American and White applicants, do they favor the White one? Some argue yes, citing either employer prejudice or employer perception that race signals lower productivity. Others argue that differential treatment by race is a relic of the past, eliminated by some combination of employer enlightenment, affirmative action programs and the profit-maximization motive. In fact, many in this latter camp even feel that stringent enforcement of affirmative action programs has produced an environment of reverse discrimination. They would argue that faced with identical candidates, employers might favor the African-American one. Data limitations make it difficult to empirically test these views. Since researchers possess

far less data than employers do, White and African-American workers that appear similar to researchers may look very different to employers. So any racial difference in labor market outcomes could just as easily be attributed to differences that are observable to employers but unobservable to researchers.

To circumvent this difficulty, we conduct a field experiment that builds on the correspondence testing methodology that has been primarily used in the past to study minority outcomes in the United Kingdom. We send resumes in response to help-wanted ads in Chicago and Boston newspapers and measure callback for interview for each sent resume. We experimentally manipulate perception of race via the name of the fictitious job applicant. We randomly assign very White-sounding names (such as Emily Walsh or Greg Baker) to half the resumes and very African-American-sounding names (such as Lakisha Washington or Jamal Jones) to the other half. Because we are also interested in how credentials affect the racial gap in callback, we experimentally vary the quality of the resumes used in response to a given ad.

Marianne Bertrand and Sendhil Mullainathan. "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination," in *The American Economic Review* (September, 2004), pp. 991–992, 1006–1007, 1009–1013. Reprinted by permission of *The American Economic Review*.

Higher-quality applicants have on average a little more labor market experience and fewer holes in their employment history; they are also more likely to have an e-mail address, have completed some certification degree, possess foreign language skills, or have been awarded some honors. In practice, we typically send four resumes in response to each ad: two higher-quality and two lower-quality ones. We randomly assign to one of the higher- and one of the lower-quality resumes an African-American-sounding name. In total, we respond to over 1,300 employment ads in the sales, administrative support, clerical, and customer services job categories and send nearly 5,000 resumes. The ads we respond to cover a large spectrum of job quality, from cashier work at retail establishments and clerical work in a mail room, to office and sales management positions.

We find large racial differences in callback rates. Applicants with White names need to send about 10 resumes to get one callback whereas applicants with African-American names need to send about 15 resumes. This 50-percent gap in callback is statistically significant. A White name yields as many more callbacks as an additional eight years of experience on a resume. Since applicants' names are randomly assigned, this gap can only be attributed to the name manipulation.

Race also affects the reward to having a better resume. Whites with higher-quality resumes receive nearly 30-percent more callbacks than Whites with lower-quality resumes. On the other hand, having a higher-quality resume has a smaller effect for African-Americans. In other words, the gap between Whites and African-Americans widens with resume quality. While one may have expected improved credentials to alleviate employers' fear that African-American applicants are deficient in some unobservable skills, this is not the case in our data.

The experiment also reveals several other aspects of the differential treatment by race. First, since we randomly assign applicants'

postal addresses to the resumes, we can study the effect of neighborhood of residence on the likelihood of callback. We find that living in a wealthier (or more educated or Whiter) neighborhood increases callback rates. But, interestingly, African-Americans are not helped more than Whites by living in a "better" neighborhood. Second, the racial gap we measure in different industries does not appear correlated to Census-based measures of the racial gap in wages. The same is true for the racial gap we measure in different occupations. In fact, we find that the racial gaps in callback are statistically indistinguishable across all the occupation and industry categories covered in the experiment. Federal contractors, who are thought to be more severely constrained by affirmative action laws, do not treat the African-American resumes more preferentially; neither do larger employers or employers who explicitly state that they are "Equal Opportunity Employers." In Chicago, we find a slightly smaller racial gap when employers are located in more African-American neighborhoods. . . .¹

Interpretation

Three main sets of questions arise when interpreting the results above. First, does a higher callback rate for White applicants imply that employers are discriminating against African-Americans? Second, does our design only isolate the effect of race or is the name manipulation conveying some other factors than race? Third, how do our results relate to different models of racial discrimination?

Interpreting Callback Rates

Our results indicate that for two identical individuals engaging in an identical job search, the one with an African-American name would receive fewer interviews. Does differential treatment within our experiment imply that employers are discriminating against African-Americans (whether it is ra-

tional, prejudice-based, or other form of discrimination)? In other words, could the lower callback rate we record for African-American resumes *within our experiment* be consistent with a racially neutral review of the *entire pool* of resumes the surveyed employers receive?

In a racially neutral review process, employers would rank order resumes based on their quality and call back all applicants that are above a certain threshold. Because names are randomized, the White and African-American resumes we send should rank similarly on average. So, irrespective of the skill and racial composition of the applicant pool, a race-blind selection rule would generate equal treatment of Whites and African-Americans. So our results must imply that employers use race as a factor when reviewing resumes, which matches the legal definition of discrimination.

But even rules where employers are not trying to interview as few African-American applicants as possible may generate observed differential treatment in our experiment. One such hiring rule would be employers trying to interview a target level of African-American candidates. For example, perhaps the average firm in our experiment aims to produce an interview pool that matches the population base rate. This rule could produce the observed differential treatment if the average firm receives a higher proportion of African-American resumes than the population base rate because African-Americans disproportionately apply to the jobs and industries in our sample.

Some of our other findings may be consistent with such a rule. For example, the fact that “Equal Opportunity Employers” or federal contractors do not appear to discriminate any less may reflect the fact that such employers receive more applications from African-Americans. On the other hand, other key findings run counter to this rule. As we discuss above, we find no systematic difference in the racial gap in callback across

occupational or industry categories, despite the large variation in the fraction of African-Americans looking for work in those categories. African-Americans are underrepresented in managerial occupations, for example. If employers matched base rates in the population, the few African-Americans who apply to these jobs should receive a higher callback rate than Whites. Yet, we find that the racial gap in managerial occupations is the same as in all the other job categories. This rule also runs counter to our findings on returns to skill. Suppose firms are struggling to find White applicants but overwhelmed with African-American ones. Then they should be less sensitive to the quality of White applicants (as they are trying to fill in their hiring quota for Whites) and much more sensitive to the quality of Black applicants (when they have so many to pick from). Thus, it is unlikely that the differential treatment we observe is generated by hiring rules such as these.

Potential Confounds

While the names we have used in this experiment strongly signal racial origin, they may also signal some other personal trait. More specifically, one might be concerned that employers are inferring social background from the personal name. When employers read a name like “Tyroné” or “Latoya,” they may assume that the person comes from a disadvantaged background. In the extreme form of this social background interpretation, employers do not care at all about race but are discriminating only against the social background conveyed by the names we have chosen.

While plausible, we feel that some of our earlier results are hard to reconcile with this interpretation. For example, we found that while employers value “better” addresses, African-Americans are not helped more than Whites by living in Whiter or more educated neighborhoods. If the African-American names we have chosen mainly signal negative social background, one might have expected

the estimated name gap to be lower for better addresses. Also, if the names mainly signal social background, one might have expected the name gap to be higher for jobs that rely more on soft skills or require more interpersonal interactions. We found no such evidence.

There is one final potential confound to our results. Perhaps what appears as a bias against African-Americans is actually the result of *reverse discrimination*. If qualified African-Americans are thought to be in high demand, then employers with average quality jobs might feel that an equally talented African-American would never accept an offer from them and thereby never call her or him in for an interview. Such an argument might also explain why African-Americans do not receive as strong a return as Whites to better resumes, since higher qualification only strengthens this argument. But this interpretation would suggest that among the better jobs, we ought to see evidence of reverse discrimination, or at least a smaller racial gap. However, we do not find any such evidence. The racial gap does not vary across jobs with different skill requirements, nor does it vary across occupation categories. Even among the better jobs in our sample, we find that employers significantly favor applicants with White names.

Relation to Existing Theories

What do these results imply for existing models of discrimination? Economic theories of discrimination can be classified into two main categories: taste-based and statistical discrimination models. Both sets of models can obviously “explain” our average racial gap in callbacks. But can these models explain our other findings? More specifically, we discuss the relevance of these models with a focus on two of the facts that have been uncovered in this [chapter]: (i) the lower returns to credentials for African-Americans; (ii) the relative uniformity of the race gap across occupations, job requirements and, to a lesser extent, employer characteristics and industries.

Taste-based models (Gary S. Becker, 1961) differ in whose prejudiced “tastes” they emphasize: customers, coworkers, or employers. Customer and co-worker discrimination models seem at odds with the lack of significant variation of the racial gap by occupation and industry categories, as the amount of customer contact and the fraction of White employees vary quite a lot across these categories. We do not find a larger racial gap among jobs that explicitly require “communication skills” and jobs for which we expect either customer or coworker contacts to be higher (retail sales for example).

Because we do not know what drives employer tastes, employer discrimination models could be consistent with the lack of occupation and industry variation. Employer discrimination also matches the finding that employers located in more African-American neighborhoods appear to discriminate somewhat less. However, employer discrimination models would struggle to explain why African-Americans get relatively lower returns to their credentials. Indeed, the cost of indulging the discrimination taste should increase as the minority applicants’ credentials increase.

Statistical discrimination models are the prominent alternative to the taste-based models in the economics literature. In one class of statistical discrimination models, employers use (observable) race to proxy for *unobservable* skills (e.g., Edmund S. Phelps, 1972; Kenneth J. Arrow, 1973). This class of models struggles to explain the credentials effect as well. Indeed, the added credentials should lead to a larger update for African-Americans and hence greater returns to skills for that group.

A second class of statistical discrimination models “emphasize[s] the precision of the information that employers have about individual productivity” (Altonji and Blank, 1999). Specifically, in these models, employers believe that the same observable signal is more precise for Whites than for African-Americans (Denis J. Aigner and Glenn G. Cain, 1977; Shelly

J. Lundberg and Richard Startz, 1983; Bradford Cornell and Ivo Welch, 1996). Under such models, African-Americans receive lower returns to observable skills because employers place less weight on these skills. However, how reasonable is this interpretation for our experiment? First, it is important to note that we are using the same set of resume characteristics for both racial groups. So the lower precision of information for African-Americans cannot be that, for example, an employer does not know what a high school degree from a very African-American neighborhood means (as in Aigner and Cain, 1977). Second, many of the credentials on the resumes are in fact externally and easily verifiable, such as a certification for a specific software.

An alternative version of these models would rely on bias in the observable signal rather than differential variance or noise of these signals by race. Perhaps the skills of African-Americans are discounted because affirmative action makes it easier for African-Americans to get these skills. While this is plausible for credentials such as an employee-of-the-month honor, it is unclear why this would apply to more verifiable and harder skills. It is equally unclear why work experience would be less rewarded since our study suggests that getting a job is more, not less, difficult for African-Americans.

The uniformity of the racial gap across occupations is also troubling for a statistical discrimination interpretation. Numerous factors that should affect the level of statistical discrimination, such as the importance of unobservable skills, the observability of qualifications, the precision of observable skills and the ease of performance measurement, may vary quite a lot across occupations.

This discussion suggests that perhaps other models may do a better job at explaining our findings. One simple alternative model is lexicographic search by employers. Employers receive so many resumes that they may use quick heuristics in reading these resumes. One such heuristic could be to simply read

no further when they see an African-American name. Thus they may never see the skills of African-American candidates and this could explain why these skills are not rewarded. This might also to some extent explain the uniformity of the race gap since the screening process (i.e., looking through a large set of resumes) may be quite similar across the variety of jobs we study.

Conclusion

This paper suggests that African-Americans face differential treatment when searching for jobs and this may still be a factor in why they do poorly in the labor market. Job applicants with African-American names get far fewer callbacks for each resume they send out. Equally importantly, applicants with African-American names find it hard to overcome this hurdle in callbacks by improving their observable skills or credentials.

Taken at face value, our results on differential returns to skill have possibly important policy implications. They suggest that training programs alone may not be enough to alleviate the racial gap in labor market outcomes. For training to work, some general-equilibrium force outside the context of our experiment would have to be at play. In fact, if African-Americans recognize how employers reward their skills, they may rationally be less willing than Whites to even participate in these programs.

NOTE

1. For further details on this experiment, see Bertrand and Mullainathan (2004).

REFERENCES

- Aigner, Dennis J. and Cain, Glenn G. "Statistical Theories of Discrimination in Labor Markets." *Industrial and Labor Relations Review*, January 1977, 30(1), pp. 175–87.
- Altonji, Joseph G. and Blank, Rebecca M. "Race and Gender in the Labor Market," in Orley Ashenfelter

- and David Card, eds., *Handbook of labor economics*, Vol. 30. Amsterdam: North-Holland, 1999, pp. 3143–259.
- Arrow, Kenneth J. “The Theory of Discrimination,” in Orley Ashenfelter and Albert Rees, eds., *Discrimination in labor markets*. Princeton, NJ: Princeton University Press, 1973, pp. 3–33.
- Becker, Gary S. *The economics of discrimination*, 2nd Ed. Chicago: University of Chicago Press, 1961.
- Bertrand, Marianne and Sendhil Mullainathan. “Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination.” *American Economic Review* 94 (September 2004): 991–1013.
- Cornell, Bradford and Welch, Ivo. “Culture, Information, and Screening Discrimination.” *Journal of Political Economy*, June 1996, 104(3), pp. 542–71.
- Council of Economic Advisers. *Changing America: Indicators of social and economic well-being by race and Hispanic origin*. September 1998, <http://w3.access.gpo.gov/eop/ca/pdfs/ca.pdf>.
- Lundberg, Shelly J. and Starz, Richard. “Private Discrimination and Social Intervention in Competitive Labor Market.” *American Economic Review*, June 1983, 73(3), pp. 340–47.
- Phelps, Edmund S. “The Statistical Theory of Racism and Sexism.” *American Economic Review*, September 1972, 62(4), pp. 659–61.

71 ■ Claude Steele

Stereotype Threat and African-American Student Achievement

Over the past four decades African-American college students have been more in the spotlight than any other American students. This is because they aren't just college students; they are a cutting edge in America's effort to integrate itself in the nearly forty years since the passage of the Civil Rights Act. These students have borne much of the burden for our national experiment in racial integration. And to a significant degree the success of the experiment will be determined by their success.

Nonetheless, throughout the 1990s the national college dropout rate for African Americans has been 20 to 25 percent higher than that for whites. Among those who finish college, the grade point average of Black

students is two-thirds of a grade below that of whites. . . .

Virtually all aspects of underperformance—lower standardized test scores, lower college grades, lower graduation rates—persist among students from the African-American middle class. This situation forces on us an uncomfortable recognition: that beyond class, something racial is depressing the academic performance of these students.

Some time ago two of my colleagues, Joshua Aronson and Steven Spencer, and I tried to see the world from the standpoint of African-American students, concerning ourselves less with features of theirs that might explain their troubles than with features of the world they see. A story I was told recently

depicts some of these. The storyteller was worried about his friend, a normally energetic Black student who had broken up with his longtime girlfriend and had since learned that she, a Hispanic, was now dating a white student. This hit him hard. Not long after hearing about his girlfriend, he sat through an hour's discussion of *The Bell Curve* in his psychology class, during which the possible genetic inferiority of his race was openly considered. Then he overheard students at lunch arguing that affirmative action allowed in too many underqualified Blacks. By his own account, this young man had experienced very little of what he thought of as racial discrimination on campus. Still, these were features of his world. Could they have a bearing on his academic life?

My colleagues and I have called such features “stereotype threat”—the threat of being viewed through the lens of a negative stereotype, or the fear of doing something that would inadvertently confirm that stereotype. Everyone experiences stereotype threat. We are all members of some group about which negative stereotypes exist, from white males and Methodists to women and the elderly. And in a situation where one of those stereotypes applies—a man talking to women about pay equity, for example, or an aging faculty member trying to remember a number sequence in the middle of a lecture—we know that we may be judged by it.

Like the young man in the story, we can feel mistrustful and apprehensive in such situations. For him, as for African-American students generally, negative stereotypes apply in many situations, even personal ones. Why was that old roommate unfriendly to him? Did that young white woman who has been so nice to him in class not return his phone call because she's afraid he'll ask her for a date? Is it because of his race or something else about him? He cannot know the answers, but neither can his rational self fully dismiss the questions. Together they raise a deeper question: Will his

race be a boundary to his experience, to his emotions, to his relationships? . . .

Measuring Stereotype Threat

Can stereotype threat be shown to affect academic performance? And if so, who would be most affected—stronger or weaker students? Which has a greater influence on academic success among Black college students—the degree of threat or the level of preparation with which they enter college?

As we confronted these questions in the course of our research, we came in for some surprises. We began with what we took to be the hardest question: Could something as abstract as stereotype threat really affect something as irrepressible as intelligence? Ours is an individualistic culture; forward movement is seen to come from within. Against this cultural faith one needs evidence to argue that something as “sociological” as stereotype threat can repress something as “individualistic” as intelligence.

To acquire such evidence, Joshua Aronson and I (following a procedure developed with Steven Spencer) designed an experiment to test whether the stereotype threat that Black students might experience when taking a difficult standardized test could depress their performance on the test to a statistically reliable degree. We brought white and Black Stanford students into the laboratory and gave them, one at a time, a very difficult thirty-minute section of a Graduate Record Exam subject test in English literature. Most of these students were sophomores, which meant that the test—designed for graduating seniors—was particularly hard for them—precisely the feature, we reasoned, that would make this simple testing situation different for our Black participants than for our white participants, despite the fact that all the participants were of equal ability levels measured by all available criteria. (The difficulty of the test guaranteed that both Black and white students would find the test frustrating. And

it is in these situations that members of ability-stereotyped groups are most likely to experience the extra burden of stereotype threat. First, the experience of frustration with the test gives credibility to the limitation alleged in the stereotype. For this reason, frustration can be especially stinging and disruptive for test-takers to whom the stereotype is relevant. Second, it is on a demanding test that one can least afford to be bothered by the thoughts that likely accompany stereotype threat.)

A significant part of the negative stereotype about African Americans concerns intellectual ability. Thus, in the stereotype threat conditions of the experiments in this series, we merely mentioned to participants that the test was a measure of verbal ability. This was enough, we felt, to make the negative stereotype about African Americans' abilities relevant to their performance on the test, and thus to put them at risk of confirming, or being seen to confirm, the negative stereotype about their abilities. If the pressure imposed by the relevance of a negative stereotype about one's group is enough to impair an important intellectual performance, then Black participants should perform worse than whites in the "diagnostic" condition of this experiment but not in the "nondiagnostic" condition. As figure 1 depicts, this is precisely what happened: Blacks performed a full standard deviation lower than whites under the stereotype threat of the test being "diagnostic" of their intellectual ability, even though we had statistically matched the two groups in ability level. Something other than ability was involved; we believed it was stereotype threat.

But maybe the Black students performed less well than the white students because they were less motivated, or because their skills were somehow less applicable to the advanced material of this test. We needed some way to determine if it was indeed stereotype threat that depressed the Black students' scores. We reasoned that if stereotype threat had impaired their performance on the test, then reducing

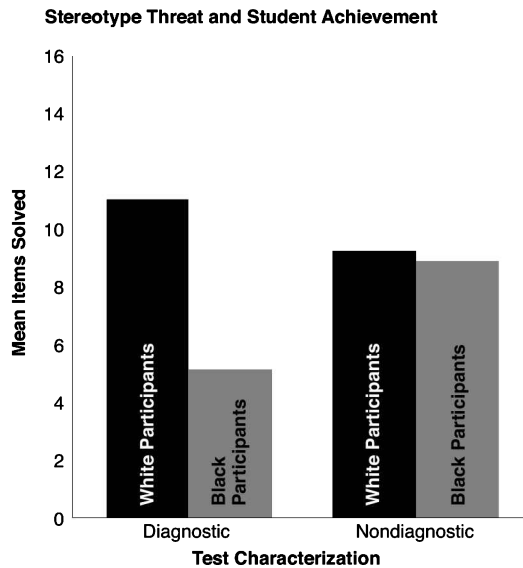


Figure 1. White and Black Participants' Score (Controlled for SAT) on a Difficult English Test as a Function of Characterization of the Test

this threat would allow their performance to improve. We presented the same test as a laboratory task that was used to study how certain problems are generally solved. We stressed that the task did not measure a person's level of intellectual ability. A simple instruction, yes, but it profoundly changed the meaning of the situation. In one stroke "spotlight anxiety," as the psychologist William Cross once called it, was turned off—and the Black students' performance on the test rose to match that of equally qualified whites (see figure 1). In the nonstereotype threat conditions, we presented the same test as an instrument for studying problem solving that was "nondiagnostic" of individual differences in ability—thus making the racial stereotype irrelevant to their performance.

Aronson and I decided that what we needed next was direct evidence of the subjective state we call stereotype threat. To seek this, we looked into whether simply sitting down to take a difficult test of ability was enough to make Black students mindful of their race and stereotypes about it. This may seem unlikely. White students I have taught over the years

have sometimes said that they have hardly any sense of even having a race. But Blacks have many experiences with the majority “other group” that make their race salient to them.

We again brought Black and white students in to take a difficult verbal test. But just before the test began, we gave them a long list of words, each of which had two letters missing. They were told to complete the words on this list as fast as they could. We knew from a preliminary survey that twelve of the eighty words we had selected could be completed in such a way as to relate to the stereotype about Blacks’ intellectual ability. The fragment “—ce,” for example, could become “race.” If simply taking a difficult test of ability was enough to make Black students mindful of stereotypes about their race, these students should complete more fragments with stereotype-related words. That is just what happened. When Black students were told that the test would measure ability, they completed the fragments with significantly more stereotype-related words than when they were told that it was not a measure of ability. Whites made few stereotype-related completions in either case. . . .

How Stereotype Threat Affects People Differently

Is everyone equally threatened and disrupted by a stereotype? One might expect, for example, that it would affect the weakest students most. But in all our research the most achievement-oriented students, who were also the most skilled, motivated, and confident, were the most impaired by stereotype threat. This fact had been under our noses all along—in our data and even in our theory. A person has to care about a domain in order to be disturbed by the prospect of being stereotyped in it. That is the whole idea of disidentification—protecting against stereotype threat by ceasing to care about the domain in which the stereotype applies. Our earlier experiments had selected Black students who

identified with verbal skills and women who identified with math. But when we tested participants who identified less with these domains, what had been under our noses hit us in the face. None of them showed any effect of stereotype threat whatsoever.

These weakly identified students did not perform well on the test: once they discovered its difficulty, they stopped trying very hard and got a low score. But their performance did not differ depending on whether they felt they were at risk of being judged stereotypically.

This finding, I believe, tells us two important things. The first is that the poorer college performance of Black students may have another source in addition to the one—lack of good preparation and, perhaps, of identification with school achievement—that is commonly understood. This additional source—the threat of being negatively stereotyped in the environment—has not been well understood. The distinction has important policy implications: different kinds of students may require different pedagogies of improvement.

The second thing is poignant: what exposes students to the pressure of stereotype threat is not weaker academic identity and skills but stronger academic identity and skills. They may have long seen themselves as good students—better than most. But led into the domain by their strengths, they pay an extra tax on their investment—vigilant worry that their future will be compromised by society’s perception and treatment of their group.

This tax has a long tradition in the Black community. The Jackie Robinson story is a central narrative of Black life, literature, and journalism. *Ebony Magazine* has run a page for fifty years featuring people who have broken down one or another racial barrier. Surely the academic vanguard among Black college students today knows this tradition—and knows, therefore, that the thing to do, as my father told me, is to buckle down, pay whatever tax is required, and disprove the damn stereotype.

That, however, seems to be precisely what these students are trying to do. In some of our

experiments we administered the test of ability by computer, so that we could see how long participants spent looking at different parts of the test questions. Black students taking the test under stereotype threat seemed to be trying too hard rather than not hard enough. They reread the questions, reread the multiple choices, rechecked their answers, more than when they were not under stereotype threat. The threat made them inefficient on a test that, like most standardized tests, is set up so that thinking long often means thinking wrong, especially on difficult items like the ones we used. . . .

In the old song about the “steel-drivin’ man,” John Henry races the new steam-driven drill to see who can dig a hole faster. When the race is over, John Henry has prevailed by digging the deeper hole—only to drop dead. The social psychologist Sherman James uses the term “John Henryism” to describe a psychological syndrome that he found to be associated with hypertension in several samples of North Carolina Blacks: holding too rigidly to the faith that discrimination and disadvantage can be overcome with hard work and persistence. Certainly this is the right attitude. But taken to extremes, it can backfire. A deterioration of performance under stereotype threat by the skilled, confident Black students in our experiments may be rooted in John Henryism.

This last point can be disheartening. Our research, however, offers an interesting suggestion about what can be done to overcome stereotype threat and its detrimental effects. The success of Black students may depend less on expectations and motivation—things that are thought to drive academic performance—than on trust that stereotypes about their group will not have a limiting effect in their school world.

Putting this idea to the test, Joseph Brown and I asked, How can the usual detrimental effect of stereotype threat on the standardized-test performance of these students be reduced? By strengthening students’ expectations and confidence, or by strengthening their trust that

they are not at risk of being judged on the basis of stereotypes? In the ensuing experiment we strengthened or weakened participants’ confidence in their verbal skills, by arranging for them to have either an impressive success or an impressive failure on a test of verbal skills, just before they took the same difficult verbal test we had used in our earlier research. When the second test was presented as a test of ability, the boosting or weakening of confidence in their verbal skills had no effect on performance: Black participants performed less well than equally skilled white participants. What does this say about the commonsense idea that black students’ academic problems are rooted in lack of self-confidence?

What did raise the level of black students’ performance to that of equally qualified whites was reducing stereotype threat—in this case by explicitly presenting the test as racially fair. When this was done, Blacks performed at the same high level as whites even if their self-confidence had been weakened by a prior failure.

These results suggest something that I think has not been made clear elsewhere: when strong Black students sit down to take a difficult standardized test, the extra apprehension they feel in comparison with whites is less about their own ability than it is about having to perform on a test and in a situation that may be primed to treat them stereotypically. We discovered the extent of this apprehension when we tried to develop procedures that would make our Black participants see the test as “race-fair.” It wasn’t easy. African-Americans have endured so much bad press about test scores for so long that, in our experience, they are instinctively wary about the tests’ fairness. We were able to convince them that our test was race-fair only when we implied that the research generating the test had been done by Blacks. When they felt trust, they performed well regardless of whether we had weakened their self-confidence beforehand. And when they didn’t feel trust, no amount of bolstering of self-confidence helped.

REFERENCES

- Aronson, J. 1997. *The Effects of Conceptions of Ability on Task Valuation*. Unpublished manuscript, New York University.
- _____. 1999. *The Effects of Conceiving Ability as Fixed or Improvable on Responses to Stereotype Threat*. Unpublished manuscript, New York University.
- Brown, J. L., and C. M. Steele. 2001. *Performance Expectations Are Not a Necessary Mediator of Stereotype Threat in African American Verbal Test Performance*. Unpublished manuscript, Stanford University.
- Cross, W. E., Jr. 1991. *Shades of Black: Diversity in African-American Identity*. Philadelphia: Temple University Press.
- Spencer, S. J., E. Iserman, P. G. Davies, and D. M. Quinn. 2001. *Suppression of Doubts, Anxiety, and Stereotypes as a Mediator of the Effect of Stereotype Threat on Women's Math Performance*. Unpublished manuscript, University of Waterloo.
- Spencer, S. J., C. M. Steele, and D. M. Quinn. 1999. "Stereotype Threat and Women's Math Performance." *Journal of Experimental Social Psychology* 35: 4–28.
- Steele, C. M. 1975. "Name-Calling and Compliance." *Journal of Personality and Social Psychology* 31: 361–69.
- _____. 1992. "Race and the Schooling of Black Americans." *Atlantic Monthly* 269: 68–78.
- _____. 1997. "A Threat in the Air: How Stereotypes Shape Intellectual Identity and Performance." *American Psychologist* 52: 613–29.
- _____. 1999. "Thin Ice: Stereotype Threat and Black College Students." *Atlantic Monthly* 248: 44–54.
- Steele, C. M., and J. Aronson. 1995. "Stereotype Threat and the Intellectual Test Performance of African Americans." *Journal of Personality and Social Psychology* 69: 797–811.
- Steele, C. M., S. J. Spencer, and J. Aronson. 2002. "Contending with Group Image: The Psychology of Stereotype and Social Identity Threat." In *Advances in Experimental Social Psychology*, vol. 34, ed. M. Zanna. Academic Press.
- Steele, C. M., S. J. Spencer, P. G. Davies, K. Harber, and R. E. Nisbett. 2001. *African American College Achievement: A "Wise" Intervention*. Unpublished manuscript, Stanford University.
- Steele, S. 1990. *The Content of Our Character: A New Vision of Race in America*. New York: St. Martin's.

72 ■ Devah Pager

Marked

Race, Crime, and Finding Work in an Era of Mass Incarceration

While stratification researchers typically focus on schools, labor markets, and the family as primary institutions affecting inequality, a new institution has emerged as central to the sorting and stratifying of young disadvantaged men: the criminal justice system. With over two million individuals currently incarcerated, and well over half a million prisoners released each year, the large and growing numbers of

men being processed through the criminal justice system raises important questions about the consequences of this massive institutional intervention. Further, large racial disparities in incarceration lead us to question the degree to which mass incarceration has particular implications for stratification along racial lines.

This [chapter] focuses on the consequences of incarceration for the employment outcomes

This is a commissioned chapter that draws heavily on material in a previous publication (Devah Pager, "The Mark of a Criminal Record," *American Journal of Sociology* 108 [2003], pp. 937–975).

of black and white men. I adopt an experimental audit approach to formally test the degree to which race and criminal background affect subsequent employment opportunities. By using matched pairs of individuals to apply for real entry-level jobs, it becomes possible to directly measure the extent to which contact with the criminal justice system—in the absence of other disqualifying characteristics—serves as a barrier to employment among equally qualified applicants. Further, by varying the race of the tester pairs, it becomes possible to assess the ways in which the effects of race and criminal record interact to produce new forms of labor market inequalities.

Mass Incarceration and the Credentialing of Stigma

Over the past three decades, the number of prison inmates has increased by more than 700 percent, leaving the United States the country with the highest incarceration rate in the world (Bureau of Justice Statistics, 2006; Mauer, 1999). During this time, incarceration changed from a punishment reserved primarily for the most heinous offenders to one extended to a much greater range of crimes and a much larger segment of the population. Recent trends in crime policy have led to the imposition of harsher sentences for a wider range of offenses, thus casting an ever widening net of penal intervention.

For each individual processed through the criminal justice system, police records, court documents, and corrections databases detail dates of arrest, charges, conviction, and terms of incarceration. Most states make these records publicly available, often through online repositories, accessible to employers, landlords, creditors, and other interested parties. With increasing numbers of occupations, public services, and other social goods becoming off-limits to ex-offenders, these records can be used as the official basis for eligibility determination or exclusion. The state, in this way, serves as a credentialing institution, providing

official and public certification of those among us who have been convicted of wrongdoing. The “credential” of a criminal record, like educational or professional credentials, constitutes a formal and enduring classification of social status, which can be used to regulate access and opportunity across numerous social, economic, and political domains. As increasing numbers of young men are marked by their contact with the criminal justice system, it becomes a critical priority to understand the costs and consequences of this now prevalent form of negative credential (Pager, 2007b).

Racial Stereotypes in an Era of Mass Incarceration

The expansion of the correctional population has been particularly consequential for blacks. Blacks today represent over 40 percent of current prison inmates relative to just 12 percent of the U.S. population. Over the course of a lifetime, nearly one in three young black men—and well over half of young black high school dropouts—will spend some time in prison. According to these estimates, young black men are more likely to go to prison than to attend college, serve in the military, or, in the case of high school dropouts, to be in the labor market (Bureau of Justice Statistics, 1997; Pettit and Western, 2004). Prison is no longer a rare or extreme event among our nation’s most marginalized groups. Rather it has now become a normal and anticipated marker in the transition to adulthood.

There is reason to believe that the implications of these trends extend well beyond the prison walls, with assumptions of criminality among blacks generalizing beyond those directly engaged in crime. Blacks in this country have long been regarded with suspicion and fear; but unlike progressive trends in other racial attitudes, associations between race and crime have changed little in recent years. Survey respondents consistently rate blacks as more prone to violence than any other American racial or ethnic group, with the stereotype

of aggressiveness and violence most frequently endorsed in ratings of African Americans (Sniderman and Piazza, 1993; Smith, 1990). The stereotype of blacks as criminals is deeply embedded in the collective consciousness of white Americans, irrespective of the perceiver's level of prejudice or personal beliefs (Devine and Elliot, 1995; Eberhardt et al., 2004; Graham and Lowery, 2004).

While it would be impossible to trace the source of contemporary racial stereotypes to any one factor, the disproportionate growth of criminal justice intervention in the lives of young black men—and corresponding media coverage of this phenomenon, which presents an even more skewed representation—has likely played an important role. Experimental research shows that exposure to news coverage of a violent incident committed by a black perpetrator not only increases punitive attitudes about crime but further increases negative attitudes about blacks generally (Gilliam and Iyengar, 2000). The more exposure we have to images of blacks in custody or behind bars, the stronger our expectations become regarding the race of assailants or the criminal tendencies of black strangers (Cole, 1995).

The consequences of mass incarceration then may in fact extend far beyond the costs to the individual bodies behind bars, and to the families that are disrupted or the communities whose residents cycle in and out. The criminal justice system may itself legitimate and reinforce deeply embedded racial stereotypes, contributing to the persistent chasm in this society between black and white.

Assessing the Impact of Race and Criminal Background

In considering the labor market impacts of race and criminal background, questions of causality loom large. For example, while employment disparities between blacks and whites have been well established, the causes of these disparities remain widely contested. Where racial discrimination may represent

one explanation, a growing number of researchers have pointed instead to individual-level factors such as skill deficits and other human capital characteristics as a key source of racial wage differentials (Neal and Johnson, 1996; Farkas and Vicknair, 1996). Likewise, assessing the impact of incarceration is not altogether straightforward. On the one hand, it's not hard to imagine that a prison record would carry a weighty stigma, with members of the general public (employers included) reluctant to associate or work with former inmates. On the other hand, criminal offenders aren't typically the image of the model employee. It's certainly possible that the poor employment outcomes of ex-offenders stem rather from characteristics of the offenders themselves, as opposed to any consequence of their criminal conviction. Poor work habits, substance abuse problems, or deficient interpersonal skills may be sufficient to explain the employment disadvantages of ex-offenders, and yet these characteristics are difficult to capture using standard data sources.

Given the difficulties inherent to evaluating the impact of race and imprisonment through conventional measures, I set out to investigate this issue by constructing an experiment. I wanted to bracket the range of personal characteristics associated with African Americans and ex-offenders in order to home in on the causal impact of race and a criminal record. The experimental audit methodology allowed me to do just that. Using this approach, I pose the questions: Given two equally qualified job applicants, how much does a criminal record affect the chances of being selected by an employer? To what extent does race condition employers' responses? In answering these questions, we can begin to understand the ways in which race and criminal background shape and constrain important economic opportunities.

The Audit Methodology

The basic design of an employment audit involves sending matched pairs of individuals

(called testers) to apply for real job openings in order to see whether employers respond differently to applicants on the basis of selected characteristics. The current study included four male testers, two blacks and two whites, matched into two teams—the two black testers formed one team, and the two white testers formed a second team.¹ The testers were college students from Milwaukee who were matched on the basis of age, race, physical appearance, and general style of self-presentation. The testers were assigned fictitious resumes that reflected equivalent levels of education and work experience.² In addition, within each team, one auditor was randomly assigned a “criminal record” for the first week; the pair then rotated which member presented himself as the ex-offender for each successive week of employment searches, such that each tester served in the criminal record condition for an equal number of cases.³ By varying which member of the pair presented himself as having a criminal record, unobserved differences within the pairs of applicants were effectively controlled.

Before initiating the fieldwork, testers participated in a common training program to become familiar with the details of their assumed profile and to ensure uniform behavior in job interviews. The training period lasted for one week, during which testers participated in mock interviews with one another and practice interviews with cooperating employers. The testers were trained to respond to common interview questions in standardized ways, and were well-rehearsed for a wide range of scenarios that emerge in hiring situations. Frequent communication between myself and the testers throughout each day of fieldwork allowed for regular supervision and troubleshooting in the event of unexpected occurrences.

A random sample of entry-level positions (those jobs requiring no previous experience and no education beyond high school) was drawn each week from the Sunday classified advertisement section of the *Milwaukee Journal Sentinel*. The most common job titles in-

cluded waitstaff, laborers and warehouse, production/operators, customer service, sales, delivery drivers, and cashiers; a handful of clerical and managerial positions were also included. I excluded from the sample those occupations with legal restrictions on ex-offenders, such as jobs in the health care industry, work with children and the elderly, jobs requiring the handling of firearms (e.g., security guards), and jobs in the public sector. Of course, any true estimate of the collateral consequences of incarceration would also need to take account of the wide range of employment fully off-limits to individuals with prior felony convictions.

Each of the audit pairs was randomly assigned fifteen job openings each week. The white pair and the black pair were assigned separate sets of jobs, with the same-race testers applying to the same jobs. One member of the pair applied first, with the second applying one day later (randomly varying whether the ex-offender was first or second). A total of 350 employers were audited during the course of this study: 150 by the white pair and 200 by the black pair. Additional tests were performed by the black pair because black testers received fewer callbacks on average, and there were thus fewer data points with which to draw comparisons. A larger sample size enables the calculation of more precise estimates of the effects under investigation (for a more in-depth discussion of the study design, see Pager, 2003).

This study focused only on the first stage of the employment process, as this is the stage most likely to be affected by the barriers of race and criminal record. Testers visited employers, filled out applications, and proceeded as far as they could during the course of one visit. If testers were asked to interview on the spot, they did so, but they did not return to the employer for a second visit. The primary dependent variable then is the proportion of applications that elicited callbacks from employers. Individual voice mail boxes were set up for each tester to record employer responses.



The effects of race and criminal record are large and statistically significant ($p < .01$). The interaction between the two is not significant in the full sample.

Figure 1. The Effects of Race and Criminal Background on Employment

An advantage of the call-back as our key outcome variable—as opposed to a job offer—is that it does not require employers to narrow their selection down to a single applicant. At the job offer stage, if presented with an ex-offender and an equally qualified nonoffender, even employers with little concern over hiring ex-offenders would likely select the applicant with no criminal record, an arguably safer choice. Equating the two applicants could in fact magnify the impact of the criminal record, as it becomes the only remaining basis for selection between the two (Heckman, 1998). The call-back, by contrast, does not present such complications. Typically employers interview multiple candidates for entry-level positions before selecting a hire. In fact, in a subsequent survey, employers in this study reported interviewing an average of eight applicants for the last entry-level position filled. At the call-back stage, then, employers need not yet choose between the ex-offender and nonoffender. If the applicants appear well-qualified, and if the employer does not view the criminal record as an automatic disqualifier, s/he can interview them both.⁴

Hiring Outcomes by Race and Criminal Background

Results are based on the proportion of applications submitted by each tester which

elicited callbacks from employers. Three main findings should be noted from the audit results, presented in Figure 1. First, there is a large and significant effect of a criminal record for all job seekers. Among whites, for example, 34 percent of those without criminal records received callbacks relative to only 17 percent of those with criminal records. A criminal record is thus associated with a 50 percent reduction in the likelihood of a callback among whites. Often testers reported seeing employers' levels of responsiveness change dramatically once they had glanced down at the criminal record question on the application form. Many employers seem to use the information as a screening mechanism, weeding out applicants at the initial stage of review.

Second, there is some indication that the magnitude of the criminal record effect may be even larger for blacks. While the interaction between race and criminal record is not statistically significant, the substantive difference is worth noting. The ratio of callbacks for nonoffenders relative to offenders for whites was two to one (34 percent vs. 17 percent), while this same ratio for blacks is close to three to one (14 percent vs. 5 percent).⁵ The estimated effect of a criminal record is thus 40 percent larger for blacks than for whites. The combination of minority status and criminal background appears to intensify employers'

negative reactions, leaving few employment prospects for black ex-offenders (200 applications resulted in only 10 callbacks).

Finally, looking at the callback rates for black and white tester pairs side by side, the fundamental importance of race becomes vividly clear. Among those without criminal records, black applicants were less than half as likely to receive callbacks compared to equally qualified whites (14 percent vs. 34 percent). This implies that young black men needed to work more than twice as hard (apply to twice as many jobs) to secure the same opportunities as whites with identical qualifications. Even more striking, the powerful effects of race rival even the strong stigma conveyed by a criminal record. In this study, a white applicant *with a criminal record* was just as likely, if not more, to receive a callback as a black applicant with no criminal history (17 percent vs. 14 percent). Despite the fact that the white applicant revealed evidence of a felony drug conviction, and despite the fact that he reported having only recently returned from a year and a half in prison, employers seemed to view this applicant as no more risky than a young black man with no history of criminal involvement. Racial disparities have been documented in many contexts, but here, comparing the two effects side by side, we are confronted with a troubling reality: Being black in America today is just about the same as having a felony conviction in terms of one's chances of finding a job.⁶

The results of this research suggest that both race and criminal background represent extremely powerful barriers to job entry. The matched design allows us to separate speculation about applicant qualifications (supply-side influences) from the preferences or biases of employers (demand-side influences). While this study remains silent on the many supply-side factors that may also contribute to the employment difficulties of blacks and ex-offenders, it speaks loud and clear about the significance of em-

ployer demand in shaping the opportunities available to job seekers on the basis of their race or criminal background. Before applicants have an opportunity to demonstrate their capabilities in person, a large proportion are weeded out on the basis of a single categorical distinction.

Discussion

There is serious disagreement among academics, policy makers, and practitioners over the extent to which contact with the criminal justice system—in itself—leads to harmful consequences for employment. The present study takes a strong stand in this debate by offering direct evidence of the causal relationship between a criminal record and employment outcomes. Using matched pairs and an experimentally assigned criminal record, this estimate is unaffected by the problems of selection which plague observational data. While certainly there are additional ways in which incarceration may affect economic outcomes, this study provides convincing evidence that mere contact with the criminal justice system, in the absence of any transformative or selective effects, severely limits subsequent employment opportunities. And while the audit study investigates employment barriers to ex-offenders from a micro-perspective, the implications are far-reaching. The finding that ex-offenders are one-half to one-third as likely to be considered by employers suggests that a criminal record indeed presents a major barrier to employment. With over two million people currently behind bars and over 12 million people with prior felony convictions, the consequences for labor market inequalities are potentially profound.

Second, the persistent effect of race on employment opportunities is painfully clear in these results. Blacks are less than half as likely to receive consideration by employers relative to their white counterparts, and black nonoffenders fall behind even whites with prior

felony convictions. While this research cannot identify the precise source of employers' reluctance to hire blacks, it is important to consider the influence of high incarceration rates among blacks as one possible source of racial bias. Indeed, the available evidence points to the pervasiveness of images associating blacks with crime, and the power of these images to strengthen negative feelings about blacks as a group. It may be the case, then, that increasing rates of incarceration among blacks, and the disproportionate coverage of criminal justice contact among blacks in the media, heighten negative reactions toward African Americans generally, irrespective of their personal involvement in crime.

No longer a peripheral institution, the criminal justice system has become a dominant presence in the lives of young disadvantaged men, playing a key role in the sorting and stratifying of labor market opportunities. The "criminal credential" now represents a common marker among young disadvantaged men, allowing for the easy identification and exclusion of those with a prior arrest or conviction. Further, because blacks are so strongly associated with the population under correctional supervision, it becomes common to assume that any given young black man is likely to have—or to be on his way to acquiring—a criminal record.

At this point in history, it is impossible to tell whether the massive presence of incarceration in today's stratification system represents a unique anomaly of the late twentieth century or part of a larger movement toward a system of stratification based on the official certification of individual character and competence. Whether this process will continue to form the basis of emerging social cleavages remains to be seen.

NOTES

1. The two teams were sent to different employers in order to minimize suspicion that might otherwise arise from two similar applicants with similar criminal backgrounds applying within a short

time interval. In the present design, the criminal record effect is estimated as a within-team effect while race is measured as a between-team effect. While the latter is less efficient, the comparison should be unbiased. Black and white testers used *identical* resumes (something that would not have been possible had they visited the same employers) and were selected and trained to present comparable profiles to employers. Likewise, employers were randomly assigned across tester pairs, minimizing heterogeneity in employer characteristics by team.

2. Testers presented themselves as high school graduates with steady work experience in entry level jobs.

3. The criminal record in all cases was a drug felony, possession with intent to distribute (cocaine), and 18 months of prison time. Testers presented the information to employers by checking the box "yes" in answer to the standard application question, "Have you ever been convicted of a crime?" As additional cues, testers also reported work experience in the correctional facility and listed their parole officer as a reference (see Pager, 2003 for a more in-depth discussion of these issues).

4. A more in-depth discussion of methodological concerns, including limits to generalizability, representativeness of testers, sample restrictions, and experimenter effects is presented in Pager (2007b), appendix 4A.

5. While not significant in the full sample, this interaction becomes significant when analyzed specifically for suburban employers or among employers with whom the testers had extended personal contact (see Pager, 2007b, chap. 7).

6. These results do not appear specific to Milwaukee. Other audit studies have revealed effects of similar magnitude (see Pager, 2007a for a summary); and a replication of the present study in New York City comes to similar conclusions (Pager and Western, 2005).

REFERENCES

- Bureau of Justice Statistics. 2002. "Recidivism of Prisoners Released in 1994," edited by Patrick Langan and David Levin. Washington, DC: U.S. Department of Justice.
- _____. 2006. "Prison and Jail Inmates at Midyear 2005." Washington, DC: U.S. Department of Justice.
- Bureau of Justice Statistics Special Report 1997. March. *Lifetime Likelihood of Going to State or Federal Prison*, by Thomas P. Bonczar and Allen

- J. Beck. Washington, DC: U.S. Department of Justice.
- Cole, David. 1995. "The paradox of race and crime: A comment on Randall Kennedy's 'politics of distinction'." *Georgetown Law Journal* 83:2547–2571.
- Devine, P. G., and A. J. Elliot. 1995. "Are Racial Stereotypes Really Fading? The Princeton Trilogy Revisited." *Personality and Social Psychology Bulletin* 21(11): 1139–1150.
- Eberhardt, Jennifer L., Phillip Atiba Goff, Valerie J. Purdie, and Paul G. Davies. 2004. "Seeing black: Race, crime, and visual processing." *Journal of Personality & Social Psychology* 87:876–893.
- Farkas, George, and Kevin Vicknair. 1996. "Appropriate tests of racial wage discrimination require controls for cognitive skill: comment on Cancio, Evans, and Maume." *American Sociological Review* 61:557–560.
- Freeman, Richard B. 1987. "The Relation of Criminal Activity to Black Youth Employment." *Review of Black Political Economy* 16(1–2): 99–107.
- Gilliam, Franklin D., and Shanto Iyengar. 2000. "Prime suspects: The influence of local television news on the viewing public." *American Journal of Political Science* 44:560–573.
- Graham, Sandra, and Brian S. Lowery. 2004. "Priming unconscious racial stereotypes about adolescent offenders." *Law and Human Behavior* 28:483–504.
- Heckman, James J. 1998. "Detecting discrimination." *The Journal of Economic Perspectives* 12:101–116.
- Heckman, James, and Peter Seligman. 1993. "The Urban Institute Audit Studies: Their Methods and Findings." In Michael Fix and Raymond J. Struyk (eds.), *Clear and Convincing Evidence: Measurement of Discrimination in America*. Washington, DC: Urban Institute Press.
- Mauer, Marc. 1999. *Race to Incarcerate*. New York: New Press.
- Neal, Derek, and William Johnson. 1996. "The Role of Pre-market Factors in Black-White Wage Differences." *Journal of Political Economy*, 104(5): 869–895.
- Pager, Devah. 2003. "The Mark of a Criminal Record." *American Journal of Sociology* 108(5): 937–975.
- _____. 2007a. "The Use of Field Experiments for Studies of Employment Discrimination: Contributions, Critiques, and Directions for the Future." *Annals of the American Academy of Political and Social Science* 609: 104–133.
- _____. 2007b. *Marked: Race, Crime, and Finding Work in an Era of Mass Incarceration*. Chicago: University of Chicago Press.
- Pager, Devah, and Bruce Western. 2005. "Discrimination in Low Wage Labor Markets: Evidence from New York City." Paper presented at Population Association of America, Philadelphia, PA.
- Pettit, Becky, and Bruce Western. 2004. "Mass Imprisonment and the Life Course: Race and Class Inequality in U.S. Incarceration." *American Sociological Review* 69: 151–169.
- Sampson, Robert J., and John H. Laub. 1993. *Crime in the Making: Pathways and Turning Points Through Life*. Cambridge, MA: Harvard University Press.
- Schwartz, Richard, and Jerome Skolnick. 1962. "Two Studies of Legal Stigma." *Social Problems*, Fall: 133–142.
- Smith, Tom W. 1991. "Ethnic Images. General Social Survey Technical Report, 19." Chicago: National Opinion Research Center, University of Chicago.
- Sniderman, Paul M., and Thomas Piazza. 1993. *The Scar of Race*. Cambridge, MA: Harvard University Press.
- Travis, Jeremy, Amy Solomon, and Michelle Waul. 2001. *From Prison to Home: The Dimensions and Consequences of Prisoner Reentry*. Washington, DC: Urban Institute Press.
- Uggen, Christopher. 2000. "Work as a Turning Point in the Life Course of Criminals: A Duration Model of Age, Employment, and Recidivism." *American Sociological Review* 65(4): 529–546.
- Uggen, Christopher, Jeff Manza, and Melissa Thompson. 2006. "Citizenship and Reintegration: The Socioeconomic, Familial, and Civic Lives of Criminal Offenders." *The Annals of the American Academy of Social and Political Science* 605:281–310.
- Western, Bruce. 2002. "The Impact of Incarceration on Wage Mobility and Inequality." *American Sociological Review* 67(4): 526–546.
- Western, Bruce, and Katherine Beckett. 1999. "How Unregulated Is the U.S. Labor Market? The Penal System as a Labor Market Institution." *American Journal of Sociology* 104(4): 1030–60.
- Wilson, William Julius. 1996. *When Work Disappears: The World of the New Urban Poor*. New York: Vintage Books.

Are Racial and Ethnic Distinctions Declining in Significance?

73 ■ William Julius Wilson

The Declining Significance of Race Blacks and Changing American Institutions

Race relations in America have undergone fundamental changes in recent years, so much so that now the life chances of individual blacks have more to do with their economic class position than with their day-to-day encounters with whites. In earlier years the systematic efforts of whites to suppress blacks were obvious to even the most insensitive observer. Blacks were denied access to valued and scarce resources through various ingenious schemes of racial exploitation, discrimination, and segregation, schemes that were reinforced by elaborate ideologies of racism. But the situation has changed. However determinative such practices were for the previous efforts of the black population to achieve racial equality, and however significant they were in the creation of poverty-stricken ghettos and a vast underclass of black proletarians—that massive population at the very bottom of the social class ladder plagued by poor education and low-paying, unstable jobs—they do not provide a meaningful explanation of the life chances of black Americans today. The traditional patterns of interaction between blacks and whites, particularly in the labor market, have been fundamentally altered.

In the antebellum period, and in the latter half of the nineteenth century through the first half of the twentieth century, the continuous and explicit efforts of whites to construct racial barriers profoundly affected the lives of black Americans. Racial oppression was deliberate, overt, and is easily documented, ranging from slavery to segregation, from the endeavors of the white economic elite to exploit black labor to the actions of the white masses to eliminate or neutralize black competition, particularly economic competition.¹ As the nation has entered the latter half of the twentieth century, however, many of the traditional barriers have crumbled under the weight of the political, social, and economic changes of the civil rights era. A new set of obstacles has emerged from basic structural shifts in the economy. These obstacles are therefore impersonal but may prove to be even more formidable for certain segments of the black population. Specifically, whereas the previous barriers were usually designed to control and restrict the entire black population, the new barriers create hardships essentially for the black underclass; whereas the old barriers were based explicitly on racial

William Julius Wilson. *The Declining Significance of Race: Blacks and Changing American Institutions*, pp. 1–9, 12–23, 183–185. Copyright © 1978 by the University of Chicago Press. Used by permission of the University of Chicago Press and the author.

motivations derived from intergroup contact, the new barriers have racial significance only in their consequences, not in their origins. In short, whereas the old barriers bore the pervasive features of racial oppression, the new barriers indicate an important and emerging form of class subordination.

It would be shortsighted to view the traditional forms of racial segregation and discrimination as having essentially disappeared in contemporary America; the presence of blacks is still firmly resisted in various institutions and social arrangements, for example, residential areas and private social clubs. However, in the economic sphere, class has become more important than race in determining black access to privilege and power. It is clearly evident in this connection that many talented and educated blacks are now entering positions of prestige and influence at a rate comparable to or, in some situations, exceeding that of whites with equivalent qualifications. It is equally clear that the black underclass is in a hopeless state of economic stagnation, falling further and further behind the rest of society. . . .

Three Stages of American Race Relations

My basic thesis is that American society has experienced three major stages of black-white contact and that each stage embodies a different form of racial stratification structured by the particular arrangement of both the economy and the polity. Stage one coincides with antebellum slavery and the early post-bellum era and may be designated the period of *plantation economy and racial-caste oppression*. Stage two begins in the last quarter of the nineteenth century and ends at roughly the New Deal era and may be identified as the period of *industrial expansion, class conflict, and racial oppression*. Finally, stage three is associated with the modern, industrial, post-World War II era, which really began to crystallize during the 1960s and 1970s, and may be characterized as the period of *progres-*

sive transition from racial inequalities to class inequalities. For the sake of brevity I shall identify the different periods respectively as the preindustrial, industrial, and modern industrial stages of American race relations.

Although this abbreviated designation of the periods of American race relations seems to relate racial change to fundamental economic changes rather directly, it bears repeating that the different stages of race relations are structured by the unique arrangements and interactions of the economy and the polity. Although I stress the economic basis of structured racial inequality in the preindustrial and industrial periods of race relations, I also attempt to show how the polity more or less interacted with the economy either to reinforce patterns of racial stratification or to mediate various forms of racial conflict. Moreover, for the modern industrial period, I try to show how race relations have been shaped as much by important economic changes as by important political changes. Indeed, it would not be possible to understand fully the subtle and manifest changes in race relations in the modern industrial period without recognizing the dual and often reciprocal influence of structural changes in the economy and political changes in the state. Thus, my central argument is that different systems of production and/or different arrangements of the polity have imposed different constraints on the way in which racial groups have interacted in the United States, constraints that have structured the relations between racial groups and that have produced dissimilar contexts not only for the manifestation of racial antagonisms but also for racial group access to rewards and privileges.

In contrast to the modern industrial period in which fundamental economic and political changes have made economic class affiliation more important than race in determining Negro prospects for occupational advancement, the preindustrial and industrial periods of black-white relations have one central feature in common, namely, overt efforts of

whites to solidify economic racial domination (ranging from the manipulation of black labor to the neutralization or elimination of black economic competition) through various forms of juridical, political, and social discrimination. Since racial problems during these two periods were principally related to group struggles over economic resources, they readily lend themselves to the economic class theories of racial antagonisms that associate racial antipathy with class conflict. A brief consideration of these theories, followed by a discussion of their basic weaknesses, will help to raise a number of theoretical issues that will be useful for analyzing the dynamics of racial conflict in the preindustrial and industrial stages of American race relations. However, in a later section of this chapter I shall attempt to explain why these theories are not very relevant to the modern industrial stage of American race relations.

Economic Class Theories

Students of race relations have paid considerable attention to the economic basis of racial antagonism in recent years, particularly to the theme that racial problems in historical situations are related to the more general problems of economic class conflict. A common assumption of this theme is that racial conflict is merely a special manifestation of class conflict. Accordingly, ideologies of racism, racial prejudices, institutionalized discrimination, segregation, and other factors that reinforce or embody racial stratification are seen as simply part of a superstructure determined and shaped by the particular arrangement of the class structure.² However, given this basic assumption, which continues to be the most representative and widely used economic class argument,³ proponents have advanced two major and somewhat divergent explanations of how class conflicts actually shape and determine racial relations—the orthodox Marxist theory of capitalist exploitation,⁴ and the *split labor-market theory* of working-class antagonisms.⁵

The orthodox Marxist theory, which is the most popular variant of the Marxists' explanations of race,⁶ postulates that because the ultimate goal of the capitalist class is to maximize profits, efforts will be made to suppress workers' demands for increased wages and to weaken their bargaining power by promoting divisions within their ranks. The divisions occur along racial lines to the extent that the capitalist class is able to isolate the lower-priced black labor force by not only supporting job, housing, and educational discrimination against blacks, but also by developing or encouraging racial prejudices and ideologies of racial subjugation such as racism. The net effect of such a policy is to insure a marginal working class of blacks and to establish a relatively more privileged position for the established white labor force. Since discrimination guarantees a situation where the average wage rate of the black labor force is less than the average wage rate of the established white labor force, the probability of labor solidarity against the capitalist class is diminished.

At the same time, orthodox Marxists argue, the members of the capitalist class benefit not only because they have created a reserved army of labor that is not united against them and the appropriation of surplus from the black labor force is greater than the exploitation rate of the white labor force,⁷ but also because they can counteract ambitious claims of the white labor force for higher wages either by threatening to increase the average wage rate of black workers or by replacing segments of the white labor force with segments of the black labor force in special situations such as strikes. The weaker the national labor force, the more likely it is that it will be replaced by lower-paid black labor especially during organized strikes demanding wage increases and improved working conditions. In short, orthodox Marxists argue that racial antagonism is designed to be a "mask for privilege" that effectively conceals the efforts of the ruling class to exploit subordinate minority groups and divide the working class.

In interesting contrast to the orthodox Marxist approach, the split labor-market theory posits the view that rather than attempting to protect a segment of the laboring class, business “supports a liberal or *laissez faire* ideology that would permit all workers to compete freely in an open market. Such open competition would displace higher paid labor. Only under duress does business yield to a labor aristocracy [i.e., a privileged position for white workers].”⁸

The central hypothesis of the split labor-market theory is that racial antagonism first develops in a labor market split along racial lines. The term “antagonism” includes all aspects of intergroup conflict, from beliefs and ideologies (e.g., racism), to overt behavior (e.g., discrimination), to institutions (e.g., segregationist laws). A split labor market occurs when the price of labor for the same work differs for at least two groups, or would differ if they performed the same work. The price of labor “refers to labor’s total cost to the employer, including not only wages, but the cost of recruitment, transportation, room and board, education, health care (if the employer must bear these), and the cost of labor unrest.”⁹

There are three distinct classes in a split labor market: (1) business or employers; (2) higher-paid labor; and (3) cheaper labor. Conflict develops between these three classes because of different interests. The main goal of business or employers is to maintain as cheap a labor force as possible in order to compete effectively with other businesses and to maximize economic returns. Employers will often import laborers from other areas if local labor costs are too high or if there is a labor shortage. Whenever a labor shortage exists, higher-paid labor is in a good bargaining position. Accordingly, if business is able to attract cheaper labor to the market place, the interests of higher-paid labor are threatened. They may lose some of the privileges they enjoy, they may lose their bargaining power, and they may even lose their jobs. Moreover, the presence of cheaper labor in a particular

job market may not only represent actual competition but potential competition as well. An “insignificant trickle” could be seen as the beginning of a major immigration. If the labor market is split along ethnic lines, for example, if higher-paid labor is white and lower-paid labor is black, class antagonisms are transformed into racial antagonisms. Thus, “while much rhetoric of ethnic antagonism concentrates on ethnicity and race, it really in large measure (though probably not entirely) expresses this class conflict.”¹⁰

In some cases members of the lower-paid laboring class, either from within the territorial boundaries of a given country or from another country, are drawn into or motivated to enter a labor market because they feel they can improve their standard of living. As Edna Bonacich points out, “the poorer the economy of the recruits, the less the inducement needed for them to enter the new labor market.”¹¹ In other cases, individuals are forced into a new labor-market situation, such as the involuntary migration of blacks into a condition of slavery in the United States. In this connection, the greater the employer’s control over lower-priced labor, the more threatening is lower-paid labor to higher-paid labor.

However, if more expensive labor is strong enough, that is, if it possesses the power resources to preserve its economic interests, it can prevent being replaced or undercut by cheaper labor. On the one hand it can exclude lower-paid labor from a given territory. “Exclusion movements clearly serve the interests of higher paid labor. Its standards are protected, while the capitalist class is deprived of cheaper labor.”¹² On the other hand, if it is not possible for higher-paid labor to rely on exclusion (cheaper labor may be indigenous to the territory or may have been imported early in business-labor relations when higher-paid labor could not prevent the move) then it will institutionalize a system of ethnic stratification which could (1) monopolize skilled positions, thereby ensuring the effectiveness of strike action; (2) prevent cheaper labor from

developing the skills necessary to compete with higher-paid labor (for example, by imposing barriers to equal access to education); and (3) deny cheaper labor the political resources that would enable them to undercut higher-paid labor through, say, governmental regulations. “In other words, the solution to the devastating potential of weak, cheap labor is, paradoxically, to weaken them further, until it is no longer in business’ immediate interest to use them as replacement.”¹³ Thus, whereas orthodox Marxist arguments associate the development and institutionalization of racial stratification with the motivations and activities of the capitalist class, the split labor-market theory traces racial stratification directly to the powerful, higher-paid working class.

Implicit in both of these economic class theories is a power-conflict thesis associating the regulation of labor or wages with properties (ownership of land or capital, monopolization of skilled positions) that determine the scope and degree of a group’s ability to influence behavior in the labor market. Furthermore, both theories clearly demonstrate the need to focus on the different ways and situations in which various segments of the dominant racial group perceive and respond to the subordinate racial group. However, as I examine the historical stages of race relations in the United States, I find that the patterns of black/white interaction do not consistently and sometimes do not conveniently conform to the propositions outlined in these explanations of racial antagonism. In some cases, the orthodox Marxian explanation seems more appropriate; in other instances, the split labor-market theory seems more appropriate; and in still others, neither theory can, in isolation, adequately explain black-white conflict.

If we restrict our attention for the moment to the struggle over economic resources, then the general pattern that seems to have characterized race relations in the United States during the preindustrial and industrial stages was that the economic elite segments of the white population have been principally re-

sponsible for those forms of racial inequality that entail the exploitation of labor (as in slavery), whereas whites in the lower strata have been largely responsible for those forms of imposed racial stratification that are designed to eliminate economic competition (as in job segregation). Moreover, in some situations, the capitalist class and white workers form an alliance to keep blacks suppressed. Accordingly, restrictive arguments to the effect that racial stratification was the work of the capitalist class or was due to the “victory” of higher-paid white labor obscure the dynamics of complex and variable patterns of black-white interaction.

However, if we ignore the more categorical assertions that attribute responsibility for racial stratification to a particular class and focus seriously on the analyses of interracial contact in the labor market, then I will be able to demonstrate that, depending on the historical situation, each of the economic class theories provides arguments that help to illuminate race relations during the preindustrial and industrial periods of black-white contact. By the same token, I hope to explain why these theories have little application to the third, and present, stage of modern industrial race relations. My basic argument is that the meaningful application of the arguments in each theory for any given historical period depends considerably on knowledge of the constraints imposed by the particular systems of production and by the particular laws and policies of the state during that period, constraints that shape the structural relations between racial and class groups and which thereby produce different patterns of intergroup interaction. . . .

The Influence of the System of Production

The term “system of production” not only refers to the technological basis of economic processes or, in Karl Marx’s terms, the “forces of production,” but it also implies the “social

relations of production,” that is, “the interaction (for example, through employment and property arrangement) into which men enter at a given level of the development of the forces of production.”¹⁴ As I previously indicated, different systems of production impose constraints on racial group interaction. In the remainder of this section I should like to provide a firmer analytical basis for this distinction as it applies specifically to the three stages of American race relations, incorporating in my discussion relevant theoretical points raised in the foregoing sections of this chapter.

It has repeatedly been the case that a non-manufacturing or plantation economy with a simple division of labor and a small aristocracy that dominates the economic and political life of a society has characteristically generated a paternalistic rather than a competitive form of race relations, and the antebellum South was no exception.¹⁵ Paternalistic racial patterns reveal close symbiotic relationships marked by dominance and subservience, great social distance and little physical distance, and clearly symbolized rituals of racial etiquette. The southern white aristocracy created a split labor market along racial lines by enslaving blacks to perform tasks at a cheaper cost than free laborers of the dominant group. This preindustrial form of race relations was not based on the actions of dominant-group laborers, who, as we shall see, were relatively powerless to effect significant change in race relations during this period, but on the structure of the relations established by the aristocracy. Let me briefly amplify this point.

In the southern plantation economy, public power was overwhelmingly concentrated in the hands of the white aristocracy. This power was not only reflected in the control of economic resources and in the development of a juridical system that expressed the class interests of the aristocracy, but also in the way the aristocracy was able to impose its viewpoint on the larger society.¹⁶ This is not to suggest that these aspects of public power

have not been disproportionately controlled by the economic elite in modern industrialized Western societies; rather it indicates that the hegemony of the southern ruling elite was much greater in degree, not in kind, than in these societies. The southern elite’s hegemony was embodied in an economy that required little horizontal or vertical mobility. Further, because of the absence of those gradations of labor power associated with complex divisions of labor, white workers in the antebellum and early postbellum South had little opportunity to challenge the control of the aristocracy. Because white laborers lacked power resources in the southern plantation economy, their influence on the form and quality of racial stratification was minimal throughout the antebellum and early postbellum periods. Racial stratification therefore primarily reflected the relationships established between blacks and the white aristocracy, relationships which were not characterized by competition for scarce resources but by the exploitation of black labor.¹⁷ Social distance tended to be clearly symbolized by rituals of racial etiquette: gestures and behavior reflecting dominance and subservience. Consequently, any effort to impose a system of public segregation was superfluous. Furthermore, since the social gap between the aristocracy and black slaves was wide and stable, ideologies of racism played less of a role in the subordination of blacks than they subsequently did in the more competitive systems of race relations following the Civil War. In short, the relationship represented intergroup paternalism because it allowed for “close symbiosis and even intimacy, without any threat to status inequalities.”¹⁸ This was in sharp contrast to the more competitive forms of race relations that accompanied the development of industrial capitalism in the late nineteenth century and first few decades of the twentieth century (the industrial period of American race relations), wherein the complex division of labor and opportunities for greater mobility not only produced inter-

action, competition, and labor-market conflict between blacks and the white working class, but also provided the latter with superior resources (relative to those they possessed under the plantation economy) to exert greater influence on the form and content of racial stratification.

The importance of the system of production in understanding race relations is seen in a comparison of Brazil and the southern United States during the post-slavery periods. In the United States, the southern economy experienced a fairly rapid rate of expansion during the late nineteenth century, thereby creating various middle level skilled and unskilled positions that working-class whites attempted to monopolize for themselves. The efforts of white workers to eliminate black competition in the South generated an elaborate system of Jim Crow segregation that was reinforced by an ideology of biological racism. The white working class was aided not only by its numerical size, but also by its increasing accumulation of political resources that accompanied changes in its relation to the means of production.

As white workers gradually translated their increasing labor power into political power, blacks experienced greater restrictions in their efforts to achieve a satisfactory economic, political, and social life. In Brazil, on the other hand, the large Negro and mulatto population was not thrust into competition with the much smaller white population over access to higher-status positions because, as Marvin Harris notes, "there was little opportunity for any member of the lower class to move upward in the social hierarchy."¹⁹ No economic class group or racial group had much to gain by instituting a rigid system of racial segregation or cultivating an ideology of racial inferiority. Racial distinctions were insignificant to the landed aristocracy, who constituted a numerically small upper class in what was basically a sharply differentiated two-class society originally shaped during slavery. The mulattoes, Negroes, and poor

whites were all in the same impoverished lower-ranking position. "The general economic stagnation which has been characteristic of lowland Latin America since the abolition of slavery," observes Marvin Harris, "tends to reinforce the pattern of pacific relationships among the various racial groups in the lower ranking levels of the social hierarchy. Not only were the poor whites outnumbered by the mulattoes and Negroes, but there was little of a significant material nature to struggle over in view of the generally static condition of the economy."²⁰ Accordingly, in Brazil, segregation, discrimination, and racist ideologies failed to crystallize in the first several decades following the end of slavery. More recently, however, industrialization has pushed Brazil toward a competitive type of race relations, particularly the southern region (for example, São Paulo) which has experienced rapid industrialization and has blacks in economic competition with many lower-status white immigrants.²¹

Whereas the racial antagonism in the United States during the period of industrial race relations (such as the Jim Crow segregation movement in the South and the race riots in northern cities) tended to be either directly or indirectly related to labor-market conflicts, racial antagonism in the period of modern industrial relations tends to originate outside the economic order and to have little connection with labor-market strife. Basic changes in the system of production have produced a segmented labor structure in which blacks are either isolated in the relatively non-unionized, low-paying, basically undesirable jobs of the noncorporate sector, or occupy the higher-paying corporate and government industry positions in which job competition is either controlled by powerful unions or is restricted to the highly trained and educated, regardless of race. If there is a basis for labor-market conflict in the modern industrial period, it is most probably related to the affirmative action programs originating from the civil rights legislation of the

1960s. However, since affirmative action programs are designed to improve job opportunities for the talented and educated, their major impact has been in the higher-paying jobs of the expanding government sector and the corporate sector. The sharp increase of the more privileged blacks in these industries has been facilitated by the combination of affirmative action and rapid industry growth. Indeed despite the effectiveness of affirmative action programs the very expansion of these sectors of the economy has kept racial friction over higher-paying corporate and government jobs to a minimum.

Unlike the occupational success achieved by the more talented and educated blacks, those in the black underclass find themselves locked in the low-paying and dead-end jobs of the noncorporate industries, jobs which are not in high demand and which therefore do not generate racial competition or strife among the national black and white labor force. Many of these jobs go unfilled, and employers often have to turn to cheap labor from Mexico and Puerto Rico. As Nathan Glazer has pointed out, "Expectations have changed, and fewer blacks and whites today will accept a life at menial labor with no hope for advancement, as their fathers and older brothers did and as European immigrants did."²²

Thus in the modern industrial era neither the corporate nor government sectors nor the noncorporate low-wage sector provide the basis for the kind of interracial competition and conflict that has traditionally plagued the labor market in the United States. This, then, is the basis for my earlier contention that the economic class theories which associate labor-market conflicts with racial antagonism have little application to the present period of modern industrial race relations.

The Polity and American Race Relations

If the patterned ways in which racial groups have interacted historically have been shaped

in major measure by different systems of production, they have also been undeniably influenced by the changing policies and laws of the state. For analytical purposes, it would be a mistake to treat the influences of the polity and the economy as if they were separate and unrelated. The legal and political systems in the antebellum South were effectively used as instruments of the slaveholding elite to strengthen and legitimate the institution of slavery. But as industrialization altered the economic class structure in the postbellum South, the organizing power and political consciousness of the white lower class increased and its members were able to gain enough control of the political and juridical systems to legalize a new system of racial domination (Jim Crow segregation) that clearly reflected their class interests.

In effect, throughout the preindustrial period of race relations and the greater portion of the industrial period the role of the polity was to legitimate, reinforce, and regulate patterns of racial inequality. However, it would be unwarranted to assume that the relationship between the economic and political aspects of race necessarily implies that the latter is simply a derivative phenomenon based on the more fundamental processes of the former. The increasing intervention, since the mid-twentieth century, of state and federal government agencies in resolving or mediating racial conflicts has convincingly demonstrated the political system's autonomy in handling contemporary racial problems. Instead of merely formalizing existing racial alignments as in previous periods, the political system has, since the initial state and municipal legislation of the 1940s, increasingly created changes leading to the erosion of traditional racial alignments; in other words, instead of reinforcing racial barriers created during the preindustrial and industrial periods, the political system in recent years has tended to promote racial equality.

Thus, in the previous periods the polity was quite clearly an instrument of the white

population in suppressing blacks. The government's racial practices varied, as I indicated above, depending on which segment of the white population was able to assert its class interests. However, in the past two decades the interests of the black population have been significantly reflected in the racial policies of the government, and this change is one of the clearest indications that the racial balance of power has been significantly altered. Since the early 1940s the black population has steadily gained political resources and, with the help of sympathetic white allies, has shown an increasing tendency to utilize these resources in promoting or protecting its group interests.

By the mid-twentieth century the black vote had proved to be a major vehicle for political pressure. The black vote not only influenced the outcome of national elections but many congressional, state, and municipal elections as well. Fear of the Negro vote produced enactment of public accommodation and fair employment practices laws in northern and western municipalities and states prior to the passage of federal civil rights legislation in 1964. This political resurgence for black Americans increased their sense of power, raised their expectations, and provided the foundation for the proliferation of demands which shaped the black revolt during the 1960s. But there were other factors that helped to buttress Negro demands and contributed to the developing sense of power and rising expectations, namely, a growing, politically active black middle class following World War II and the emergence of the newly independent African states.

The growth of the black middle class was concurrent with the growth of the black urban population. It was in the urban areas, with their expanding occupational opportunities, that a small but significant number of blacks were able to upgrade their occupations, increase their income, and improve their standard of living. The middle-class segment of an oppressed minority is most

likely to participate in a drive for social justice that is disciplined and sustained. In the early phases of the civil rights movement, the black middle class channeled its energies through organizations such as the National Association for the Advancement of Colored People, which emphasized developing political resources and successful litigation through the courts. These developments were paralleled by the attack against traditional racial alignments in other parts of the world. The emerging newly independent African states led the assault. In America, the so-called "leader of the free world," the manifestation of racial tension and violence has been a constant source of embarrassment to national government officials. This sensitivity to world opinion made the national government more vulnerable to pressures of black protest at the very time when blacks had the greatest propensity to protest.

The development of black political resources that made the government more sensitive to Negro demands, the motivation and morale of the growing black middle class that resulted in the political drive for racial equality, and the emergence of the newly independent African states that increased the federal government's vulnerability to civil rights pressures all combined to create a new sense of power among black Americans and to raise their expectations as they prepared to enter the explosive decade of the 1960s. The national government was also aware of this developing sense of power and responded to the pressures of black protest in the 1960s with an unprecedented series of legislative enactments to protect black civil rights.

The problem for blacks today, in terms of government practices, is no longer one of legalized racial inequality. Rather the problem for blacks, especially the black underclass, is that the government is not organized to deal with the new barriers imposed by structural changes in the economy. With the passage of equal employment legislation and the authorization of affirmative action programs the

government has helped clear the path for more privileged blacks, who have the requisite education and training, to enter the mainstream of American occupations. However, such government programs do not confront the impersonal economic barriers confronting members of the black underclass, who have been effectively screened out of the corporate and government industries. And the very attempts of the government to eliminate traditional racial barriers through such programs as affirmative action have had the unintentional effect of contributing to the growing economic class divisions within the black community.

Class Stratification and Changing Black Experiences

The problems of black Americans have always been compounded because of their low position in both the economic order (the average economic class position of blacks as a group) and the social order (the social prestige or honor accorded individual blacks because of their ascribed racial status). It is of course true that the low economic position of blacks has helped to shape the categorical social definitions attached to blacks as a racial group, but it is also true that the more blacks become segmented in terms of economic class position, the more their concerns about the social significance of race will vary.

In the preindustrial period of American race relations there was of course very little variation in the economic class position of blacks. The system of racial-caste oppression relegated virtually all blacks to the bottom of the economic class hierarchy. Moreover, the social definitions of racial differences were heavily influenced by the ideology of racism and the doctrine of paternalism, both of which clearly assigned a subordinate status for blacks vis-à-vis whites. Occasionally, a few individual free blacks would emerge and accumulate some wealth or property, but they were the overwhelming exception. Thus

the uniformly low economic class position of blacks reinforced and, in the eyes of most whites, substantiated the social definitions that asserted Negroes were culturally and biogenetically inferior to whites. The uniformly low economic class position of blacks also removed the basis for any meaningful distinction between race issues and class issues within the black community.

The development of a black middle class accompanied the change from a preindustrial to an industrial system of production. Still, despite the fact that some blacks were able to upgrade their occupation and increase their education and income, there were severe limits on the areas in which blacks could in fact advance. Throughout most of the industrial period of race relations, the growth of the black middle class occurred because of the expansion of institutions created to serve the needs of a growing urbanized black population. The black doctor, lawyer, teacher, minister, businessman, mortician, excluded from the white community, was able to create a niche in the segregated black community. Although the income levels and life-styles of the black professionals were noticeably and sometimes conspicuously different from those of the black masses, the two groups had one basic thing in common, a racial status contemptuously regarded by most whites in society. If E. Franklin Frazier's analysis of the black bourgeoisie is correct, the black professionals throughout the industrial period of American race relations tended to react to their low position in the social order by an ostentatious display of material possessions and a conspicuous effort to disassociate themselves from the black masses.²³

Still, as long as the members of the black middle class were stigmatized by their racial status; as long as they were denied the social recognition accorded their white counterparts; more concretely, as long as they remained restricted in where they could live, work, socialize, and be educated, race would continue to be a far more salient and impor-

tant issue in shaping their sense of group position than their economic class position. Indeed, it was the black middle class that provided the leadership and generated the momentum for the civil rights movement during the mid-twentieth century. The influence and interests of this class were clearly reflected in the way the race issues were defined and articulated. Thus, the concept of “freedom” quite clearly implied, in the early stages of the movement, the right to swim in certain swimming pools, to eat in certain restaurants, to attend certain movie theaters, and to have the same voting privileges as whites. These basic concerns were reflected in the 1964 Civil Rights Bill which helped to create the illusion that, when the needs of the black middle class were met, so were the needs of the entire black community.

However, although the civil rights movement initially failed to address the basic needs of the members of the black lower class, it did increase their awareness of racial oppression, heighten their expectations about improving race relations, and increase their impatience with existing racial arrangements. These feelings were dramatically manifested in a series of violent ghetto outbreaks that rocked the nation throughout the late 1960s. These outbreaks constituted the most massive and sustained expression of lower-class black dissatisfaction in the nation’s history. They also forced the political system to recognize the problems of human survival and de facto segregation in the nation’s ghettos—problems pertaining to unemployment and underemployment, inferior ghetto schools, and deteriorated housing.

However, in the period of modern industrial race relations, it would be difficult indeed to comprehend the plight of inner-city blacks by exclusively focusing on racial discrimination. For in a very real sense, the current problems of lower-class blacks are substantially related to fundamental structural changes in the economy. A history of discrimination and oppression created a huge

black underclass, and the technological and economic revolutions have combined to insure it a permanent status.

As the black middle class rides on the wave of political and social changes, benefiting from the growth of employment opportunities in the growing corporate and government sectors of the economy, the black underclass falls behind the larger society in every conceivable respect. The economic and political systems in the United States have demonstrated remarkable flexibility in allowing talented blacks to fill positions of prestige and influence at the same time that these systems have shown persistent rigidity in handling the problems of lower-class blacks. As a result, for the first time in American history class issues can meaningfully compete with race issues in the way blacks develop or maintain a sense of group position.²⁴

Conclusion

The foregoing sections of this chapter present an outline and a general analytical basis for the arguments that will be systematically explored [elsewhere].²⁵ I have tried to show that race relations in American society have been historically characterized by three major stages and that each stage is represented by a unique form of racial interaction which is shaped by the particular arrangement of the economy and the polity. My central argument is that different systems of production and/or different policies of the state have imposed different constraints on the way in which racial groups interact—constraints that have structured the relations between racial groups and produced dissimilar contexts not only for the manifestation of racial antagonisms but also for racial-group access to rewards and privileges. I emphasized in this connection that in the preindustrial and industrial periods of American race relations the systems of production primarily shaped the patterns of racial stratification and the role of the polity was to legitimate, reinforce,

or regulate these patterns. In the modern industrial period, however, both the system of production and the polity assume major importance in creating new patterns of race relations and in altering the context of racial strife. Whereas the preindustrial and industrial stages were principally related to group struggles over economic resources as different segments of the white population overtly sought to create and solidify economic racial domination (ranging from the exploitation of black labor in the preindustrial period to the elimination of black competition for jobs in the industrial period) through various forms of political, juridical, and social discrimination; in the modern industrial period fundamental economic and political changes have made economic class position more important than race in determining black chances for occupational mobility. Finally, I have outlined the importance of racial norms or belief systems, especially as they relate to the general problem of race and class conflict in the preindustrial and industrial periods.

My argument that race relations in America have moved from economic racial oppression to a form of class subordination for the less-privileged blacks is not meant to suggest that racial conflicts have disappeared or have even been substantially reduced. On the contrary, the basis of such conflicts have shifted from the economic sector to the sociopolitical order and therefore do not play as great a role in determining the life chances of individual black Americans as in the previous periods of overt economic racial oppression.

NOTES

1. See William J. Wilson, *Power, Racism and Privilege: Race Relations in Theoretical and Sociobistorical Perspectives* (New York: The Free Press, 1973).

2. In Marxist terminology, the “superstructure” refers to the arrangements of beliefs, norms, ideologies, and noneconomic institutions.

3. However, not all theorists who emphasize the importance of economic class in explanations of race relations simply relegate problems of race to the super-

structure. The Marxist scholars Michael Burawoy and Eugene Genovese recognize the reciprocal influence between the economic class structure and aspects of the superstructure (belief systems, political systems, etc.), a position which I also share and which is developed more fully in subsequent sections of this chapter. See Eugene D. Genovese, *Roll, Jordan, Roll: The World the Slaves Made* (New York: Pantheon, 1974); idem, *In Red and Black: Marxian Explorations in Southern and Afro-American History* (New York: Vintage Press, 1971); and Michael Burawoy, “Race, Class, and Colonialism,” *Social and Economic Studies* 23 (1974): 521–50.

4. Oliver C. Cox, *Caste, Class and Race: A Study in Social Dynamics* (Garden City, New York: Doubleday, 1948); Paul A. Baran and Paul M. Sweezy, *Monopoly Capital: An Essay on the American Economic and Social Order* (Harmondsworth: Penguin, 1966); Michael Reich, “The Economics of Racism,” in *Problems in Political Economy*, ed. David M. Gordon (Lexington, Mass.: Heath, 1971); and M. Nikolinos, “Notes on an Economic Theory of Racism,” *Race: A Journal of Race and Group Relations* 14 (1973): 365–81.

5. Edna Bonacich, “A Theory of Ethnic Antagonism: The Split Labor Market,” *American Sociological Review* 37 (October 1972): 547–59; idem, “Abolition, the Extension of Slavery and the Position of Free Blacks: A Study of Split Labor Markets in the United States,” *American Journal of Sociology* 81 (1975): 601–28.

6. For examples of alternative and less orthodox Marxist explanations of race, see Eugene D. Genovese, *The Political Economy of Slavery: Studies in the Economy and Society of the Slave South* (New York: Pantheon, 1966); idem, *The World the Slaveholders Made: Two Essays in Interpretation* (New York: Pantheon, 1969); idem, *In Red and Black*; idem, *Roll, Jordan, Roll*; and Burawoy, “Race, Class, and Colonialism.”

7. “Exploitation,” in Marxist terminology, refers to the difference between the wages workers receive and the value of the goods they produce. The size of this difference, therefore, determines the degree of exploitation.

8. Bonacich, “A Theory of Ethnic Antagonism,” p. 557.

9. *Ibid.*, p. 549.

10. *Ibid.*, p. 553.

11. *Ibid.*, p. 549.

12. *Ibid.*, p. 555.

13. *Ibid.*, p. 556.

14. Neil J. Smelser, *Karl Marx on Society and Social Change* (Chicago: University of Chicago Press, 1974), p. xiv. According to Smelser, Marx used the notions “forces of production” and “social relations of production” as constituting the “mode of produc-

tion.” However, in Marx’s writings the mode of production is often discussed as equivalent only to the “forces of production.” To avoid confusion, I have chosen the term “system of production” which denotes the interrelation of the forces of production and the mode of production.

15. Pierre L. van den Berghe, *Race and Racism: A Comparative Perspective* (New York: John Wiley and Sons, 1967), p. 26.

16. See, for example, Genovese, *Roll, Jordan, Roll*.

17. An exception to this pattern occurred in the cities of the antebellum South, where non-slaveholding whites played a major role in the development of urban segregation. However, since an overwhelming majority of the population resided in rural areas, race relations in the antebellum southern cities were hardly representative of the region.

18. van den Berghe, *Race and Racism*, p. 27.

19. Marvin Harris, *Patterns of Race in the Americas* (New York: Walker, 1964), p. 96.

20. *Ibid.*, p. 96.

21. van den Berghe, *Race and Racism*, p. 28.

22. Nathan Glazer, “Blacks and Ethnic Groups: The Difference, and the Political Difference It Makes,” in *Key Issues in the Afro-American Experience*, ed. Nathan I. Huggins, Martin Kilson, and Daniel M. Fox (New York: Harcourt Brace Jovanovich, 1971), 2: 209.

23. E. Franklin Frazier, *Black Bourgeoisie* (New York: The Free Press, 1957). See also Nathan Hare, *Black Anglo-Saxons* (New York: Collier, 1965).

24. The theoretical implications of this development for ethnic groups in general are discussed by Milton Gordon under the concept “ethclass.” See Milton M. Gordon, *Assimilation in American Life* (New York: Oxford University Press, 1964).

25. See William Julius Wilson, *The Declining Significance of Race: Blacks and Changing American Institutions* (Chicago and London: University of Chicago Press, 1978).

74 ■ Joe R. Feagin

The Continuing Significance of Race

Antiblack Discrimination in Public Places

Racial discrimination as a continuing and major problem for middle-class blacks has been downplayed as analysts have turned to the various problems of the “underclass.” For example, Wilson (1978, pp. 110–1) has argued that the growth of the black middle class since the 1960s is the result of improving economic conditions and of government civil rights laws, which virtually eliminated overt discrimination in the workplace and public accommodations. According to Wilson, the major problem of the 1964 Civil Rights Act is its failure to meet the problems of the black underclass (Wilson 1987, pp. 146–7).

Here I treat these assertions as problematic. Do middle-class black Americans still face hostile treatment in public accommodations and other public places? If so, what form does this discrimination take? Who are the perpetrators of this discrimination? What is the impact of the discrimination on its middle-class victims? How do middle-class blacks cope with such discrimination?

Discrimination can be defined in social-contextual terms as “actions or practices carried out by members of dominant racial or ethnic groups that have a differential and negative impact on members of subordinate racial and ethnic groups” (Feagin and

Eckberg 1980, pp. 1–2). This differential treatment ranges from the blatant to the subtle (Feagin and Feagin 1986). Here I focus primarily on blatant discrimination by white Americans targeting middle-class blacks. . . .

To examine discrimination, I draw primarily on 37 in-depth interviews from a larger study of 135 middle-class black Americans in Boston, Buffalo, Baltimore, Washington, D.C., Detroit, Houston, Dallas, Austin, San Antonio, Marshall, Las Vegas, and Los Angeles. The interviewing was done in 1988–1990; black interviewers were used. . . .

Responses to Discrimination: Public Accommodations

Two Fundamental Strategies: Verbal Confrontation and Withdrawal

In the following account, a black news director at a major television station shows the interwoven character of discriminatory action and black response. The discrimination took the form of poor restaurant service, and the responses included both suggested withdrawal and verbal counterattack.

He [her boyfriend] was waiting to be seated. . . . He said, “You go to the bathroom and I’ll get the table. . . .” He was standing there when I came back; he continued to stand there. The restaurant was almost empty. There were waiters, waitresses, and no one seated. And when I got back to him, he was ready to leave, and said, “Let’s go.” I said, “What happened to our table?” He wasn’t seated. So I said, “No, we’re not leaving, please.” And he said, “No, I’m leaving.” So we went outside, and we talked about it. And what I said to him was, you have to be aware of the possibilities that this is not the first time that this has happened at this restaurant or at other restaurants, but this is the first time it has happened to a black news director here or someone who could make an issue of it, or someone who is prepared to make an issue of it.

So we went back inside after I talked him into it and, to make a long story short, I had the manager come. I made most of the people who were there (while conducting myself professionally the whole time) aware that I was incensed at being treated this way. . . . I said, “Why do you think we weren’t seated?” And the manager said, “Well, I don’t really know.” And I said, “Guess.” He said, “Well I don’t know, because you’re black?” I said, “Bingo. Now isn’t it funny that you didn’t guess that I didn’t have any money” (and I opened up my purse) and I said, “because I certainly have money. And isn’t it odd that you didn’t guess that it’s because I couldn’t pay for it because I’ve got two American Express cards and a Master Card right here. I think it’s just funny that you would have assumed that it’s because I’m black.”. . . And then I took out my card and gave it to him and said, “If this happens again, or if I hear of this happening again, I will bring the full wrath of an entire news department down on this restaurant.” And he just kind of looked at me. “Not [just] because I am personally offended. I am. But because you have no right to do what you did, and as a people we have lived a long time with having our rights abridged. . . .” There were probably three or four sets of diners in the restaurant and maybe five waiters/waitresses. They watched him standing there waiting to be seated. His reaction to it was that he wanted to leave. I understood why he would have reacted that way, because he felt that he was in no condition to be civil. He was ready to take the place apart and . . . sometimes it’s appropriate to behave that way. We hadn’t gone the first step before going on to the next step. He didn’t feel that he could comfortably and calmly take the first step, and I did. So I just asked him to please get back in the restaurant with me, and then you don’t have to say a word, and let me handle it from there. It took some convincing, but I had to appeal to his sense of this is not just you, this is not just for you. We are finally in a position as black people where there are some of us who

can genuinely get their attention. And if they don't want to do this because it's right for them to do it, then they'd better do it because they're afraid to do otherwise. If it's fear, then fine, instill the fear.

This example provides insight into the character of modern discrimination. The discrimination was not the “No Negroes” exclusion of the recent past, but rejection in the form of poor service by restaurant personnel. The black response indicates the change in black-white interaction since the 1950s and 1960s, for discrimination is handled with vigorous confrontation rather than deference. It is possible that the white personnel defined the couple as “poor blacks” because of their jeans, although the jeans were fashionable and white patrons wear jeans. In comments not quoted here the news director rejects such an explanation. She forcefully articulates a theory of rights—a response that signals the critical impact of civil rights laws on the thinking of middle-class blacks. The news director articulates the American dream: she has worked hard, earned the money and credit cards, developed the appropriate middle-class behavior, and thus has under the law a *right* to be served. There is defensiveness in her actions too, for she feels a need to legitimate her status by showing her purse and credit cards. One important factor that enabled her to take such assertive action was her power to bring a TV news team to the restaurant. This power marks a change from a few decades ago when very few black Americans had the social or economic resources to fight back successfully.

This example underscores the complexity of the interaction in such situations, with two levels of negotiation evident. The negotiation between the respondent and her boyfriend on withdrawal vs. confrontation highlights the process of negotiating responses to discrimination and the difficulty in crafting such responses. Not only is there

a process of dickering with whites within the discriminatory scene but also a negotiation between the blacks involved. . . .

The confrontation response is generally so costly in terms of time and energy that acquiescence or withdrawal are common options. An example of the exit response was provided by a utility company executive in an east coast city:

I can remember one time my husband had picked up our son . . . from camp; and he'd stopped at a little store in the neighborhood near the camp. It was hot, and he was going to buy him a snowball. And the proprietor of the store—this was a very old, white neighborhood, and it was just a little sundry store. But the proprietor said he had the little window where people could come up and order things. Well, my husband and son had gone into the store. And he told them, “Well, I can't give it to you here, but if you go outside to the window, I'll give it to you.” And there were other [white] people in the store who'd been served [inside]. So, they just left and didn't buy anything.

Here the act seems a throwback to the South of the 1950s, where blacks were required to use the back or side of a store. This differential treatment in an older white neighborhood is also suggestive of the territorial character of racial relations in many cities. The black response to degradation here was not to confront the white person or to acquiesce abjectly, but rather to reject the poor service and leave. . . .

Careful Situation Assessments

We have seen in the previous incidents some tendency for blacks to assess discriminatory incidents before they act. . . .

The complex process of evaluation and response is described by a college dean, who commented generally on hotel and restaurant discrimination encountered as he travels across the United States:

When you're in a restaurant and . . . you notice that blacks get seated near the kitchen. You notice that if it's a hotel, your room is near the elevator, or your room is always way down in a corner somewhere. You find that you are getting the undesirable rooms. And you come there early in the day and you don't see very many cars on the lot and they'll tell you that this is all we've got. Or you get the room that's got a bad television set. You know that you're being discriminated against. And of course you have to act accordingly. You have to tell them, "Okay, the room is fine, [but] this television set has got to go. Bring me another television set." So in my personal experience, I simply cannot sit and let them get away with it [discrimination] and not let them know that I know that that's what they are doing. . . .

When I face discrimination, first I take a long look at myself and try to determine whether or not I am seeing what I think I'm seeing in 1989, and if it's something that I have an option [about]. In other words, if I'm at a store making a purchase, I'll simply walk away from it. If it's at a restaurant where I'm not getting good service, I first of all let the people know that I'm not getting good service, then I [may] walk away from it. But the thing that I have to do is to let people know that I know that I'm being singled out for a separate treatment. And then I might react in any number of ways—depending on where I am and how badly I want whatever it is that I'm there for.

This commentary adds another dimension to our understanding of public discrimination, its cumulative aspect. Blacks confront not just isolated incidents—such as a bad room in a luxury hotel once every few years—but a lifelong series of such incidents. Here again the omnipresence of careful assessments is underscored. The dean's interview highlights a major difficulty in being black—one must be constantly prepared to assess accurately and then decide on the ap-

propriate response. This long-look approach may indicate that some middle-class blacks are so sensitive to white charges of hypersensitivity and paranoia that they err in the opposite direction and fail to see discrimination when it occurs. In addition, as one black graduate student at a leading white university in the Southeast put it: "I think that sometimes timely and appropriate responses to racially motivated acts and comments are lost due to the processing of the input." The "long look" can result in missed opportunities to respond to discrimination. . . .

Responses to Discrimination: The Street

Reacting to White Strangers

As we move away from public accommodations settings to the usually less protected street sites, racial hostility can become more fleeting and severer, and thus black responses are often restricted. The most serious form of street discrimination is violence. Often the reasonable black response to street discrimination is withdrawal, resigned acceptance, or a quick verbal retort. The difficulty of responding to violence is seen in this report by a man working for a media surveying firm in a southern industrial city:

I was parked in front of this guy's house. . . . This guy puts his hands on the window and says, "Get out of the car, nigger." . . . So, I got out, and I thought, "Oh, this is what's going to happen here." And I'm talking fast. And they're, "What are you doing here?" And I'm, "This is who I am. I work with these people. This is the man we want to put in the survey." And I pointed to the house. And the guy said, "Well you have an out-of-state license tag, right?" "Yea." And he said, "If something happened to you, your people at home wouldn't know for a long time, would they?" . . . I said, "Look, I deal with a company that deals with television. [If] something happens to me, it's going to be a na-

tional thing. . . . So, they grab me by the lapel of my coat, and put me in front of my car. They put the blade on my zipper. And now I'm thinking about this guy that's in the truck [behind me], because now I'm thinking that I'm going to have to run somewhere. Where am I going to run? Go to the police? [laughs] So, after a while they bash up my headlight. And I drove [away].

Stigmatized and physically attacked solely because of his color, this man faced verbal hostility and threats of death with courage. Cautiously drawing on his middle-class resources, he told the attackers his death would bring television crews to the town. This resource utilization is similar to that of the news director in the restaurant incident. Beyond this verbal threat his response had to be one of caution. For most whites threatened on the street, the police are a sought-after source of protection, but for black men this is often not the case.

At the other end of the street continuum is nonverbal harassment such as the "hate stare." . . . A middle-class student with dark skin reported that on her way to university classes she had stopped at a bakery in a white residential area where very few blacks live or shop. A white couple in front of the store stared intently and hatefully at her as she crossed the sidewalk and entered and left the bakery. She reported that she had experienced this hate stare many times. The incident angered her for some days thereafter, in part because she had been unable to respond more actively to it. . . .

It seems likely that for middle-class blacks the street is the site of recurring encounters with various types of white malevolence. A vivid example of the cumulative character and impact of this discrimination was given by another black student at a white university, who recounted his experiences walking home at night from a campus job to his apartment in a predominantly white residential area:

So, even if you wanted to, it's difficult just to live a life where you don't come into conflict with others. Because every day you walk the streets, it's not even like once a week, once a month. It's every day you walk the streets. Every day that you live as a black person you're reminded how you're perceived in society. You walk the streets at night; white people cross the streets. I've seen white couples and individuals dart in front of cars to not be on the same side of the street. Just the other day, I was walking down the street, and this white female with a child, I saw her pass a young white male about 20 yards ahead. When she saw me, she quickly dragged the child and herself across the busy street. What is so funny is that this area has had an unknown white rapist in the area for about four years. [When I pass] white men tighten their grip on their women. I've seen people turn around and seem like they're going to take blows from me. The police constantly make circles around me as I walk home, you know, for blocks. I'll walk, and they'll turn a block. And they'll come around me just to make sure, to find out where I'm going. So, every day you realize [you're black]. Even though you're not doing anything wrong; you're just existing. You're just a person. But you're a black person perceived in an unblack world. (This quote includes a clarification sentence from a follow-up interview.)

Unable to "see" his middle-class symbols of college dress and books, white couples (as well as individuals) have crossed the street in front of cars to avoid walking near this modest-build black student, in a predominantly white neighborhood. Couples moving into defensive postures are doubtless reacting to the stigma of "black maleness." The student perceives such avoidance as racist, however, not because he is paranoid, but because he has previously encountered numerous examples of whites taking such defensive measures. . . .

Conclusion

I have examined the sites of discrimination, the types of discriminatory acts, and the responses of the victims and have found the color stigma still to be very important in the public lives of affluent black Americans. The sites of racial discrimination range from relatively protected home sites, to less protected workplace and educational sites, to the even less protected public places. The 1964 Civil Rights Act guarantees that black Americans are “entitled to the full and equal enjoyment of the goods, services, facilities, privileges, advantages, and accommodations” in public accommodations. Yet the interviews indicate that deprivation of full enjoyment of public facilities is not a relic of the past; deprivation and discrimination in public accommodations persist. Middle-class black Americans remain vulnerable targets in public places. . . .

Particular instances of discrimination may seem minor to outside white observers when considered in isolation. But when blatant acts of avoidance, verbal harassment, and physical attack combine with subtle and covert slights, and these accumulate over months, years, and lifetimes, the impact on a black person is far more than the sum of the individual instances. . . .

The cumulative impact of racial discrimination accounts for the special way that blacks have of looking at and evaluating interracial incidents. One respondent, a clerical employee at an adoption agency, described the “second eye” she uses:

I think that it causes you to have to look at things from two different perspectives. You have to decide whether things that are done or slights that are made are made because you are black or they are made because the person is just rude, or unconcerned and uncaring. So it's kind of a situation where you're always kind of looking to see with a

second eye or a second antenna just what's going on.

The language of “second eye” suggests that blacks look at white-black interaction through a lens colored by personal and group experience with cross-institutional and cross-generational discrimination.

Blacks must be constantly aware of the repertoire of possible responses to chronic and burdensome discrimination. One older respondent spoke of having to put on her “shield” just before she leaves the house each morning. When quizzed, she said that for more than six decades, as she leaves her home, she has tried to be prepared for insults and discrimination in public places, even if nothing happens that day. This extraordinary burden of discrimination was eloquently described by the female professor who resented having to worry about life-threatening incidents that her “very close white friends . . . simply don't have to worry about.” Another respondent was articulate on this point:

. . . if you can think of the mind as having one hundred ergs of energy, and the average man uses fifty percent of his energy dealing with the everyday problems of the world—just general kinds of things—then he has fifty percent more to do creative kinds of things that he wants to do. Now that's a white person. Now a black person also has one hundred ergs; he uses fifty percent the same way a white man does, dealing with what the white man has [to deal with], so he has fifty percent left. But he uses twenty-five percent fighting being black, [with] all the problems being black and what it means. Which means he really only has twenty-five percent to do what the white man has fifty percent to do, and he's expected to do just as much as the white man with that twenty-five percent. . . . So, that's kind of what happens. You just don't have as much energy left to do

as much as you know you really could if you were free, [if] your mind were free.

The individual cost of coping with racial discrimination is great, and, as he says, you cannot accomplish as much as you could if you retained the energy wasted on discrimination. This is perhaps the most tragic cost of persisting discrimination in the United States. In spite of decades of civil rights legislation, black Americans have yet to attain the full promise of the American dream.

REFERENCES

- Feagin, Joe R. and Douglas Eckberg. 1980. "Prejudice and Discrimination." *Annual Review of Sociology* 6:1–20.
- Feagin, Joe R. and Clairece Booher Feagin. 1986. *Discrimination American Style* (rev. ed.). Melbourne, FL: Krieger Publishing Co.
- Wilson, William J. 1978. *The Declining Significance of Race*. Chicago: University of Chicago Press.
- _____. 1987. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press.

75 ■ Melvin L. Oliver and Thomas M. Shapiro

Black Wealth/White Wealth

A New Perspective on Racial Inequality

Each year two highly publicized news reports capture the attention and imagination of Americans. One lists the year's highest income earners. Predictably, they include glamorous and highly publicized entertainment, sport, and business personalities. For the past decade that list has included many African Americans: musical artists such as Michael Jackson, entertainers such as Bill Cosby and Oprah Winfrey, and sports figures such as Michael Jordan and Magic Johnson. During the recent past as many as half of the "top ten" in this highly exclusive rank have been African Americans.

Another highly publicized list, by contrast, documents the nation's wealthiest Americans. The famous *Forbes* magazine profile of the nation's wealthiest 400 focuses not on income, but on wealth.¹ This list includes those people whose assets—or com-

mand over monetary resources—place them at the top of the American economic hierarchy. Even though this group is often ten times larger than the top earners list, it contains few if any African Americans. An examination of these two lists creates two very different perceptions of the well-being of America's black community on the eve of the twenty-first century. The large number of blacks on the top income list generates an optimistic view of how black Americans have progressed economically in American society. The near absence of blacks in the *Forbes* listing, by contrast, presents a much more pessimistic outlook on blacks' economic progress.

This chapter develops a perspective on racial inequality that is based on the analysis of private wealth. Just as a change in focus from income to wealth in the discussion

Melvin L. Oliver and Thomas M. Shapiro. *Black Wealth/White Wealth: A New Perspective on Racial Inequality*, pp. 1–10, 12. Copyright © 1997 by Routledge, Inc./Taylor & Francis, Inc. Reproduced by permission of Routledge, Inc./Taylor & Francis, Inc. The excerpts included here are not presented in the same order as they appear in the original.

above provides a different perspective on racial inequality, our analysis reveals deep patterns of racial imbalance not visible when viewed only through the lens of income. This analysis provides a new perspective on racial inequality by exploring how material assets are created, expanded, and preserved.

The basis of our analysis is the analytical distinction between wealth and other traditional measures of economic status, of how people are “making it” in America (for example, income, occupation, and education). Wealth is a particularly important indicator of individual and family access to life chances. Income refers to a flow of money over time, like a rate per hour, week, or year; wealth is a stock of assets owned at a particular time. Wealth is what people own, while income is what people receive for work, retirement, or social welfare. Wealth signifies the command over financial resources that a family has accumulated over its lifetime along with those resources that have been inherited across generations. Such resources, when combined with income, can create the opportunity to secure the “good life” in whatever form is needed—education, business, training, justice, health, comfort, and so on. Wealth is a special form of money not used to purchase milk and shoes and other life necessities. More often it is used to create opportunities, secure a desired stature and standard of living, or pass class status along to one’s children. In this sense the command over resources that wealth entails is more encompassing than is income or education, and closer in meaning and theoretical significance to our traditional notions of economic well-being and access to life chances.

More important, wealth taps not only contemporary resources but material assets that have historic origins. Private wealth thus captures inequality that is the product of the past, often passed down from generation to generation. Given this attribute, in attempting to understand the economic status of

blacks, a focus on wealth helps us avoid the either-or view of a march toward progress or a trail of despair. Conceptualizing racial inequality through wealth revolutionizes our conception of its nature and magnitude, and of whether it is declining or increasing. While most recent analyses have concluded that contemporary class-based factors are most important in understanding the sources of continuing racial inequality, our focus on wealth sheds light on both the historical and the contemporary impacts not only of class but of race. . . .

The argument for class, most eloquently and influentially stated by William Julius Wilson in his 1978 book *The Declining Significance of Race*, suggests that the racial barriers of the past are less important than present-day social class attributes in determining the economic life chances of black Americans. Education, in particular, is the key attribute in whether blacks will achieve economic success relative to white Americans. Discrimination and racism, while still actively practiced in many spheres, have marginally less effect on black Americans’ economic attainment than whether or not blacks have the skills and education necessary to fit in a changing economy. In this view, race assumes importance only as the lingering product of an oppressive past. As Wilson observes, this time in his *The Truly Disadvantaged*, racism and its most harmful injuries occurred in the past, and they are today experienced mainly by those on the bottom of the economic ladder, as “the accumulation of disadvantages . . . passed from generation to generation.”²

We believe that a focus on wealth reveals a crucial dimension of the seeming paradox of continued racial inequality in American society. Looking at wealth helps solve the riddle of seeming black progress alongside economic deterioration. Black wealth has grown, for example, at the same time that it has fallen further behind that of whites.

Wealth reveals an array of insights into black and white inequality that challenge our conception of racial and social justice in America.

The empirical heart of our analysis resides in an examination of differentials in black and white wealth holdings. This focus paints a vastly different empirical picture of social inequality than commonly emerges from analyses based on traditional inequality indicators. The burden of our claim is to demonstrate not simply the taken-for-granted assumption that wealth reveals “more” inequality— $\text{income multiplied } x \text{ times}$ is not the correct equation. More importantly we show that wealth uncovers a qualitatively different pattern of inequality on crucial fronts. Thus the goal of this work is to provide an analysis of racial differences in wealth holding that reveals dynamics of racial inequality otherwise concealed by income, occupational attainment, or education. It is our argument that wealth reveals a particular network of social relations and a set of social circumstances that convey a unique constellation of meanings pertinent to race in America. This perspective significantly adds to our understanding of public policy issues related to racial inequality; at the same time it aids us in developing better policies for the future. In stating our case, we do not discount the important information that the traditional indicators provide, but we argue that by adding to the latter an analysis of wealth a more thorough, comprehensive and powerful explanation of social inequality can be elaborated. . . .

Economists argue that racial differences in wealth are a consequence of disparate class and human capital credentials (age, education, experience, skills), propensities to save, and consumption patterns. A sociology of wealth seeks to properly situate the social context in which wealth generation occurs. Thus the sociology of wealth accounts for racial differences in wealth holding by

demonstrating the unique and diverse social circumstances that blacks and whites face. One result is that blacks and whites also face different structures of investment opportunity, which have been affected historically and contemporaneously by both race and class. We develop three concepts to provide a sociologically grounded approach to understanding racial differentials in wealth accumulation. These concepts highlight the ways in which this opportunity structure has disadvantaged blacks and helped contribute to massive wealth inequalities between the races.

Our first concept, “racialization of state policy,” refers to how state policy has impaired the ability of many black Americans to accumulate wealth—and discouraged them from doing so—from the beginning of slavery throughout American history. From the first codified decision to enslave African Americans to the local ordinances that barred blacks from certain occupations to the welfare state policies of today that discourage wealth accumulation, the state has erected major barriers to black economic self-sufficiency. In particular, state policy has structured the context within which it has been possible to acquire land, build community, and generate wealth. Historically, policies and actions of the United States government have promoted homesteading, land acquisition, home ownership, retirement, pensions, education, and asset accumulation for some sectors of the population and not for others. Poor people—blacks in particular—generally have been excluded from participation in these state-sponsored opportunities. In this way, the distinctive relationship between whites and blacks has been woven into the fabric of state actions. The modern welfare state has racialized citizenship, social organization, and economic status while consigning blacks to a relentlessly impoverished and subordinate position within it.

Our second focus, on the “economic detour,” helps us understand the relatively low level of entrepreneurship among and the small scale of the businesses owned by black Americans. While blacks have historically sought out opportunities for self-employment, they have traditionally faced an environment, especially from the postbellum period to the middle of the twentieth century, in which they were restricted by law from participation in business in the open market. Explicit state and local policies restricted the rights of blacks as free economic agents. These policies had a devastating impact on the ability of blacks to build and maintain successful enterprises. While blacks were limited to a restricted African American market to which others (for example, whites and other ethnics) also had easy access, they were unable to tap the more lucrative and expansive mainstream white markets. Blacks thus had fewer opportunities to develop successful businesses. When businesses were developed that competed in size and scope with white businesses, intimidation and ultimately, in some cases, violence were used to curtail their expansion or get rid of them altogether. The lack of important assets and indigenous community development has thus played a crucial role in limiting the wealth-accumulating ability of African Americans.

The third concept we develop is synthetic in nature. The notion embodied in the “sedimentation of racial inequality” is that in central ways the cumulative effects of the past have seemingly cemented blacks to the bottom of society’s economic hierarchy. A history of low wages, poor schooling, and segregation affected not one or two generations of blacks but practically all African Americans well into the middle of the twentieth century. Our argument is that the best indicator of the sedimentation of racial inequality is wealth. Wealth is one indicator of material disparity that captures the historical legacy of low wages, personal and organiza-

tional discrimination, and institutionalized racism. The low levels of wealth accumulation evidenced by current generations of black Americans best represent the economic status of blacks in the American social structure.

To argue that blacks form the sediment of the American stratificational order is to recognize the extent to which they began at the bottom of the hierarchy during slavery, and the cumulative and reinforcing effects of Jim Crow and de facto segregation through the mid-twentieth century. Generation after generation of blacks remained anchored to the lowest economic status in American society. The effect of this inherited poverty and economic scarcity for the accumulation of wealth has been to “sediment” inequality into the social structure. The sedimentation of inequality occurred because the investment opportunity that blacks faced worked against their quest for material self-sufficiency. In contrast, whites in general, but well-off whites in particular, were able to amass assets and use their secure financial status to pass their wealth from generation to generation. What is often not acknowledged is that the same social system that fosters the accumulation of private wealth for many whites denies it to blacks, thus forging an intimate connection between white wealth accumulation and black poverty. Just as blacks have had “cumulative disadvantages,” many whites have had “cumulative advantages.” Since wealth builds over a lifetime and is then passed along to kin, it is, from our perspective, an essential indicator of black economic well-being. By focusing on wealth we discover how a black’s socioeconomic status results from a socially layered accumulation of disadvantages passed on from generation to generation. In this sense we uncover a racial wealth tax.

Our empirical analysis enables us to raise and answer several key questions about wealth: How has wealth been distributed in

American society over the twentieth century? What changes in the distribution of wealth occurred during the 1980s? And finally, what are the implications of these changes for black-white inequality?

During the eighties the rich got much richer, and the poor and middle classes fell further behind. Why? The Reagan tax cuts provided greater discretionary income for middle- and upper-class taxpayers. One asset whose value grew dramatically during the eighties was real estate, an asset that is central to the wealth portfolio of the average American. Home ownership makes up the largest part of wealth held by the middle class, whereas the upper class more commonly hold a greater degree of their wealth in financial assets. Owning a house is the hallmark of the American Dream, but it is becoming harder and harder for average Americans to afford their own home and fewer are able to do so.

In part because of the dramatic rise in home values, the wealthiest generation of elderly people in America's history is in the process of passing along its wealth. Between 1987 and 2011 the baby boom generation stands to inherit approximately \$7 trillion. Of course, all will not benefit equally, if at all. One-third of the worth of all estates will be divided by the richest 1 percent, each legatee receiving an average inheritance of \$6 million. Much of this wealth will be in the form of property, which, as the philosopher Robert Nozick is quoted as saying in a 1990 *New York Times* piece, "sticks out as a special kind of unearned benefit that produces unequal opportunities."³ Kevin, a seventy-five-year-old retired homeowner interviewed for this study, captures the dilemma of unearned inheritance:

You heard that saying about the guy with a rich father? The kid goes through life thinking that he hit a triple. But really he was born on third base. He didn't hit no triple at all,

but he'll go around telling everyone he banged the fucking ball and it was a triple. He was born there!

Inherited wealth is a very special kind of money imbued with the shadows of race. Racial difference in inheritance is a key feature of our story. For the most part, blacks will not partake in divvying up the baby boom bounty. America's racist legacy is shutting them out. The grandparents and parents of blacks under the age of forty toiled under segregation, where education and access to decent jobs and wages were severely restricted. Racialized state policy and the economic detour constrained their ability to enter the post-World War II housing market. Segregation created an extreme situation in which earlier generations were unable to build up much, if any, wealth. The average black family headed by a person over the age of sixty-five has no net financial assets to pass down to its children. Until the late 1960s there were few older African Americans with the ability to save much at all, much less invest. And no savings and no inheritance meant no wealth.

The most consistent and strongest common theme to emerge in interviews conducted with white and black families was that family assets expand choices, horizons, and opportunities for children while lack of assets limits opportunities. Because parents want to give their children whatever advantages they can, we wondered about the ability of the average American household to expend assets on their children. We found that the lack of private assets intrudes on the dreams that many Americans have for their children. Extreme resource deficiency characterizes several groups. It may surprise some to learn that 62 percent of households headed by single parents are without savings or other financial assets, or that two of every five households without a high school degree lack a financial nest egg. Nearly one-third of

all households—and 61 percent of all black households—are without financial resources. These statistics lead to our focus on the most resource-deficient households in our study—African Americans.

We argue that, materially, whites and blacks constitute two nations. One of the analytic centerpieces of this work tells a tale of two middle classes, one white and one black. Most significant, the claim made by blacks to middle-class status depends on income and not assets. In contrast, a wealth pillar supports the white middle class in its drive for middle-class opportunities and a middle-class standard of living. Middle-class blacks, for example, earn seventy cents for every dollar earned by middle-class whites but they possess only fifteen cents for every dollar of wealth held by middle-class whites. For the most part, the economic foundation of the black middle class lacks one of the pillars that provide stability and security to middle-class whites—assets. The black middle class position is precarious and fragile with insubstantial wealth resources. This analysis means it is entirely premature to celebrate the rise of the black middle class. The glass is both half empty and half full, because the wealth data reveal the paradoxical situation in which blacks' wealth has grown while at the same time falling further behind that of whites.

The social distribution of wealth discloses a fresh and formidable dimension of racial inequality. Blacks' achievement at any given level not only requires that greater effort be expended on fewer opportunities but also bestows substantially diminished rewards. Examining blacks and whites who share similar socioeconomic characteristics brings to light persistent and vast wealth discrepancies. Take education as one prime example: the most equality we found was among the college educated, but even here at the pinnacle of achievement whites control four times as much wealth as blacks with the same degrees.

This predicament manifests a disturbing break in the link between achievement and results that is essential for democracy and social equality.

The central question of this study is, Why do the wealth portfolios of blacks and whites vary so drastically? The answer is not simply that blacks have inferior remunerable human capital endowments—substandard education, jobs, and skills, for example—or do not display the characteristics most associated with higher income and wealth. We are able to demonstrate that even when blacks and whites display similar characteristics—for example, are on a par educationally and occupationally—a potent difference of \$43,143 in home equity and financial assets still remains. Likewise, giving the average black household the same attributes as the average white household leaves a \$25,794 racial gap in financial assets alone.

The extent of discrimination in institutions and social policy provides a persuasive index of bias that undergirds the drastic differences between blacks and whites. We show that skewed access to mortgage and housing markets and the racial valuing of neighborhoods on the basis of segregated markets result in enormous racial wealth disparity. Banks turn down qualified blacks much more often for home loans than they do similarly qualified whites. Blacks who do qualify, moreover, pay higher interest rates on home mortgages than whites. Residential segregation persists into the 1990s, and we found that the great rise in housing values is color-coded.⁴ Why should the mean value of the average white home appreciate at a dramatically higher rate than the average black home? Home ownership is without question the single most important means of accumulating assets. The lower values of black homes adversely affect the ability of blacks to utilize their residences as collateral for obtaining personal, business, or educational loans. We estimate that institutional biases in

the residential arena have cost the current generation of blacks about \$82 billion. Passing inequality along from one generation to the next casts another racially stratified shadow on the making of American inequality. Institutional discrimination in housing and lending markets extends into the future the effects of historical discrimination within other institutions.

Placing these findings in the larger context of public policy discussions about racial and social justice adds new dimensions to these discussions. A focus on wealth changes our thinking about racial inequality. The more one learns about wealth differences, the more mistaken current policies appear. To take these findings seriously, as we do, means not shirking the responsibility of seeking alternative policy ideas with which to address issues of inequality. We might even need to think about social justice in new ways. In some key respects our analysis of disparities in wealth between blacks and whites forms an agenda for the future, the key principle of which is to link opportunity structures to policies promoting asset formation that begin to close the racial wealth gap.

Closing the racial gap means that we have to target policies at two levels. First, we need policies that directly address the situation of African Americans. Such policies are necessary to speak to the historically generated disadvantages and the current racially based policies that have limited the ability of blacks, as a group, to accumulate wealth resources.

Second, we need policies that directly promote asset opportunities for those on the bottom of the social structure, both black and white, who are locked out of the wealth accumulation process. More generally, our analysis clearly suggests the need for massive

redistributional policies in order to reforge the links between achievement, reward, social equality, and democracy. These policies must take aim at the gross inequality generated by those at the very top of the wealth distribution. Policies of this type are the most difficult ones on which to gain consensus but the most important in creating a more just society.

This chapter's underlying goal is to establish a way to view racial inequality that will serve as a guide in securing racial equality in the twenty-first century. Racial equality is not an absolute or idealized state of affairs, because it cannot be perfectly attained. Yet the fact that it can never be perfectly attained in the real world is a wholly insufficient excuse for dismissing it as utopian or impossible. What is important are the bearings by which a nation chooses to orient its character. We can choose to let racial inequality fester and risk heightened conflict and violence. Americans can also make a different choice, a commitment to equality and to closing the gap as much as possible.

NOTES

1. Harold Senecker, 1993, "The *Forbes* 400: The Richest People in America," *Forbes*, 18 October, pp. 110–113.
2. William J. Wilson, 1987, *The Truly Disadvantaged*, Chicago: University of Chicago Press.
3. Nick Ravo, 1990, "A Windfall Nears in Inheritances from the Richest Generation," *New York Times*, 22 July E4.
4. Reynolds Farley and William H. Frey, 1994, "Changes in the Segregation of Whites from Blacks During the 1980s: Small Steps Toward a More Integrated Society," *American Sociological Review* 59: 23–45. Douglas S. Massey and Nancy A. Denton, 1993, *American Apartheid: Segregation and the Making of the Underclass*, Cambridge: Harvard University Press.

76 ■ William Julius Wilson and Richard P. Taub

There Goes the Neighborhood**Race, Ethnic, and Class Tensions in Four Chicago Neighborhoods and Their Meaning in America**

This chapter is an investigation into ethnic, racial, and class dynamics in four neighborhoods in Chicago, a city that has experienced a steep drop in its white population and a sharp rise in Latino residents. Chicago's Latino population grew by nearly 38 percent between 1990 and 2000, while its white population declined by almost 15 percent (see Figure 1). Whites constituted just 31 percent of the population in 2000, down from 38 percent in 1990.¹ African Americans remained Chicago's largest single group at 36 percent, but their share of the population had also dropped slightly—by 1.9 percent—after rising steadily through most of the twentieth century.

To fully capture this ethnic diversity, we felt that the most representative neighborhoods would be those that were neither poor nor affluent. We chose neighborhoods that consisted mainly of the working and lower middle classes—neighborhoods, in short, that best represented ordinary Americans and that were more likely to be the destination of outside racial and ethnic groups seeking desirable and affordable places to live. We selected areas that were populated by different ethnic groups to capture variations in responses to neighborhood change. In 1992, after much preliminary investigation, four

were chosen—Beltway, Dover, Archer Park, and Groveland—on Chicago's South and West sides. Because some of the materials used in this study are quite sensitive, the names of these four neighborhoods are pseudonyms. Beltway was chosen as the white neighborhood, Dover as the white neighborhood in transition, Archer Park as the Latino neighborhood, and Groveland as the African American neighborhood.

The choice of neighborhoods represents what sociologist David Willer calls theoretical sampling; that is, the selection of natural cases that include the necessary conditions for the application of theoretical assumptions, assumptions that steer the research and that are used to interpret the findings.² The field research in this study enabled us to examine certain theoretical assumptions involving race and the social organization of neighborhoods, including Albert Hirschman's general theory of exit, voice, and loyalty—which in this instance is applied to neighborhoods.³ . . .

This investigation was undertaken partly to better understand the factors that produced or prevented what social scientists call “the tipping point”—rapid ethnic turnover.⁴ In many neighborhoods, an infusion of minorities prompted whites to leave and discouraged other whites from replacing them.

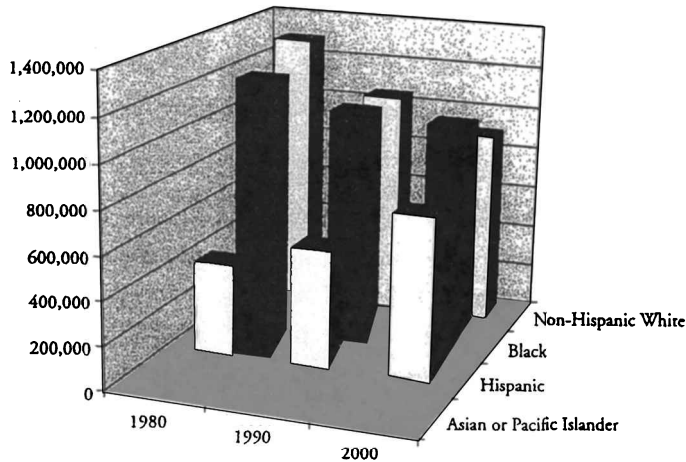


Figure 1. City of Chicago: Total Population by Race and Hispanic Origin, 1980–2000

Thus the proportion of minority-group members grew quickly, particularly after the minority group became a major force in the community. Real estate agents have historically abetted this process by “steering” purchasers to “live with their own kind.”

Why do some neighborhoods experiencing an immigration of outside ethnic groups reach the tipping point more quickly than others, and why do other neighborhoods never reach the tipping point at all? In addressing such questions we can reflect on economist Albert Hirschman’s theory of exit, voice, and loyalty, a theory that was applied to firms and organizations but that is relevant to understanding ethnic changes in neighborhoods. Hirschman argues that when people become dissatisfied with changes in their surroundings they can *exit*—move or withdraw from further participation—or they can exercise *voice*. Hirschman defines “voice” as any attempt “to change, rather than to escape from,” an undesirable situation.⁵

The more willing people are to try to exercise voice—that is, to change, correct, or prevent a particular situation—the less likely

they are to exit. In situations where both exit and voice options are available, past experience will largely determine whether people overcome their biases in favor of exit, the easier option. The view that a neighborhood is on the path of inexorable change, even when these changes have yet to occur, can trigger an exodus. Indeed, Americans maintain a strong bias toward the exit alternative when confronting ethnic and racial changes.

“When general conditions in a neighborhood deteriorate,” writes Hirschman, “those who value most highly neighborhood qualities such as safety, cleanliness, good schools, and so forth will be the first to move out; they will search for housing in somewhat more expensive neighborhoods or in the suburbs and will be lost to the citizens’ groups and community action programs that would attempt to stem and reverse the tide of deterioration.”⁶

The late sociologist Hubert Blalock suggested that when minorities penetrate a neighborhood, existing residents are more likely to remain if new residents bring attractive resources such as high social status, as

when in-migrants include actors, sports celebrities, or other highly prized professionals.⁷ But such neighborhoods are more likely to be affluent to begin with. Minorities who move into neighborhoods populated by ordinary white citizens usually lack high-status social resources. And given that such new residents represent a larger group of people with similar occupational and educational levels and similar language skills, white residents fear that a small trickle could unleash a major invasion.

Of course, in a democratic society people are free to move. However, as Blalock pointed out, some residents might find it too difficult or too expensive to retreat to areas where few or no minorities reside.⁸ Thus, when residents view exit as problematic or costly, they are likely to turn instead to voice. Hirschman offers another explanation for why residents decide to exercise voice: the more they are loyal—that is, attached—to an organization or, in this case, a neighborhood, the less likely they are to exit. Loyalty reflects the extent to which residents are willing “to trade off the certainty of exit against the uncertainties” of improving local conditions.⁹

Loyalty becomes particularly important when it reduces the likelihood that the residents most concerned about neighborhood quality will be the first to depart. Such residents tend to be those with superior social resources and more options; they are also likely to be influential. The longer these residents stay, the less likely it is that other residents in the neighborhood will leave.

Our investigation explored to what extent—and why—residents in selected urban neighborhoods reacted to looming racial, ethnic, or class changes, and how those actions affected neighborhood stability. How residents reacted in turn reflected the dynamics of ethnic, racial, and class tensions—themselves indicative of larger forces at play.

A crucial question that we ask in this chapter is: To what extent are our findings consis-

tent with the assumptions of Hirschman’s theory? Our findings led us to augment his theoretical formulations—that is, to broaden the theoretical sketch and thus provide clearer direction for further research.

The neighborhoods of Beltway, Dover, and Archer Park have one attribute in common: a growing Latino population. In 1980, Beltway—a neighborhood of some 22,000 residents featuring carefully manicured lawns and gardens dotted with stately—was 95 percent white. However, by 2000 the white population had declined to 76 percent and the Latino population had risen to 21 percent.

Dover, with almost 45,000 residents, had experienced a remarkable ethnic transformation: in 1980, 83 percent of the population was white; in 1990, 60 percent of residents were white. However, by 2000 the white population had plummeted to 19 percent, and some 79 percent of Dover’s population was Latino. Indeed, the selection of Dover provided us with the opportunity to observe intergroup tensions in a neighborhood undergoing rapid ethnic change at the time of our research.

Archer Park, a once exclusively white and largely Bohemian neighborhood, had been a Mexican enclave for several decades, and in 2000, 83 percent of the neighborhood’s residents were Latino. Because its Spanish-speaking population included a high number of recent immigrants, Archer Park was poorer than the other neighborhoods. Yet even though the Mexicans who arrived in all these neighborhoods reported lower incomes, on average, than the whites they replaced, their influx shored up a sagging real estate market, and the population in these three communities grew despite white flight.

The only one of the four neighborhoods to register a population decline was Groveland, which experienced virtually no rise in its small number of Latinos, while losing almost 5 percent of its overall population since

1990. A community of close to 12,000 residents and neat single-family houses featuring trim yards, Groveland, in 2000, was 97 percent African American and featured a high percentage of long-term residents. . . .

The Changing Face of Urban Ethnic Neighborhoods

The following is a summary of how exit, voice, and loyalty were exercised as these neighborhoods transformed.

Beltway

Faced with the options of both exit and voice, residents in Beltway have, for the most part, resisted the former alternative. Many had already exited neighborhoods on the South Side of Chicago in response to the growing presence of African Americans. Although high rates of residential stability and racial and class homogeneity have, over the years provided a solid foundation for neighborhood social organization and the maintenance of social control, residents nonetheless worried about the future of their community. Our discussions with Beltwayites revealed that the possible arrival of significant numbers of African Americans, Latinos, and Asians was viewed as a threat to the neighborhood's way of life. Residents saw other neighborhoods change from white-ethnic strongholds to minority enclaves, and those transformations were often accompanied by rising rates of crime, declining property values, and the erosion of neighborhood social institutions.

As has already been mentioned, the exit possibility does not depend solely on an actual threat; it may also be based on a perceived or potential threat. The frequently held view was that the neighborhood could be in danger if certain developments were to materialize. Aware that their neighborhood, like many other formerly white enclaves, could be transformed into a haven for mi-

norities, Beltway residents saw their neighborhood as “the last garden spot in Chicago”; that is, the last of the stable, white, working-class neighborhoods. “Beltway is the last stand” was a popular slogan. It revealed that voice was favored over exit. There are two main reasons for this.

The more obvious reason is that a good percentage of the employed adults in Beltway were on the city's payroll. Their employment status required them to live within the city of Chicago. Thus, leaving their neighborhood was not a viable option. Indeed, in terms of neighborhood stability and social amenities, there were no affordable white working-class areas in the city that constituted a better alternative to Beltway. But voice also took precedence over exit because Beltwayites were fully aware that the neighborhood's strong foundation for stability provided a more than reasonable chance to prevent an ethnic turnover.

Accordingly, there was intense community pressure to maintain a high level of social organization: residents reminded one another to work vigilantly to keep the neighborhood strong and thereby reduce the threat of advancement by outside groups. The prevailing view was that if the neighborhood were allowed to deteriorate, some residents would be encouraged to leave; African Americans and Latinos from nearby communities could then fill the resulting vacancies. Residents were called upon to keep yards clean, to curb gangs, to supervise their children, to remove graffiti, and to work actively to attract new businesses.

Beltway's ability to achieve and sustain a relatively high level of social organization over the years could be attributed to several factors, including residential stability, racial and class homogeneity, a network of dense acquaintanceships, extended families, vibrant community institutions, and powerful connections to the city's political structure. Consequently, the common goal of maintaining

stability in the ethnic and racial makeup of the neighborhood was more easily achieved.

A strong common belief system also contributed to the social organization of Beltway. Integrating sentiments concerning the adherence to neighborhood standards with patriotism and social conservatism, this belief system helped to generate a sense of what Hirschman calls loyalty. In fact, our fieldworkers found that expressions of loyalty were not uncommon in Beltway. As one married woman emphatically put it, “We’re raising our kids here, and I don’t know about you, but we ain’t going nowhere.” If loyalty reinforces voice and discourages exit, Beltway has benefited from the attachment of many residents to the neighborhood.

Nonetheless, despite the expressions of loyalty and vigorous efforts to keep the neighborhood stable, during the 1990s the white population of Beltway decreased (13.5 percent), while the number of minority residents increased, especially Latinos (from 7.5 percent to 21 percent). Our research was conducted in Beltway during the first half of the 1990s, and the residents were fully aware of the demographic changes in the neighborhood, especially the rise in the number of Latino homeowners. However, as Hubert Blalock has pointed out, white residents in neighborhoods with an influx of outside groups and who find it too expensive or too difficult to retreat to areas where few or no nonwhites reside will likely remain.¹⁰ Thus, given the significant presence of municipal workers required to live within the city’s boundaries, combined with Beltway’s relatively high level of social organization, voice may continue to trump exit in the short run, if not over the long haul.

Dover

In sharp contrast to Beltway’s relative stability stands the large exit of whites from Dover. Once an area of working-class Eastern European immigrants, Dover has become a Mexi-

can American enclave. Since 1990 whites have overwhelmingly chosen the exit option and are now represented by only 19 percent of the population. Before the arrival of Mexican Americans in large numbers, Dover was struggling through economic decline. Industrial layoffs combined with the out-migration of many who had acquired the means to buy homes in the suburbs created housing vacancies and lowered property values. Beginning in the 1980s the en masse in-migration of Latinos, many from adjacent Mexican American enclaves, counteracted this trend by sharply increasing housing demand and population density.

Rapid social change influences institutional arrangements and patterns of behavior. The Mexican arrivals were younger than the white residents who remained, and language and cultural differences further separated the groups. Tensions between whites and Latinos simmered beneath the surface and were occasionally vented in public situations, most frequently in conflicts over language use.

During the period of our research, white business and civic organizations were in decline mainly due to a reluctance or an inability to attract Latino participants to fill vacancies created by an aging or fleeing white population. Local churches, on the other hand, drew large numbers of Latino residents, but cultural clashes resulted in ethnic divisions, and the two groups tended to segregate themselves. These developments were indicative of weak neighborhood social organization that encouraged whites to exit the neighborhood in increasing numbers throughout the 1990s. However, whites and Latinos in Dover did find common ground in their response to African Americans. Latinos in Dover—even the recent migrants—were no more open to living with black Americans than were the white residents. Anti-black sentiment was fueled by concerns over Dover’s close proximity to black ghetto

neighborhoods that featured high rates of joblessness and the problems that accompany an acute lack of employment. Although Latino enclaves provided a buffer between Dover and the poor African American neighborhoods, it was a thin cushion at best.

Given the neighborhood's proximity to poor black areas, efforts to prevent African Americans from residing in Dover and the busing of students from overcrowded Dover schools to adjacent poor black neighborhoods presented a rare opportunity for wide-scale white-Latino cooperation. Schools then, provided a setting for ethnic integration and cooperation as Mexican and white parents joined forces to pressure officials to build more classrooms in Dover and thus eliminate the practice of busing.

In the final analysis, the rapid exodus of whites from Dover will very likely continue. What distinguished whites in Dover from whites in Beltway was that prejudiced remarks in Beltway accompanied feelings that they had to prepare to confront an impending threat, whereas prejudiced comments among whites in Dover were associated with feelings that they had "lost the battle." Aside from the cross-ethnic efforts of Dover parents to combat school overcrowding, the exit option has overwhelmed the potential for whites' use of voice—that is, taking steps to correct or address what they perceive to be undesirable changes in the neighborhood. And despite the substantial decline in the number of white residents during the 1990s, the total population of Dover increased from 32,000 to 45,000 during the same period, due to the growth of Latino residents.

Indeed, given the overall increase of the Latino population in Chicago, Latino families that leave Dover for other communities will be replaced by other Latino families moving in. For all these reasons Dover will remain a Latino enclave for the foreseeable future. Given the burgeoning Latino popula-

tion's demand for housing, issues of loyalty, which were so crucial in maintaining the dominant white population in Beltway, were not as important for preserving the Mexican makeup of Dover.

Archer Park

Issues of loyalty are even less important for preserving the ethnic makeup of Archer Park. In the latter decades of the twentieth century, Archer Park changed from a neighborhood overwhelmingly populated by whites, many of them of Bohemian heritage, to a stronghold for Mexican Americans. Indeed, Mexican Americans have been the dominant population in Archer Park since the 1970s. And unlike in Dover and Beltway, residents expressed little concern about neighborhood ethnic change or ethnic or racial challenges to their prerogatives. Except for a few immobile elderly residents, whites have very nearly vanished from the neighborhood. Although our study found that the African American population had increased on the outer fringes of Archer Park, and although a continuous stream of upwardly mobile Latino residents were exiting Archer Park, these trends were counterbalanced by the steady flow of foreign-born Mexicans into the neighborhood. Mexicans were, therefore, firmly in control of the neighborhood and its major institutions, including schools.

Yet it was not a neighborhood that held even the Mexican residents. Upwardly mobile Mexicans exited as soon as they acquired sufficient financial resources to do so. Of our four neighborhoods, Archer Park most clearly distinguished itself as a "stepping-stone" community. Many of the upwardly mobile residents moved to the suburbs or to higher-status Chicago communities, such as Dover, at the first opportunity. Indeed, Dover, which ranked above Archer Park in terms of housing prices and class status, included many Mexican residents who once lived in Archer Park.

Signs of transience rather than permanence pervaded Archer Park; during our time there, even longtime residents expressed a feeling of longing for another neighborhood. This lack of attachment or loyalty to the neighborhood reduced the use of the voice option. Conditions in the neighborhood were less than desirable due in large measure to the extreme overcrowding created by the steady and heavy stream of immigrants from Mexico, as well as by the in-migration of residents from the nearby Mexican community of Pilsen. Yet residents invested little effort in the social organization of the neighborhood with the hope of improving the quality of life. In other words, there was little collective supervision of community activities, and little participation in voluntary organizations such as block clubs, civic and business clubs, parent-teacher organizations, and political groups. The residents simply left Archer Park as soon as they got the chance.

Accordingly, the lack of a perceived threat to Mexican dominance in Archer Park and the low commitment to the neighborhood as a permanent place of residence decreased concerns among residents about the future of the neighborhood and its institutions. Indeed, issues of race and ethnicity were not reflected in the social organization of Archer Park. Yet ethnic antagonisms did prevail. The few remaining elderly whites, although resigned to the idea that Archer Park had become a Mexican enclave, evinced bitterness about the ethnic and cultural changes in the neighborhood and exhibited hostility toward their Mexican neighbors. Moreover, competition among Latinos, blacks, and elderly whites for the use of facilities at the parks and playgrounds contributed to ethnic tensions. Finally, and most fundamentally, Mexican animus toward African Americans in Archer Park revolved around attempts to differentiate themselves from blacks in terms of symbols of social prestige in the larger society.

The more loyal or attached residents are to a neighborhood, the greater the degree of ethnic tensions when the neighborhood is threatened with the possibility of an ethnic turnover. Not only was Archer Park free of any threat of ethnic transformation, but our research suggested that there was relatively little loyalty among the residents in the neighborhood. Accordingly, unlike in Beltway and in Dover, where hostility toward blacks had been associated with concerns about neighborhood ethnic preservation or control of institutions such as local schools, expressions of racial antagonism among Archer Park Latinos derived mainly from their efforts to differentiate themselves from African Americans and thereby ensure that they avoided the stigmas associated with race.

Groveland

If residential stability is a sign of loyalty to a neighborhood, the residents of Groveland were perhaps the most loyal of our four communities. Relatively few families exited the neighborhood once they arrived. The residential turnover, measured in terms of those living in the same house over a five-year period, was the lowest among the four neighborhoods in this study. Moreover, in contrast to the palpable racial and ethnic antagonisms in Beltway and Dover and the evident, if not as intense, animosities in Archer Park, ethnic and racial antagonisms in the African American neighborhood of Groveland were far less obvious.

Nonetheless, this lower-middle-class black neighborhood continually struggled with the issues associated with being African American in contemporary U.S. society. Race remained a topic of discussion even though interracial contact within Groveland was minimal. However, although the racial dialogue in Groveland included some anti-white sentiments, it did not reverberate with the hostile expressions about other groups so

typically seen in statements made by Dover, Beltway, and Archer Park residents. In fact, discussions about race observed by our field researchers overwhelmingly reflected concerns about forging a positive black identity rather than overtly contesting the dominant white population.

Several factors accounted for the relatively neutral racial attitudes in Groveland. A substantial number of the neighborhood's lower-middle-class blacks worked in government jobs, whose rules of hiring and employment restricted competition that might be attributed solely to racial reasons. The employment conditions of many Groveland residents reduced concerns that consume upper-middle-class black professionals in the larger society, where the more subjective nature of employment leads to anxieties about racial bias in salary differentials and promotions. Furthermore, since Groveland was a segregated lower-middle-class community, the involvement and interaction of the residents with whites and other ethnic groups were limited, and the adverse effects of prejudice and other racial indignities were thereby minimized. Finally, Groveland was not experiencing an immigration of other ethnic groups. In fact, the combined population of whites, Latinos, and Asians was small and declining. Accordingly, institutions in Groveland were not only firmly controlled by African Americans, there was virtually no ethnic competition for public space and housing.

Yet the fact that ethnicity was not a factor in the day-to-day activities in Groveland did not remove anxieties about neighborhood change. The residents were concerned about the pressures from surrounding ghetto neighborhoods. As an attractive lower-middle-class community, Groveland enjoyed a quality of life that far exceeded that of the surrounding, less economically advantaged black communities. Families from nearby poor neighborhoods had begun to trickle in,

and their behavior was carefully scrutinized. In block-club meetings residents expressed concern about the future of the neighborhood and what could happen if significant segments of the ghetto poor were able to gain a foothold in the community.¹¹

In some respects, Groveland residents' expression of concern about the penetration of surrounding lower-class blacks into their neighborhood was not dissimilar to Beltway whites' apprehension about the entry of people of color into theirs. The fundamental difference, of course, was that the intergroup anxieties in Groveland represented class, not racial, antagonisms. Nonetheless, both neighborhoods were concerned about the impact of outside groups on community stability, and both neighborhoods feared lower-class populations. Although blacks in Groveland did not express their feelings about lower-class or ghetto blacks in racist terms, and although whites in Beltway were in many respects mixing race and class antagonisms in their sentiments about the penetration of outsiders, both communities were anxious about lower-income populations settling in their neighborhoods and concerned with issues raised by living in close proximity to poor people. Indeed, field notes compiled by our researchers revealed that the residents of each neighborhood expressed concerns about renters, who did not, in their judgment, share the values of homeowners.

Confronted with the options of exit and voice, Groveland residents seemed to be relying on the latter as they reacted to the behavior of recent residents, which sometimes challenged the traditions established by the older residents. Pressure was put on the new residents to conform to neighborhood standards. Loyalty to the neighborhood may play a significant role in this connection. A key factor in determining the future demographic makeup of Groveland is whether neighborhood loyalty will help reinforce voice and minimize exit if the in-migration

from adjacent poor communities increases. What is clear, however, is that Groveland will remain racially segregated regardless of the class composition of the neighborhood. Whereas ethnic separation in the other three neighborhoods reflected a dynamic sifting process as different groups moved in and out, Groveland was unique in the sense that the in-migration was restricted to one group—African Americans.

The Racial and Ethnic Separation of Neighborhoods and Social Policy Dilemmas

The findings of this research suggest several general principles that have implications for social policies in addressing the problems of racial and ethnic antagonism in large American metropolitan areas. In general, when residents perceive that in-migration presents a threat to their neighborhood, they will react either by exiting or by joining forces with other neighbors to resist the change. The stronger the social organization of the neighborhood, the more likely it is that local residents will select the voice option and take steps to keep the area stable. Residents are more likely to choose the exit option when they feel that a neighborhood's resources, including the social organization of the community, are insufficient to stem the tide of ethnic change.

The greater this feeling among residents, the more quickly the neighborhood reaches the "tipping point," the beginning of a very rapid ethnic turnover.¹² However, the literature on neighborhood racial or ethnic change does not provide a general set of arguments that relate the tipping point to neighborhood social organization, arguments that would help explain more fully why some neighborhoods reach the tipping point more rapidly than do others, or why some neighborhoods undergoing racial or ethnic change never reach the tipping point.¹³

The findings of this study suggest a positive relationship between the strength of neighborhood social organization and the length of time it takes to reach the tipping point after an ethnic invasion occurs. In our comparisons of Beltway and Dover, we highlighted the fact that Dover, unlike Beltway, had neither the institutional strength nor the population stability to prevent rapid ethnic turnover. Dover struggled through a period of economic decline in the 1970s, prior to the en masse in-migration of Mexicans, which resulted in significant out-migration of white residents, housing vacancies, and declining property values. The influx of Mexicans shored up the housing market in the 1980s. Nonetheless, by 1990 whites still constituted a substantial majority of the neighborhood's population. But weak neighborhood social organization—as reflected in public conflicts over the use of language, declining business and civic organizations, and cultural clashes between whites and Latinos in local churches—cleared the path for a rapid ethnic turnover. By contrast, Beltway had been a model of population and institutional stability until 1990. Although Beltway experienced some population movements in the 1990s, it remained a predominantly white community and continued to feature organizations that effectively resisted unwanted neighborhood change. . . .

Two parallel themes emerge from this examination of the four neighborhoods. One is our exploration of what makes for a stable and strong community, which everyone wants. We've attempted to shed light on how exit, voice, and loyalty determine whether neighborhoods continue strong and relatively stable. Strong neighborhoods often remain so in opposition to other groups of people, and this is the parallel theme of our book. Together, these forces work against the notion of intergroup harmony and integration in neighborhoods, schools, and the overall society that underlies the U.S. ideal.

In other words, strong neighborhoods and community identities are a double-edged sword. Efforts to develop and sustain strong communities like Beltway and Groveland create resources that can be used to prevent or impede unwanted neighborhood integration—whether it be racial, as in Beltway, or class-based, as in Groveland.

On the other hand, neighborhoods that feature a weak social organization, such as Dover, will see their residents confront an unwanted ethnic invasion not by resorting to voice to prevent or stem the tide, but by choosing to exit, as whites in Dover did following a Latino invasion. This will eventually create a tipping point, or a rapid ethnic/racial transformation of the neighborhood.

Although the immediate future of Beltway is uncertain, a few decades ago Groveland changed from white to African American and Archer Park shifted from white to Latino. At the time of our study, Dover had only recently undergone an ethnic turnover, from white to Latino. For the most part, these population transformations have generated incredible racial and ethnic tensions. However, once a neighborhood has been transformed ethnically and the prospects for further ethnic change are low, racial and ethnic tensions within the neighborhood subside. Groveland, and to a lesser extent Archer Park, are cases in point.

Nonetheless, since urban neighborhoods are divided racially, ethnically, and culturally, the potential for ethnic conflict in the larger city is always present because groups are far more likely to focus on their differences than on their commonalities. Although they may be comfortable in their own monolithic neighborhoods, and although many of their members work in integrated employment settings, these groups are likely to view one another as competitors for political, social, and cultural resources. . . .

If there was a common concern among the residents of all four neighborhoods, it was

the prevalence of crime and other social dislocations in nearby black ghetto neighborhoods. We cannot continue to ignore the needs of inner-city ghetto residents because these problems will not go away. If unaddressed, they may worsen, further driving middle-class people—including middle-class blacks—to abandon central cities, and further heightening racial tensions in ethnic neighborhoods such as Beltway, Dover, and Archer Park, and class tensions in non-poor African American neighborhoods such as Groveland. Thus, underlying much of the racial and class tensions in urban areas is the unresolved question of how to address the needs of low-income urban black residents.

The problems of social dislocation in inner-city ghettos—including poverty, joblessness, failing schools, and family breakups—are so massive that they require federal support on a significant scale if they are to be meaningfully addressed. One of the challenges facing efforts to achieve effective multiracial coalition building is mustering sufficient political pressure to revive funding for cities so that resources would be available to attack these problems.

However, if one of the explicit aims is to ameliorate problems in inner-city black ghettos, there are two major and related obstacles to overcome. The first is the general belief—prominently held by white and black conservatives—that racism is no longer an issue in this country;¹⁴ the second is that government programs directed at improving conditions of people of color, including African Americans, are no longer warranted. Underlying both views is the following belief: if individual minorities are not succeeding today, it is due in large measure to their own shortcomings or inadequacies, not to lack of opportunity or to racism in the larger society. It is reasonable to assume that both views help to explain why white Americans overwhelmingly object to government assistance targeted at blacks. Today, whereas eight out of

every ten African Americans feel that the government is not spending enough to assist black people, only slightly more than three of every ten white Americans feel this way.

The idea that the federal government “has a special obligation to help improve the living standards of blacks” because they “have been discriminated against for so long” was supported by only one in five whites in 2001, and has never exceeded more than one in four since 1975. Significantly, the lack of white support for this idea is not related to background factors such as level of education and age.

How much of this opposition to government assistance can be attributed to stereotypes about the cultural traits of blacks, including attitudes, orientations, worldviews, habits, and behavioral styles? In other words, how much of the opposition takes the form of racial ideology that Lawrence Bobo and his colleagues refer to as “laissez-faire racism,” a racism in which blacks are perceived as being responsible for their own economic predicament and therefore undeserving of special government support?¹⁵ In this connection, James Kluegel’s study of trends in white explanations of the black-white economic gap provides persuasive testimony to continued racial stereotyping: between 1977 and 1989, the most frequently expressed reason for the economic gap was lack of motivation on the part of African Americans.¹⁶ Such attitudes lead one to overlook the devastating cumulative effects of problems such as failing public schools that service children in inner-city ghetto neighborhoods and employer discrimination that continues to exacerbate the employment woes of low-income adult blacks, especially black males.¹⁷

Accordingly, if one of the aims is to address conditions in inner-city ghettos, coalition leaders will have to take into account the underlying and persistent laissez-faire racism that is prevalent today. In doing so, the issues would have to be framed in such a way that

government programs for cities would be seen as benefiting all segments of the urban population—whites, African Americans, Latinos, Asians, and Native Americans. This doesn’t mean that special programs would not be targeted at improving life in inner-city ghetto neighborhoods—such as public-service jobs and job training, economic development assistance, and neighborhood revitalization—but rather that such programs would be part of a more comprehensive package of reforms. Programs designed to improve the overall quality of life in the larger city would include those that are targeted at specific neighborhoods, including inner-city ghetto neighborhoods.

The American ideal of integrated neighborhoods may not be achieved for all the reasons stated in this chapter, but coalition-building would at least have the potential to create a sense of group interdependence, reduce racial and ethnic conflict, and enable diverse groups to live side by side in harmony, not fear.

NOTES

1. The racial category “whites,” as used in this book, refers to the census classification of “non-Hispanic whites.” Census questions prior to 2000 allowed respondents to self-identify their race as White, Black, Native American, Asian or Pacific Islander, or Other. A follow-up question then asked respondents if they were of Hispanic ethnicity. The most recent census included an additional option for the race question, which allowed respondents to self-identify as bi- or multiracial. In previous years, these latter groups would have been counted in one of the single race categories. Therefore, population totals by race for 1980 and 1990 are not directly comparable to those for 2000. For the 2000 population figures presented in this book, respondents of Hispanic origin can be of any race, including those using the new multiracial category. The other racial categories, such as Non-Hispanic White, refer to people who classify themselves as being of only one race and not of Hispanic origin. In this book, the word “Latino” is used when referring to a person described in the census as Hispanic.

2. David Willer, *Scientific Sociology: Theory and Method* (Englewood Cliffs, N.J.: Prentice Hall, 1967).

3. Albert O. Hirschman, *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States* (Cambridge: Harvard University Press, 1970).

4. Thomas C. Schelling, "Dynamic Models of Segregation," *Journal of Mathematical Sociology* 1 (1971): 143–86; Thomas C. Shelling, *Micromotives and Macrobehavior* (New York: Norton, 1978), pp. 135–66; Richard P. Taub, D. Garth Taylor, and Jan D. Durham, *Paths of Neighborhood Change: Race and Crime in Urban America* (Chicago: University of Chicago Press, 1984); William A. V. Clark, "Residential Preferences and Neighborhood Racial Segregation: A Test of the Schelling Segregation Model," *Demography* 28 (1991): 1–19; and John Yinger, *Closed Doors, Opportunities Lost: The Continuing Costs of Housing Discrimination* (New York: Russell Sage Foundation, 1995).

5. Hirschman, *Exit, Voice, and Loyalty*, p. 30.

6. *Ibid.*, p. 51.

7. Hubert M. Blalock, *Toward a Theory of Minority-Group Relations* (New York: John Wiley & Sons, 1967). Recent research indicates that residential segregation has declined since 1970 mainly for non-poor families. On the other hand, poor black families remain uniquely segregated. See Mary J. Fischer, "The Relative Importance of Income and Race in Determining Residential Outcomes in U.S. Urban Areas," *Urban Affairs Review* 38, no. 5 (May 2001): 669–96. Also see Ingrid Gould Ellen, *Sharing America's Neighborhoods: The Prospects for Stable Racial Integration* (Cambridge: Harvard University Press, 2000).

8. Blalock, *Toward a Theory of Minority-Group Relations*.

9. Hirschman, *Exit, Voice, and Loyalty*.

10. Hubert M. Blalock, *Toward a Theory of Minority-Group Relations* (New York: John Wiley & Sons, 1967).

11. For a discussion of middle-class blacks' concerns about low-income blacks moving into their

neighborhood in another Chicago community, see Richard P. Taub, D. Garth Taylor, and Jan D. Durham, *Paths of Neighborhood Change: Race and Crime in Urban America* (Chicago: University of Chicago Press, 1984).

12. Thomas C. Schelling, "Dynamic Models of Segregation," *Journal of Mathematical Sociology* 1 (1971): 143–86; Thomas C. Shelling, *Micromotives and Macrobehavior* (New York: Norton, 1978), pp. 135–66; Taub, Taylor, and Durham, *Paths of Neighborhood Change*; William A. V. Clark, "Residential Preferences and Neighborhood Racial Segregation: A Test of the Schelling Segregation Model," *Demography* 28 (1991): 1–19; and John Yinger, *Closed Doors, Opportunities Lost: The Continuing Costs of Housing Discrimination* (New York: Russell Sage Foundation, 1995).

13. For a comprehensive discussion of this literature, see Ingrid Gould Ellen, *Sharing America's Neighborhoods: The Prospects for Stable Racial Integration* (Cambridge: Harvard University Press, 2000).

14. See, for example, Dinesh D'Souza, *The End of Racism: Principles for a Multiracial Society* (New York: Free Press, 1995); and Walter E. Williams, *More Liberty Means Less Government: Our Founders Knew This Well* (Stanford, Calif.: Hoover Institution, 1999).

15. See Lawrence Bobo and James R. Kluegel, "Status, Ideology, and Dimensions of Whites' Racial Beliefs and Attitudes: Progress and Stagnation," in *Racial Attitudes in the 1990s*, ed. Steven A. Tuch and Jack K. Martin (Westport, Conn.: Praeger, 1997), pp. 93–120.

16. James R. Kluegel, "Trends in Whites' Explanations of the Gap in Black-White Socioeconomic Status, 1977–1989," *American Sociological Review* 55 (1990): 512–25.

17. Wilson, *When Work Disappears*; and Jean Anyon, *Ghetto Schooling: A Political Economy of Urban Educational Reform* (New York: Teachers College Press, 1997).

PART VII

Gender Inequality

Labor Force Participation

77 ■ Arlie Russell Hochschild

The Time Bind

When Work Becomes Home and Home Becomes Work

The Waving Window

It is 6:45 A.M. on a fine June day in the mid-western town of Spotted Deer. At a childcare center in the basement of the Baptist church, Diane Caselli, a childcare worker in blue jeans and loose shirt, methodically turns over small upended chairs that rest on a Lilliputian breakfast table. She sets out small bowls, spoons, napkins, and a pitcher of milk around a commanding box of Cheerios. The room is cheerful, clean, half-asleep. Diane moves slowly past neatly shelved puzzles and toys, a hat rack hung with floppy, donated dress-up hats and droopy pocketbooks, a tub filled with bits of colored paper. Paintings of swerving trains and tippy houses are taped to the wall.

At seven, a tall, awkward-looking man peers hesitantly into the room, then ventures a few steps forward looking for Diane. His son Timmy tromps in behind him. Diane walks over, takes Timmy's hand, and leads him to the breakfast table, where she seats him and helps him pour cereal and milk into his bowl. Timmy's dad, meanwhile, hurries toward the door.

One wall of the room has four large windows that overlook a sidewalk. In front of

the second window is a set of small wooden steps children climb to wave good-bye to their departing parents. It's called "the waving window." Timmy dashes from the breakfast table, climbs up the wooden steps, and waits.

His dad, an engineer, briskly strides past the first window toward his red Volvo parked down the street. He stops for a moment in front of the waving window, tilts his head, eyebrows lifted clownishly, then walks on without a backward glance. Timmy returns to his cereal, sighs, and declares excitedly, "My Dad sawed me wave!" . . .

At 7:40 A.M., four-year-old Cassie sidles in, her hair half-combed, a blanket in one hand, a fudge bar in the other. "I'm late," her mother explains to Diane. "Cassie wanted the fudge bar so bad, I gave it to her," she adds apologetically—though Diane has said nothing. Gwen Bell is a sturdy young woman, with short-cropped dark hair. Lightly made up and minimally adorned with gold stud earrings, she is neatly dressed in khaki slacks and jacket. Some Amerco mothers don business suits as soldiers don armor while a few wear floral dresses suggest-

ing festivity and leisure. But Cassie's mother is dressed in a neutral way, as if she were just getting the job of self-presentation done.

"Pleese, can't you take me with you?" Cassie pleads.

"You know I can't take you to work," Gwen replies in a tone that suggests she's heard this request before. Cassie's shoulders droop in defeat. She's given it a try, but now she's resigned to her mother's imminent departure, and she's agreed, it seems, not to make too much fuss about it. Aware of her mother's unease about her long day at childcare, however, she's struck a hard bargain. Every so often she gets a morning fudge bar. This is their deal, and Cassie keeps her mother to it. As Gwen Bell later explained to me, she continually feels that she owes Cassie more time than she actually gives her. She has a time-debt to her daughter. If many busy parents settle such debts on evenings or weekends when their children eagerly "collect" promised time, Cassie insists on a morning down payment, a fudge bar that makes her mother uneasy but saves her the trouble and embarrassment of a tantrum. Like other parents at the center, Gwen sometimes finds herself indulging her child with treats or softened rules in exchange for missed time together. Diane speaks quietly to Cassie, trying to persuade her to stop sulking and join the others.

The center works on "child time." Its rhythms are child-paced, flexible, mainly slow. Teachers patiently oversee the laborious task of tying a shoelace, a prolonged sit on the potty, the scrambled telling of a tall tale. In this and other ways it is an excellent childcare center, one of a dozen islands of child time I was to discover in my three summers of field research at Amerco, a Fortune 500 company headquartered in Spotted Deer.¹ Scattered throughout the town, such islands—a playground, a pediatrician's waiting room, the back of a family van—stand out against the faster paced, more bu-

reaucratically segmented blocks of adult work time. . . .

At the Spotted Deer Childcare Center, the group of young breakfasters gradually expands, early arrivals watching the entertainment provided by yet more newcomers. Sally enters sucking her thumb. Billy's mother carries him in even though he's already five. Jonathan's mother forgets to wave, and soon after, Jonathan kicks the breakfast table from below, causing milk to spill and children to yell. Marie ushers him away to dictate a note to his mother explaining that it hurts his feelings when she doesn't wave.

Cassie still stands at the front door holding her fudge bar like a flag, the emblem of a truce in a battle over time. Every now and then, she licks one of its drippy sides while Diane, uncertain about what to do, looks on disapprovingly. The cereal eaters watch from their table, fascinated and envious. Gwen Bell turns to leave, waving goodbye to Cassie, car keys in hand. By our prior arrangement, I am her shadow for this day, and I follow her out the door and into the world of Amerco.

Arriving at her office, as always, exactly at 7:50 A.M., Gwen finds on her desk a cup of coffee in her personal mug, milk no sugar (exactly as she likes it), prepared by a coworker who has managed to get in ahead of her. The remaining half of a birthday cake has been left on a table in the hall outside her office by a dieting coworker who wants someone else to eat it before she does. Gwen prepares materials for her first meeting (having e-mailed messages to the other participants from home the previous night), which will inaugurate her official 8 A.M. to 5:45 P.M. workday. As she does so, she nibbles at a sliver of the cake while she proofreads a memo that must be Xeroxed and handed out at a second meeting scheduled for 9 A.M.

As the assistant to the head of the Public Relations Office, Gwen has to handle Amerco's responses to any press reports that

may appear about the company. This time the impending media attention is positive, and her first meeting has been called to discuss how to make the most of it. As the members of the publicity team straggle into her office and exchange friendly greetings, she sighs. She's ready.

Gwen loves her job at Amerco, and she is very good at it. Whatever the daily pressures, she also feels remarkably at home there. Her boss, a man who she says reminds her of the best aspects of her father, helps her deal with work problems and strongly supports her desire to rise in the company. In many ways, her "Amerco dad" has been better to her than her own father, who, when she was small, abruptly walked out on her mother and her. She feels lucky to have such a caring boss, and working for him has reinforced her desire to give her all to work—insofar as she can. She and her husband need her salary, she tells me, since his job, though more lucrative than hers, is less steady.

Gradually, over the last three years Gwen's workday has grown longer. She used to work a straight eight-hour day. Now it is regularly eight and a half to nine hours, not counting the work that often spills over into life at home. Gwen is not happy about this. She feels Cassie's ten-hour day at Spotted Deer is too long, but at the same time she is not putting energy into curbing her expanding workday. What she does do is complain about it, joke about it, compare stories with friends at work. Hers are not the boastful "war stories" of the older men at Amerco who proudly proclaim their ten-hour workdays and biweekly company travel schedules. Rather, Gwen's stories are more like situation comedies: stories about forgetting to shop and coming home to a refrigerator containing little more than wilted lettuce and a jar of olives, stories told in a spirit of hopeless amusement. Gwen is reasonably well informed about Amerco's flextime and reduced-hours policies, which

are available to white-collar employees like her. But she has not talked with her boss about cutting back her hours, nor have her joking coworkers, and her boss hasn't raised the possibility himself. There is just so much to get done at the office.

At 5:45 P.M. on the dot, Gwen arrives back at Spotted Deer. Cassie is waiting eagerly by the door, her coat over her arm, a crumpled picture she has drawn in her hand. Gwen gives Cassie a long, affectionate hug. By the time Gwen and Cassie roll into the driveway of their two-story white frame house, surrounded by a border of unruly shrubs, it is 6:25 P.M. John Bell is already there, having shopped, taken the messages off the phone machine, set the table, and heated the oven. This is one of two days a week he leaves home and returns earlier than his wife. He has eaten an ample lunch, knowing that they usually have a late, light dinner, but this evening he's hungrier than he means to be. He plays with Cassie while Gwen makes dinner.

To protect the dinner "hour"—8:00 to 8:30—Gwen checks that the phone machine is on, and we hear a series of abbreviated rings several times during dinner. John says grace, and we all hold hands. It is time, it seems, to let go and relax. Right after dinner, though, it's Cassie's bath time. Cassie both loves and protests her nightly bath. Gwen has come to expect Cassie to dawdle as she undresses or searches for a favorite bath toy, trying to make a minivacation of it. Gwen lets Cassie linger, scans through her phone messages, and sets them aside.

At 9 P.M., the bath over, Gwen and Cassie have "quality time" or "QT," as John affectionately calls it. This they see as their small castle of time protected from the demands of the outside world. Half an hour later, at 9:30 P.M., Gwen tucks Cassie into bed.

Later, as Gwen and John show me around their home, John points out in passing an expensive electric saw and drill set he bought

two years earlier with the thought of building a tree house for Cassie, a bigger hutch for her rabbit Max, and a guest room for visiting friends. “I have the tools,” John confides. “I just don’t have the time to use them.” Once, those tools must have represented the promise of future projects. Now they seemed to be there in place of the projects. Along with the tools, perhaps John has tried to purchase the illusion of leisure they seemed to imply. Later, as I interviewed other working parents at Amerco, I discovered similar items banished to attics and garages. Timmy’s father had bought a boat he hadn’t sailed all year. Jarod and Tylor’s parents had a camper truck they had hardly driven. Others had cameras, skis, guitars, encyclopedia sets, even the equipment to harvest maple syrup, all bought with wages that took time to earn.

John’s tools seemed to hold out the promise of another self, a self he would be “if only I had time.” His tools had become for him what Cassie’s fudge bar was for her—a magical substitute for time, a talisman.

There were, in a sense, two Bell households: the rushed family they actually were and the relaxed family they imagined they could be if only they had time. Gwen and Bill complained that they were in a time bind: they wanted more time for life at home than they had. It wasn’t that they wanted more small segments of “quality time” added into their over-busy days. They wanted a quality life, and Gwen, at least, worked for a family-friendly company whose policies seemed to hold out hope for just that. So what was preventing them from getting it?

Time in the Balance

We’ve gotten ourselves into a time bind. Feeling that we are always late, having no free time, trying to adapt as best we can to the confines of our time prisons—these are all symptoms of what has become a national way of life. There are several reasons for this.

Over the past two decades, global competition and inflation have lowered the buying power of the male wage. In response, many women have gone to work in order to maintain the family income. But the legacy of patriarchy has given cultural shape to this economic story. As women have joined men at work, they have absorbed the views of an older, male-oriented work world—its views of “an honest day’s work”—at a much faster rate than men have absorbed their share of domestic work and culture. One reason women have changed more than men is that the world of “male” work seems more honorable and valuable than the “female” world of home and children.

There is another factor too—we are increasingly anxious about our “culture of care.” Where do we turn when we are down and out? This is a question we ask, understandably, even when we are up and in. With recent welfare “reform,” the government is cutting off aid to women and the poor. With the growth of the benefit-less contingency labor force, many corporations are doing the same thing to middle-class men. Meanwhile, modern families have grown more ambiguous: It’s a little less clear than it once was who’s supposed to take care of whom, how much and for how long. So we’ve grown more anxious. Given our tradition of individualism, many of us feel alone in this anxiety, and insofar as work is our rock, we cling to it.

In this context, the idea of cutting back the workday is an idea that seems to have died, gone to heaven and become an angel of an idea. We dream about it, but it’s something we’d never really expect to do. No matter how a movement for work-family balance is structured, in the long run, no such balance will ever take hold if the social conditions that would make it possible—men who are willing to share parenting and housework, supportive communities, and policy-makers and elected officials who are prepared to

demand family-friendly reforms—remain out of reach. . . .

As my study of Amerco has shown, however, even when the jobs of working parents are secure, pay a sufficient wage and provide family-friendly programs, many working parents are still reluctant to spend more time at home. American fathers spend less time at home than mothers do, expand their work hours when children are small and, if Amerco is any indication, are reluctant to take paternity leaves. We know from previous research that many men have found a haven at work. This isn't news. The news is that growing numbers of working women are leery of spending more time at home as well. They feel guilty and stressed out by long hours at work, but they feel ambivalent about cutting back on those hours.

Women fear losing their place at work; having such a place has become a source of security, pride and a powerful sense of being valued. As a survey conducted by Bright Horizons (a Boston-based company that runs on-site daycare centers in twenty-three states) indicates, women are just as likely to feel appreciated as men at the workplace; as likely as men to feel underappreciated at home; and even more likely than men to have friends at work. To cut back on work hours means risking loosening ties to a world that, tension-filled as it is, offers insurance against even greater tension and uncertainty at home. For a substantial number of time-bound working parents, the stripped-down home and the community-denuded neighborhood are simply losing out to the pull of the workplace.

Many women are thus joining men in a flight from the “inner city” of home to the “suburbs” of the workplace. In doing so, they have absorbed the views of an older, male-oriented work world about what a “real career” and “full commitment to the job” really mean. Women now make up nearly half the labor force. The vast majority of them need and want to be there. There is definitely no going back. But women have entered the workplace on “male” terms. It would be less problematic for women to adopt a male model of work—to finally enjoy privileges formerly reserved for men—if the male model of work were one of balance. But it is not.

All this is unsettling news, in part because the children of working parents are being left to adjust to or resist the time bind—and all of its attendant consequences—more or less on their own. It is unsettling because while children remain precious to their parents, the “market value” of the world in which they are growing up has declined drastically. One need not compare their childhoods with a perfect childhood in a mythical past to conclude that our society needs to face up to a serious problem.

NOTE

1. To protect the privacy of the people in this chapter, I have given the company a fictional name and declined to specify what its workers produce or where they live. I've also altered the names, occupations, and defining details of the personal lives of individuals. I have, however, tried to document as accurately as possible their experiences of life at home and at work and to capture the essence of the culture that infuses both worlds.

78 ■ Lisa Belkin

The Opt-Out Revolution

The scene in this cozy Atlanta living room would—at first glance—warm an early feminist’s heart. Gathered by the fireplace one recent evening, sipping wine and nibbling cheese, are the members of a book club, each of them a beneficiary of all that feminists of 30-odd years ago held dear.

The eight women in the room have each earned a degree from Princeton, which was a citadel of everything male until the first co-educated class entered in 1969. And after Princeton, the women of this book club went on to do other things that women once were not expected to do. They received law degrees from Harvard and Columbia. They chose husbands who could keep up with them, not simply support them. They waited to have children because work was too exciting. They put on power suits and marched off to take on the world.

Yes, if an early feminist could peer into this scene, she would feel triumphant about the future. Until, of course, any one of these polished and purposeful women opened her mouth.

“I don’t want to be on the fast track leading to a partnership at a prestigious law firm,” says Katherine Brokaw, who left that track in order to stay home with her three children. “Some people define that as success. I don’t.”

“I don’t want to be famous; I don’t want to conquer the world; I don’t want that kind of life,” says Sarah McArthur Amsbary, who was a theater artist and teacher and earned her master’s degree in English, then stepped

out of the work force when her daughter was born. “Maternity provides an escape hatch that paternity does not. Having a baby provides a graceful and convenient exit.”

Wander into any Starbucks in any Starbucks kind of neighborhood in the hours after the commuters are gone. See all those mothers drinking coffee and watching over toddlers at play? If you look past the Lycra gym clothes and the Internet-access cell phones, the scene could be the ’50s, but for the fact that the coffee is more expensive and the mothers have M.B.A.s. . . .

Arguably, the barriers of 40 years ago are down. Fifty percent of the undergraduate class of 2003 at Yale was female; this year’s graduating class at Berkeley Law School was 63 percent women; Harvard was 46 percent; Columbia was 51. Nearly 47 percent of medical students are women, as are 50 percent of undergraduate business majors (though, interestingly, about 30 percent of M.B.A. candidates). They are recruited by top firms in all fields. They start strong out of the gate.

And then, suddenly, they stop. Despite all those women graduating from law school, they comprise only 16 percent of partners in law firms. Although men and women enter corporate training programs in equal numbers, just 16 percent of corporate officers are women, and only eight companies in the Fortune 500 have female C.E.O.s. Of 435 members of the House of Representatives, 62 are women; there are 14 women in the 100-member Senate.

Measured against the way things once were, this is certainly progress. But measured against the way things were expected to be, this is a revolution stalled. During the '90s, the talk was about the glass ceiling, about women who were turned away at the threshold of power simply because they were women. The talk of this new decade is less about the obstacles faced by women than it is about the obstacles faced by mothers. As Joan C. Williams, director of the Program on WorkLife Law at American University, wrote in the *Harvard Women's Law Journal* last spring, "Many women never get near" that glass ceiling, because "they are stopped long before by the maternal wall."

Look, for example, at the Stanford class of '81. Fifty-seven percent of mothers in that class spent at least a year at home caring for their infant children in the first decade after graduation. One out of four have stayed home three or more years. Look at Harvard Business School. A survey of women from the classes of 1981, 1985 and 1991 found that only 38 percent were working full time. Look at professional women in surveys across the board. Between one-quarter and one-third are out of the work force, depending on the study and the profession. Look at the United States Census, which shows that the number of children being cared for by stay-at-home moms has increased nearly 13 percent in less than a decade. At the same time, the percentage of new mothers who go back to work fell from 59 percent in 1998 to 55 percent in 2000.

Look, too, at the mothers who have not left completely but have scaled down or redefined their roles in the crucial career-building years (25 to 44). Two-thirds of those mothers work fewer than 40 hours a week—in other words, part time. Only 5 percent work 50 or more hours weekly. Women leave the workplace to strike out on their own at equally telling rates; the number of businesses owned or co-owned by women jumped 11 percent since 1997, nearly twice the rate of businesses in general.

Look at how all these numbers compare with those of men. Of white men with

M.B.A.'s, 95 percent are working full time, but for white women with M.B.A.'s, that number drops to 67 percent. (Interestingly, the numbers for African-American women are closer to those for white men than to those for white women.)

And look at the women of this Atlanta book club. A roomful of Princeton women each trained as well as any man. Of the 10 members, half are not working at all; one is in business with her husband; one works part time; two freelance; and the only one with a full-time job has no children.

Social scientists—most of them women—have made a specialty in recent years of studying why all this is so. Joan Williams ("Unbending Gender"), Sylvia Ann Hewlett ("Creating a Life"), Arlie Russell Hochschild (who coined the phrase "the second shift") and Felice N. Schwartz (who made popular the phrase "the mommy track"), to name just a few, have done important work about how the workplace has failed women.

But to talk to the women of the book club—or to the women of a San Francisco mothers' group with whom I also spent time, or the dozens of other women I interviewed, or the countless women I have come to know during the four years I have reported on the intersection of life and work—is to sense that something more is happening here. It's not just that the workplace has failed women. It is also that women are rejecting the workplace. . . .

As these women look up at the "top," they are increasingly deciding that they don't want to do what it takes to get there. Women today have the equal right to make the same bargain that men have made for centuries—to take time from their family in pursuit of success. Instead, women are redefining success. And in doing so, they are redefining work.

Time was when a woman's definition of success was said to be her apple-pie recipe. Or her husband's promotion. Or her well-turned-out children. Next, being successful required becoming a man. Remember those awful padded-shoulder suits and floppy ties?

Success was about the male definition of money and power.

There is nothing wrong with money or power. But they come at a high price. And lately when women talk about success they use words like satisfaction, balance and sanity.

That's why a recent survey by the research firm Catalyst found that 26 percent of women at the cusp of the most senior levels of management don't want the promotion. And it's why *Fortune* magazine found that of the 108 women who have appeared on its list of the top 50 most powerful women over the years, at least 20 have chosen to leave their high-powered jobs, most voluntarily, for lives that are less intense and more fulfilling.

It's why President Bush's adviser Karen Hughes left the White House, saying her family was homesick and wanted to go back to Austin. It's why Brenda C. Barnes, who was the president and C.E.O. of Pepsi-Cola North America, left that job to move back to Illinois with her family. And it's why Wendy Chamberlin, who was ambassador to Pakistan, resigned, because security concerns meant she never saw her two young daughters.

Why don't women run the world?

Maybe it's because they don't want to.

Attitudes cluster in place and time. This is particularly true of a college campus, where one-quarter of the student population turns over every year. Undergraduates tend to think that the school they find is the one that always was, with no knowledge of the worldview of those even a few short years before. Looked at that way, the women of the Atlanta book club are a panoramic snapshot of change.

Sally Sears, the oldest of the group, entered Princeton in the fall of 1971. Women had been fully admitted two years earlier, and the school was still very much a boys club. Sears had gone to a small public school in Alabama and entered college "very conscious of being a representative of women and a representative of the South." As she describes it the air was electric with feminism. "Margaret Mead came to talk one night, and I was stunned by how penetrat-

ing her questions were about what it was like to be the first women," she says. "I thought, my God, she's thinking of us as Samoans."

Upon graduation in 1975, Sears felt both entitled and obligated to make good. "The clear message was, 'You've been given the key to a kingdom that used to be denied to people like you,'" she says. "It never occurred to me that my choices would be proscribed. I could have anything I wanted."

What she wanted, at first, was to be "a confirmed single person, childless, a world traveler." She spent a couple of years running *The Childersburg Star*, a small Alabama newspaper owned by her family, and then, in 1978, she took a job on the air at a television station in Birmingham. That led to a job in Memphis, followed by a yearlong trip around the world, then another TV job in Dallas. By 1984 she was on the air in Atlanta, where she became a local celebrity and where she met Richard Belcher, a fellow reporter and now a local anchor. They were married in 1988, when Sears was 35.

Three years later, their son, Will, was born. Soldiers of feminism take only the shortest of maternity leaves, and as soon as Sears recovered from her C-section she was back at work. The O.J. Simpson trial was the first real test of what she calls "work plus love plus a child," because both she and her husband were sent out to California for the duration. "I got my mom and dad to bring Will out, and we all camped out at the New Otani Hotel for a few weeks," she says. "I was determined not to blink." . . .

Sears took nine years to quit. And she did so with great regret. "I would have hung in there, except the days kept getting longer and longer," she explains. "My five-day 50-hour week was becoming a 60-hour week." As news reports could be transmitted farther and farther from the "mother ship," she found herself an hour or two from home when the nightly news was done. "Will was growing up, and I was driving home from a fire," she says. "I knew there would always be wrecks and fires, but there wouldn't always be his childhood."

First she tried to reduce her schedule. “The station would not give me a part-time contract,” she says. “They said it was all or nothing.” So in August 2000, she walked away from her six-figure income and became a homeroom mom at her son’s school.

“It was wrenching for me to leave Channel 2,” she says. “I miss being the lioness in the newsroom—to walk through and have the interns say, ‘There she goes.’ It kills me that I’m not contributing to my 401(k) anymore. I do feel somehow that I let the cause down.” . . .

Sarah McArthur Amsbary of the Atlanta group leads a much-examined life. Back in college, she says, she gave no thought to melding life and work, but now, “I think about it almost constantly.”

And what she has concluded, after all this thinking, is that the exodus of professional women from the workplace isn’t really about motherhood at all. It is really about work. “There’s a misconception that it’s mostly a pull toward motherhood and her precious baby that drives a woman to quit her job, or apparently, her entire career,” she says. “Not that the precious baby doesn’t magnetize many of us. Mine certainly did. As often as not, though, a woman would have loved to maintain some version of a career, but that job wasn’t cutting it anymore. Among women I know, quitting is driven as much from the job-dissatisfaction side as from the pull-to-motherhood side.”

She compares all this to a romance gone sour. “Timing one’s quitting to coincide with a baby is like timing a breakup to coincide with graduation,” she says. “It’s just a whole lot easier than breaking up in the middle of senior year.”

That is the gift biology gives women, she says. It provides pauses, in the form of pregnancy and childbirth, that men do not have. And as the workplace becomes more stressful and all-consuming, the exit door is more attractive. “Women get to look around every few years and say, ‘Is this still what I want to be doing?’” she says. “Maybe they have higher

standards for job satisfaction because there is always the option of being their child’s primary caregiver. When a man gets that dissatisfied with his job, he has to stick it out.”

This, I would argue, is why the workplace needs women. Not just because they are 50 percent of the talent pool, but for the very fact that they are more willing to leave than men. That, in turn, makes employers work harder to keep them. It is why the accounting firm Deloitte & Touche has more than doubled the number of employees on flexible work schedules over the past decade and more than quintupled the number of female partners and directors (to 567, from 97) in the same period. It is why I.B.M. employees can request up to 156 weeks of job-protected family time off. It is why Hamot Medical Center in Erie, Pa., hired a husband and wife to fill one neonatology job, with a shared salary and shared health insurance, then let them decide who stays home and who comes to the hospital on any given day. It is why, everywhere you look, workers are doing their work in untraditional ways.

Women started this conversation about life and work—a conversation that is slowly coming to include men. Sanity, balance and a new definition of success, it seems, just might be contagious. And instead of women being forced to act like men, men are being freed to act like women. Because women are willing to leave, men are more willing to leave, too—the number of married men who are full-time caregivers to their children has increased 18 percent. Because women are willing to leave, 46 percent of the employees taking parental leave at Ernst & Young last year were men.

Looked at that way, this is not the failure of a revolution, but the start of a new one. It is about a door opened but a crack by women that could usher in a new environment for us all.

Why don’t women run the world?

“In a way,” Amsbary says, “we really do.”

79 ■ Heather Boushey

Is the Opt-Out Hypothesis Convincing?

A front page article in the *New York Times* (Story 2005) recently reported that women at Yale University are planning to quit their jobs when they enter their thirties and have children. Both *Time* (Wallis 2004) and the *New York Times Magazine* (Belkin 2003) have recently featured cover articles arguing that some mothers—especially older, highly educated new mothers—are increasingly likely to stop working when they have kids. These articles reflect the popular notion that women are increasingly “opting out” of employment when they have children. The basic argument is that mothers are choosing to stay at home in greater numbers due to the stresses of living in two-earner families or making it in the labor market.

Such news coverage may lead people to believe that there is a growing trend toward this sort of “opt-out” by highly educated mothers. However, economic data provides no evidence to support these anecdotal accounts. In 2004, the latest year for which a full year of data is available, the impact of having children in the home on women’s labor force participation (the “child penalty”) fell compared to prior years.

This analysis of the Current Population Survey’s Outgoing Rotation Group data, a Bureau of Labor Statistics nationally representative survey, shows that the child penalty on labor force participation for prime-age women, aged 25 to 44, averaged –14.4 percentage points over the period from 1984 to 2004. This means that labor force participa-

tion by women in this age group with children at home averaged 14.4 percentage points less than for women without children at home. The penalty was 20.7 percentage points in 1984 and has fallen consistently over the last two decades, down to 8.2 percentage points in 2004.

In terms of the current labor market situation, the data shows that, after controlling for changes in demographics and the labor market, the negative effects of children on women’s labor supply fell between 2000 and 2004. In 2004, prime-age women with children at home were 8.2 percentage points less likely to be in the labor force than were women without children, but this was down from a 9.9 percentage-point penalty in 2000.

The data stand in opposition to the media frenzy on this topic. In spite of the personal anecdotes highlighted in various news stories, women are not increasingly dropping out of the labor force because of their kids. The main reasons for declining labor force participation rates among women over the last four years appears to be the weakness of the labor market.

The recession of the early 2000s led to sustained job losses for all women—those with and without children at home—and by early 2005 the labor market had only just returned to its 2000 employment level, almost exactly four years after the recession began. During this recession, women experienced their largest employment losses in decades and, once this is controlled for, the presence of

children at home plays a smaller role in women's labor force participation than it did in previous years, going back to 1984.

The effect of children on women's labor force participation

This analysis addresses whether or not a woman with a child in the home is any less likely to be in the labor force today than she was at earlier points in the last two decades simply because there was a child in her household. To examine this question, this analysis uses data from the Center for Economic and Policy Research (CEPR) Outgoing Rotation Group (ORG) Extracts for years from 1984 to 2004.¹ The analysis examines the effects of children on the labor force participation rates of prime-age women, aged 25 to 44, regardless of their biological relationship to the child, including un-adopted stepchildren and foster children, along with biological children. Thus, the paper will refer to “mothers” as “women with children at home.”

The focus is on women's labor force participation rates (LFPR), which is the share of a population either at work (employed) or actively searching for work (unemployed). To the extent that journalists and other commentators on the child penalty have used nationally representative data to support their claims, they have used “raw” (or “unadjusted”) LFPR rates, which compare the LFPR rates over time without controlling for the changing demographic characteristics of women or the cyclical nature of the overall labor market. For example, over the past two decades, the share of women who identify themselves as Hispanic, not black, has more than doubled. The rise in population of Hispanic women could pull down mothers' LFPRs because Hispanic mothers typically are much more likely than other mothers to be full-time homemakers. Thus, this demographic shift alone could, if the magnitudes were large enough, explain changes in

women's LFPRs over time. This analysis seeks to isolate the effects of children on women's LFPRs, independent of any changes in the composition of the population of women or cyclical changes in the labor market. Therefore, the focus will be on LFPRs that “control” for factors, such as race and ethnicity, age, education, and the business cycle.

Several demographic and labor market changes over the past two decades may have influenced women's and mother's LFPRs. Immigration, changes in educational attainment among women, increased divorce rates, an aging labor force, and increases in mothers' age at first birth are all factors that may affect a woman's decision to work. Table 1 shows that over the two decades from 1984 to 2004, the average age among women in the 25 to 44 age range increased about 1.5 years, while the share of women in that same age range with children at home decreased (down about 0.5 percentage points for women with children under six; and over 3 percentage points for women with children under 18). Overall, women in all age ranges are also now better educated, though few (8.6 percent) have a graduate degree.

Changes in the composition of the population of women took place against normal cyclical changes in the U.S. economy. Between 2000 and 2004, for example, the labor market went from a peak through a prolonged period of slow economic recovery. As of October 2005, the employment rate—the share of the total population at work—remained 1.8 percentage points below its cyclical peak in 2000. The lackluster labor market and demographic changes have both affected women's LFPRs, leading to the illusion in the raw data that children have caused women to drop out of the labor market.

The tables and figures that follow show the raw LFPR, not adjusted for demographics or other factors, alongside estimates of the specific effects of children on women's LFPR, controlling for characteristics of the mother and the business cycle. The probability that a

Table 1. Characteristics of women aged 25 to 44

| | Year | |
|----------------------------|------|------|
| | 1984 | 2004 |
| <i>Average age</i> | 34.0 | 35.5 |
| Has children under age six | 30.2 | 29.7 |
| Has any children | 66.9 | 63.1 |
| Married | 73.8 | 62.6 |
| Race/ethnicity | | |
| White | 78.0 | 64.8 |
| Black | 12.0 | 13.4 |
| Hispanic | 6.7 | 15.1 |
| Other | 6.3 | 6.8 |
| Educational attainment | | |
| Less than high-school | 14.7 | 9.8 |
| High-school | 40.5 | 30.1 |
| Some college | 23.3 | 29.3 |
| College degree | 15.8 | 22.2 |
| Graduate degree | 5.6 | 8.6 |

Source: Author's analysis of CEPR Outgoing Rotation Group Extracts, various years.

prime-age woman is in the labor force if she has a child is estimated using a standard statistical procedure (logit model), which controls for the effects of age, race and ethnicity, educational attainment, marital status, whether there is a prime-age male in the household in the labor force, and the year (to capture the effects of the business cycle on all women, those with and without children at home).² This model provides a separate estimate for each year of the “marginal effect” of having a child on women’s likelihood of being in the labor force. The paper will refer to this effect, which is the percentage point change in the probability of being in the labor force due specifically to having a child in each year, as the “child penalty.”

Women’s labor force participation rates have not fallen due to the presence of children at home

Figure 1 shows the unadjusted LFPRs for all women, and women with and without children at home. All three groups experienced a

sharp drop in the unadjusted LFPR between 2000 and 2004. The percentage-point decline was actually slightly larger for women without children—2.2 percentage points—than it was for women with children—1.6 percentage points. The drop in the LFPR for women with children has led to claims that mothers are opting out of employment. Figure 1, however, does not isolate the effects of children on women’s LFPR; it only says that women who have children are less likely to be in the labor force.

Figure 2a shows the unadjusted LFPR for all women alongside estimates of women’s LFPR, including the “child penalty” (calculated using the logit model that controls for demographic characteristics and the business cycle). The model that estimates the child penalty by controlling for demographics only (age, education, marital status, and a man working in the household) continues to show a decline in LFPR between 2000 and 2004. However, the model that estimates the child penalty including demographics and the year (as a proxy for the business cycle) shows no

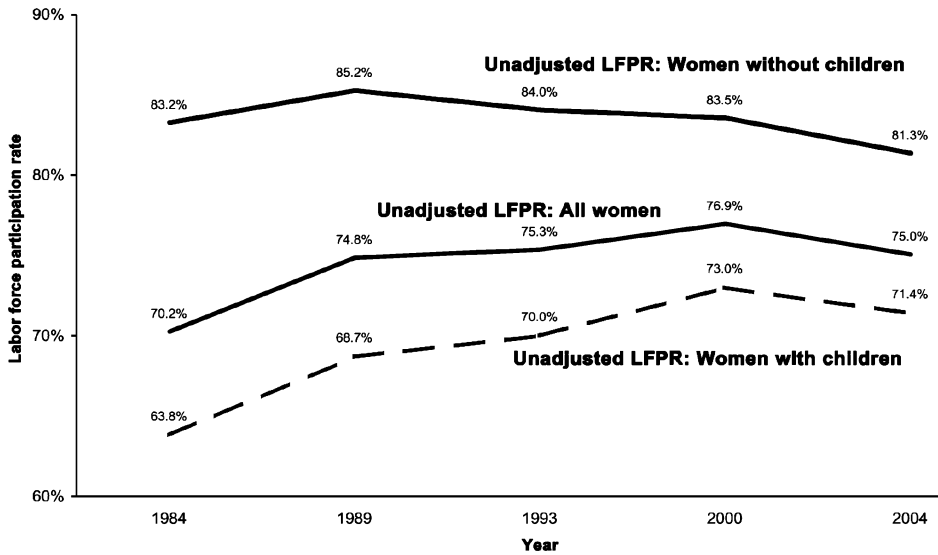


Figure 1. Women's unadjusted labor force participation rate (LFPR), aged 25 to 44, by presence of children at home, 1984 to 2004.

Source: Author's analysis of CEPR Outgoing Rotation Group Extracts, various years.

increase in women opting out of the labor force due specifically to the presence of children in the home.

Figure 2b presents the same information as 2a, but highlights the statistically adjusted effects of children on women's LFPR. Between 2000 and 2004, after controlling for changes in the composition of the population of women and the general decline in the labor market, the child penalty actually fell from 9.9 to 8.2 percentage points (statistically significant at the one percent level). All women have seen their labor supply fall between 2000 and 2004, which is something that economists need to explain. However, the claim that this decline is explained by the decision of women to stay home with their children is simply not true. Children had no more impact—if anything a smaller impact—on women's decisions to join the labor force in 2004 than they did at any earlier point in the preceding 20 years.

The trends are similar for women with young children (under age six) at home (see Figure 3). As would be expected, fewer women

with young children are in the labor force than is the case for women with any children under age 18. The effect of young children on women's participation decision, however, was smaller in 2004 than in 2000 (statistically significant at the one percent level), or any other period examined since 1984.

Women's labor force participation rates have fallen due to the early 2000s recession

Because the raw data do not take into account demographic and critical business cycle effects, the raw data mask the fact that women's LFPRs have been pushed downwards by the prolonged period of slow labor market recovery. The recession of the early 2000s was harder on women than the recessions of the 1980s or 1990s, and in particular, harder on younger women (Boushey, Rosnick, and Baker 2005). The effect of higher job losses on women's labor force participation can be seen in Table 2, which shows the relative effects of the business cycle on prime-age and younger

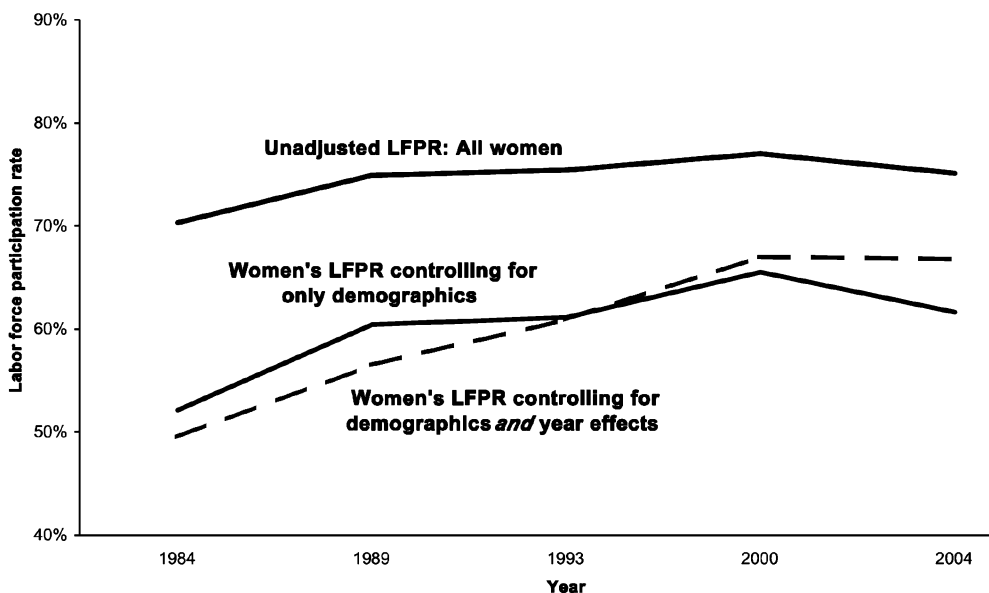


Figure 2a. Labor force participation rates (LFPR), adjusted and unadjusted, women, aged 25 to 44, with any children at home, 1984 to 2004

Source: Author's analysis of CEPR Outgoing Rotation Group Extracts, various years.

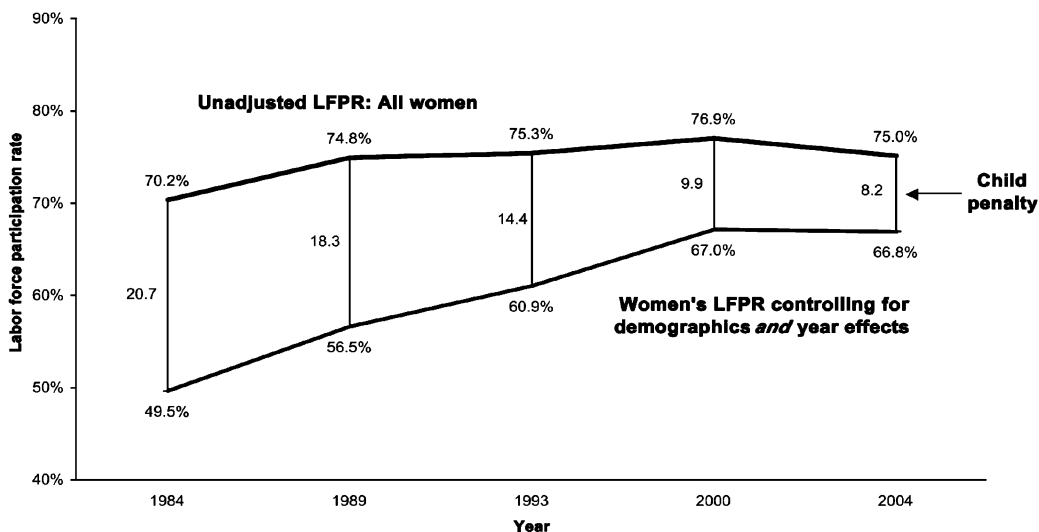


Figure 2b. Labor force participation rates (LFPR), adjusted and unadjusted, women, aged 25 to 44, with any children at home, 1984 to 2004, including the "child penalty" by year

Source: Author's analysis of CEPR Outgoing Rotation Group Extracts, various years.

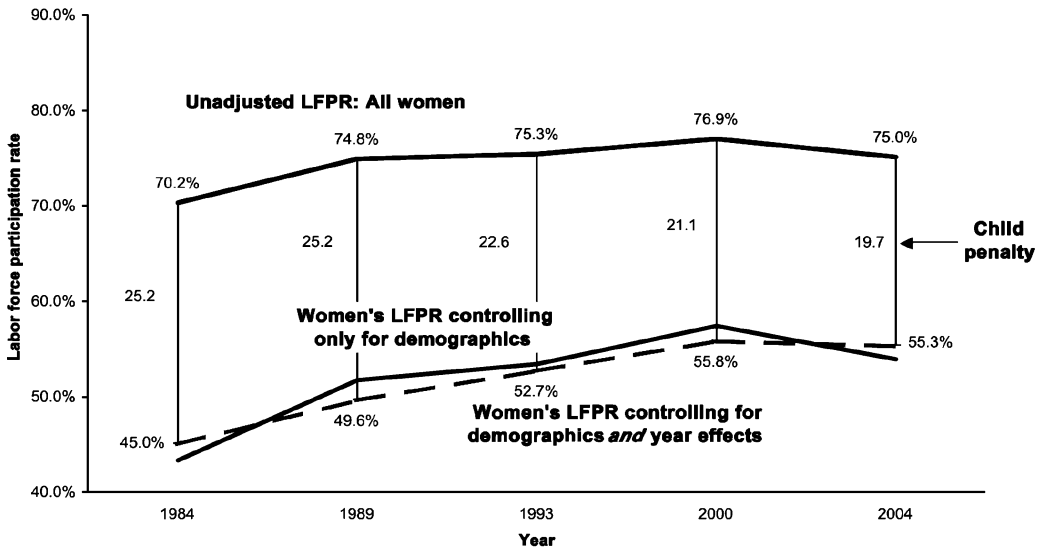


Figure 3. Labor force participation rates (LFPR), adjusted and unadjusted, women, aged 25 to 44, with young children at home, 1984 to 2004, including “child penalty” by year

Source: Author’s analysis of CEPR Outgoing Rotation Group Extracts, various years.

women’s labor force participation rates, controlling for women’s demographic characteristics. Compared to 1984, prime-age women were 3.2 percentage points more likely to be employed in 2000, but only 1.1 percentage points more likely to be employed in 2004. Thus, women were overall less likely to be in the labor force in 2004, compared to 2000 (3.2 percent versus 1.1 percent).

For younger women, the early 2000s recession had a stronger dampening effect on labor force participation. Compared to 1984, in 2004, women aged 25 to 32 were no more or less likely to be in the labor force. In 2000, women were 1.0 percentage points more likely to be in the labor force in 2000 compared to 1984, but only 1.0 percentage points more likely to be in the labor force in 2004.

Higher job losses in the recession of the early 2000s have had the effect of making it appear that women—and especially women with children—are opting out of employment. Table 3 shows the effects of having children in the home on women’s labor supply, with and without controls for the year.

Column 1 shows that the overall effect of having a child under age 18 in the home on prime-age women is to reduce the probability of being in the labor force by 14.4 percentage points. Column 2 looks at the child penalty by year, including only demographics, but not a separate business cycle effect. Here, we see that the child penalty decreases from 1984 to 2000, from 18.2 percentage points down to 11.5 percentage points. Between 2000 and 2004, however, the child penalty rises back up to 13.5 percent. Column 3 shows that there is no such decline once the model includes a year effect, to control for the business cycle, as well as cultural and other changes affecting all women (not just mothers) that may have occurred over time. The year effects shown in Column 3 are the effects of the business cycle of women’s LFPR: in 2004, compared to 1984, women were 7.4 percentage points less likely to work because of the weak labor market. This dampening effect on women’s LFPRs is nearly as large as the 8.2 percentage point child penalty in 2004. . . .

Table 2. Year effects on women's labor force participation rates (Standard errors in parentheses)

| | Women's labor force participation, controlling for demographics (percentage point) | | |
|--------------------------------|---|----------------------------|----------------------------|
| | <i>Prime-age women (aged 25 to 44)</i> | <i>Women aged 25 to 32</i> | <i>Women aged 33 to 39</i> |
| Year effects, relative to 1984 | | | |
| 1989 | 3.2 (0.3)* | 2.3 (0.5)* | 2.9 (0.5)* |
| 1993 | 2.5 (0.3)* | 0.6 (0.5) | 3.1 (0.5)* |
| 2000 | 3.6 (0.4)* | 3.0 (0.6)* | 3.2 (0.6) |
| 2004 | 1.1 (0.4)* | -0.0 (0.6) | 1.0 (0.6)^ |

Source: Author's analysis of CEPR Outgoing Rotation Group Extracts, various years.

* significant at the 1 percent level; ^ significant at the 10 percent level.

Notes: Marginal effects are derived from a logit regression on women's labor force participation. Other independent variables in the model are dummies for age, educational attainment, race/ethnicity, and whether there is a prime-age male in the labor force in the household. Full results available from the author.

Table 3. Child penalty on women's labor force participation rate (Standard errors in parentheses)

| | <i>Prime-age women (aged 25 to 44)</i> | | |
|--------------------------------|--|-----------------|-----------------|
| | (1) | (2) | (3) |
| Has child at home | -14.4 (0.3)* | | |
| Has child at home, by year | | | |
| 1984 | | -18.2 (0.4)* | -20.7 (0.6)* |
| 1989 | | -14.5 (0.4)* | -18.3 (0.6)* |
| 1993 | | -14.3 (0.4)* | -14.4 (0.6)* |
| 2000 | | -11.5 (0.4)* | -9.9 (0.6)* |
| 2004 | | -13.5 (0.4)* | -8.2 (0.6)* |
| Year effects, relative to 1984 | | | |
| 1989 | | | 1.2 (0.6)^ |
| 1993 | | | -2.1 (0.6)* |
| 2000 | | | -3.8 (0.7)* |
| 2004 | | | -7.4 (0.7)* |

Source: Author's analysis of CEPR Outgoing Rotation Groups Extracts, various years.

* significant at the 1 percent level; ^ significant at the 10 percent level.

See notes to Table 2.

Conclusion

The important trend that this paper has explored is that the lackluster performance of the labor market since 2001 is the real reason that LFPRs have been falling among women. Women's employment rates were hit exceptionally hard by this recession and they have yet to fully recover. While women had previously been more insulated from cyclical unemployment, compared to men, now they appear to be nearly as vulnerable, although it remains the case that men's employment rates fell further than women's over the past few years. Future analysis should focus on demand-side factors, rather than assuming that most women either want to or are able to choose to stay at home.

NOTES

1. The CEPR ORG Extracts are publicly available and fully documented at www.ceprdata.org. These Extracts are from the Bureau of Labor Statistics Current population Survey.

2. Full model results are available from the author upon request.

REFERENCES

- Belkin, Lisa. 2003. "The Opt-Out Revolution," *New York Times Magazine*, October 26.
- Boushey, Heather, David Rosnick, and Dean Baker. 2005. *Gender Bias in the Current Economic Recovery? Declining Employment Rates for Women in the 21st Century*. Washington, DC: Center for Economic and Policy Research. Briefing Paper.
- Story, Louise. 2005. "Many Women at Elite Colleges Set Career Path to Motherhood," *New York Times*, September 20.
- Wallis, Claudia. 2004. "The Case for Staying Home," *Time*, May 10, p. 44.

Discrimination

80 ■ Claudia Goldin and Cecilia Rouse

Orchestrating Impartiality

The Impact of “Blind” Auditions on Female Musicians

Sex-biased hiring has been alleged for many occupations but is extremely difficult to prove. The empirical literature on discrimination, deriving from the seminal contributions of Gary Becker (1971) and Kenneth Arrow (1973), has focused mainly on disparities in earnings between groups (e.g., males and females), given differences in observable productivity-altering characteristics. With the exception of various audit studies (e.g., Genevieve Kenney and Douglas A. Wissoker, 1994; David Neumark et al., 1996) and others, few researchers have been able to address directly the issue of bias in hiring practices. A change in the way symphony orchestras recruit musicians provides an unusual way to test for sex-biased hiring.

Until recently, the great symphony orchestras in the United States consisted of members who were largely handpicked by the music director. Although virtually all had auditioned for the position, most of the contenders would have been the (male) students of a select group of teachers. In an attempt to overcome this seeming bias in the hiring of musicians, most major U.S. orchestras changed their audition policies in the 1970s and 1980s, making them more open and routinized. Openings became widely advertised in the union papers, and many positions attracted more than 100 appli-

cants where fewer than 20 would have been considered before. Audition committees were restructured to consist of members of the orchestra, not just the conductor and section principal. The audition procedure became democratized at a time when many other institutions in America did as well.

But democratization did not guarantee impartiality, because favorites could still be identified by sight and through resumes. Another set of procedures was adopted to ensure, or at least give the impression of, impartiality. These procedures involve hiding the identity of the player from the jury. Although they take several forms, we use the terms “blind” and “screen” to describe the group. The question we pose is whether the hiring process became more impartial through the use of blind auditions. Because we are able to identify sex, but no other characteristics, for a large sample, we focus on the impact of the screen on the employment of women.

Screens were not adopted by all orchestras at once. Among the major orchestras, one still does not have any blind round to their audition procedure (Cleveland) and one adopted the screen in 1952 for the preliminary round (Boston Symphony Orchestra), decades before the others. Most other orchestras shifted

to blind preliminaries from the early 1970s to the late 1980s. The variation in screen adoption at various rounds in the audition process allows us to assess its use as a treatment.

The change in audition procedures with the adoption of the screen allows us to test whether bias exists in its absence. In both our study and studies using audits, the issue is whether sex (or race or ethnicity), apart from objective criteria (e.g., the sound of a musical performance, the content of a resume), is considered in the hiring process. Why sex might make a difference is another matter. . . .

Sex Composition of Orchestras

Symphony orchestras consist of about 100 musicians and, although the number has varied between 90 and 105, it is rarely lower or higher. The positions, moreover, are nearly identical between orchestras and over time. As opposed to firms, symphony orchestras do not vary much in size and have virtually identical numbers and types of jobs. Thus we can easily look at the proportion of women in an orchestra without being concerned about changes in the composition of occupations and the number of workers. An increase in the number of women from, say, 1 to 10 cannot arise because the number of harpists (a female-dominated instrument) has greatly expanded. It must be because the proportion female within many groups has increased.

Among the five highest-ranked orchestras in the nation (known as the “Big Five”)—the Boston Symphony Orchestra (BSO), the Chicago Symphony Orchestra, the Cleveland Symphony Orchestra, the New York Philharmonic (NYPhil), and the Philadelphia Orchestra—none contained more than 12 percent women until about 1980.¹ As can be seen in Figure 1A, each of the five lines (giving the proportion female) greatly increases after some point. For the NYPhil, the line steeply ascends in the early 1970s. For the BSO, the turning point appears to be a bit earlier. The percentage female in the NYPhil is currently

35 percent, the highest among all 11 orchestras in our sample after being the lowest (generally at zero) for decades. Thus the increase of women in the nation’s finest orchestras has been extraordinary. The increase is even more remarkable because, as we discuss below, turnover in these orchestras is exceedingly low. The proportion of new players who were women must have been, and indeed was, exceedingly high.

Similar trends can be discerned for four other orchestras—the Los Angeles Symphony Orchestra (LA), the San Francisco Philharmonic (SF), the Detroit Symphony Orchestra, and the Pittsburgh Symphony Orchestra (PSO)—given in Figure 1B.² The upward trend in the proportion female is also obvious in Figure 1B, although initial levels are higher than in Figure 1A. There is somewhat more choppiness to the graph, particularly during the 1940s. Although we have tried to eliminate all substitute, temporary, and guest musicians, especially during World War II and the Korean War, this was not always possible.

The only way to increase the proportion women is to hire more female musicians and turnover during most periods was low. The number of new hires is graphed in Figure 2 for five orchestras. Because “new hires” is a volatile construct, we use a centered five-year moving average. In most years after the late 1950s, the top-ranked orchestras in the group (Chicago and NYPhil) hired about four musicians a year, whereas the other three hired about six. Prior to 1960 the numbers are extremely high for LA and the PSO, because, it has been claimed, their music directors exercised their power to terminate, at will, the employment of musicians. Also of interest is that the number of new hires trends down, even excluding years prior to 1960. The important points to take from Figure 2 are that the number of new hires was small after 1960 and that it declined over time.

The proportion female among the new hires must have been sizable to increase the

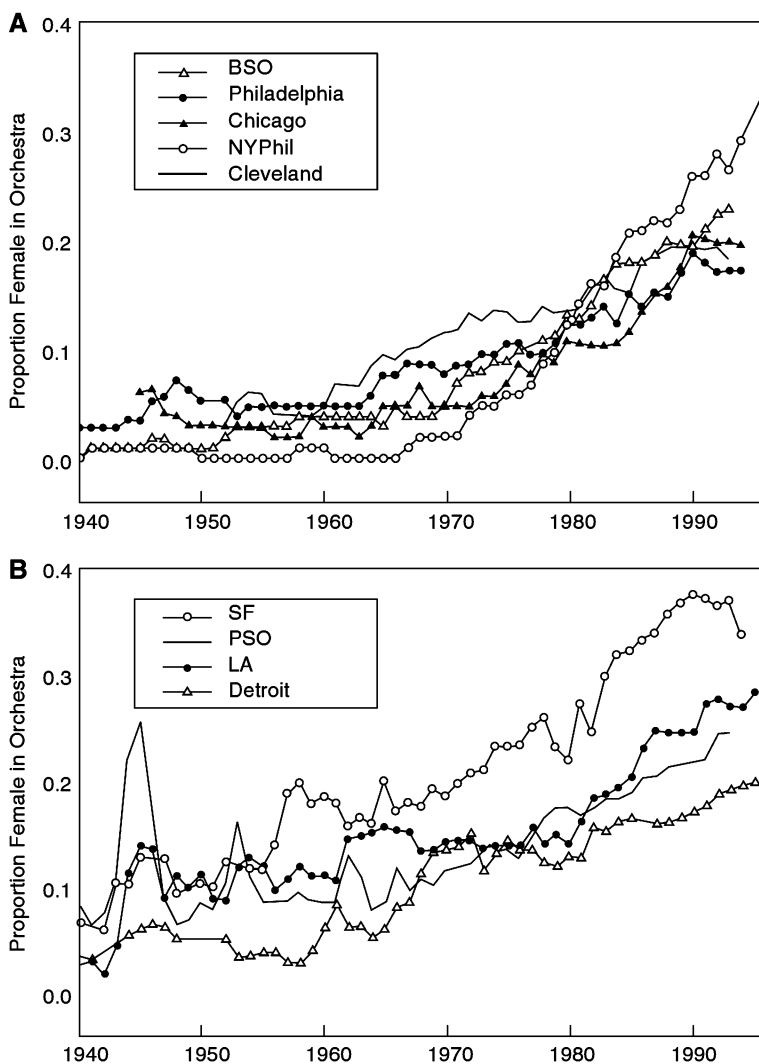


Figure 1. Proportion Female in Nine Orchestras, 1940 to 1990s. A: The “Big Five”; B: Four Others

Source: Roster sample. See text.

proportion female in the orchestras. Figure 3 shows the trend in the share of women among new hires for four of the “Big Five” (Figure 3A) and four other orchestras (Figure 3B).³ In both groups the female share of new hires rose over time, at a somewhat steeper rate for the more prestigious orchestras. Since the early 1980s the share female among new hires has been about 35 percent for the BSO and Chicago, and about 50 percent for the NYPhil, whereas before

1970 less than 10 percent of new hires were women.

Even though the fraction of new hires who are female rises at somewhat different times across the orchestras, there is a discernible increase for the group as a whole from the late 1970s to early 1980s, a time when the labor force participation of women increased generally and when their participation in various professions greatly expanded. The question, therefore, is whether the screen mattered in a

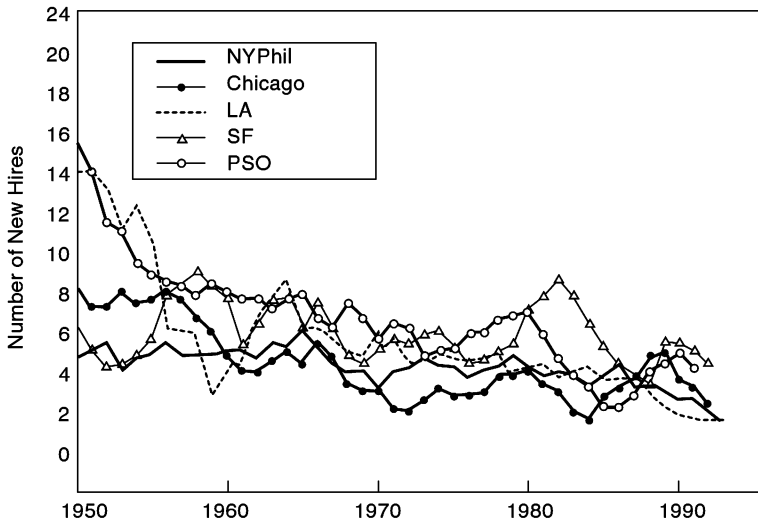


Figure 2. Number of New Hires in Five Orchestras, 1950 to 1990s

Source: Roster Sample. See text.

direct manner or whether the increase was the result of a host of other factors, including the appearance of impartiality or an increased pool of female contestants coming out of music schools. Because the majority of new hires are in their late twenties and early thirties, the question is whether the most selective music schools were producing considerably more female students in the early 1970s. We currently have information by instrument for only the Juilliard School of Music. With the exception of the brass section, the data, given in Figure 4, do not reveal any sharp breaks in the fraction of all graduates who are female. Thus, it is not immediately obvious that an expansion in the supply of qualified female musicians explains the marked increase in female symphony orchestra members; it could, therefore, be because of changes in the hiring procedures of orchestras.

But why would changes in audition procedures alter the sex mix of those hired? Many of the most renowned conductors have, at one time or another, asserted that female musicians are not the equal of male musicians. Claims abound in the world of music that “women have smaller techniques than

men,” “are more temperamental and more likely to demand special attention or treatment,” and that “the more women [in an orchestra], the poorer the sound.”⁴ Zubin Mehta, conductor of the Los Angeles Symphony from 1964 to 1978 and of the New York Philharmonic from 1978 to 1990, is credited with saying, “I just don’t think women should be in an orchestra.”⁵ Many European orchestras had, and some continue to have, stated policies not to hire women. The Vienna Philharmonic has only recently admitted its first female member (a harpist). Female musicians, it can be convincingly argued, have historically faced considerable discrimination. Thus a blind hiring procedure, such as a screen that conceals the identity of the musician auditioning, could eliminate the possibility of discrimination and increase the number of women in orchestras.

Orchestral Auditions

To understand the impact of the democratization of the audition procedure and the screen, we must first explain how orchestra auditions are now conducted. After determining that an audition must be held to fill

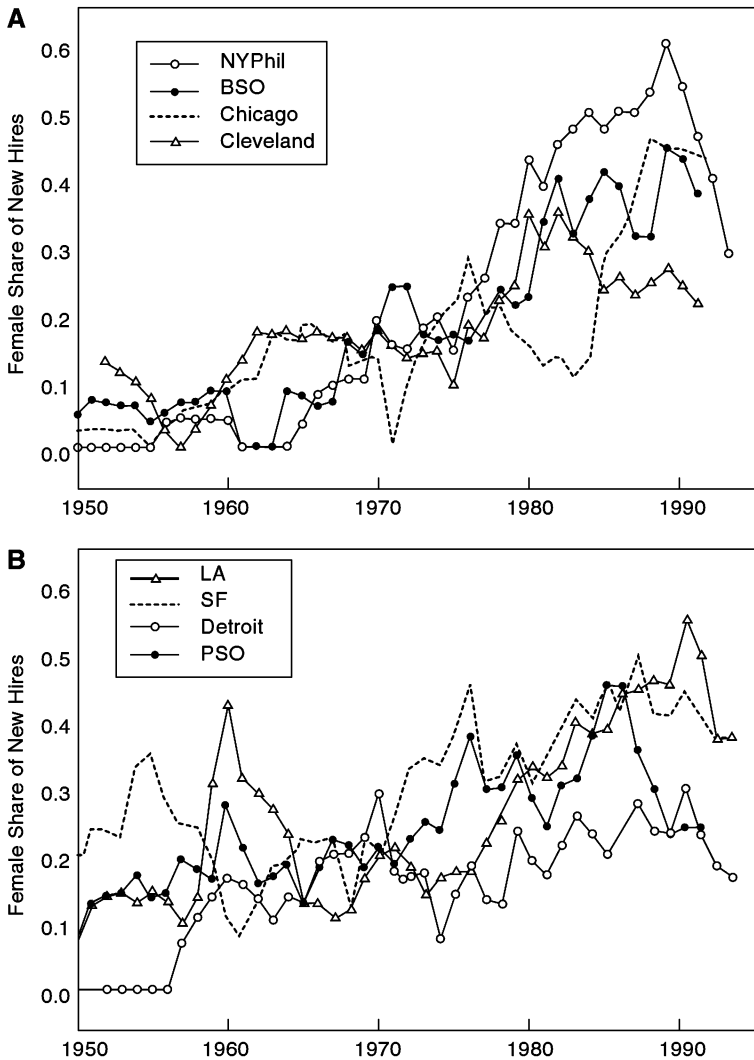


Figure 3. Female Share of New Hires in Eight Orchestras, 1950 to 1990s.
A: Four of the “Big Five”; B: Four Others

Source: Roster sample. See text.

Note: A five-year centered moving average is used. New hires are musicians who were not with the orchestra the previous year, who remain for at least one additional year and who were not substitute musicians in the current year.

an opening, the orchestra advertises that it will hold an audition. Each audition attracts musicians from across the country and, often, from around the world. Musicians interested in auditioning are required to submit a resume and often a tape of compulsory music (recorded according to specific guidelines) to be judged by members of the orchestra. In some orchestras this prescreening

is dispositive; in others the musician has the right to audition live in the preliminary round, even if the audition committee rejects the candidate on the basis of the tape. All candidates are given, in advance, most of the music they are expected to perform at the live audition.

Live auditions today generally consist of three rounds: preliminary, semifinal, and

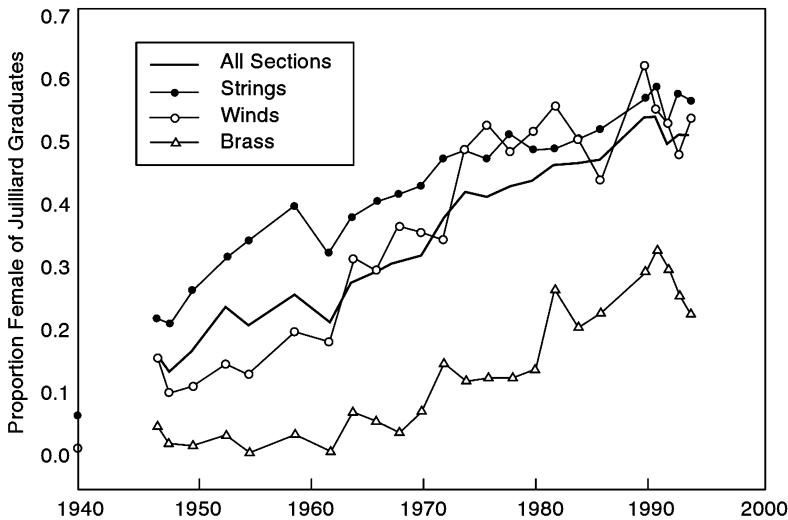


Figure 4. Proportion Female of Juilliard Graduates, Total and by Section: 1947 to 1995.

Source: Juilliard Music School files

final. But there is considerable variation. Although all orchestras now have a preliminary round, some have two final rounds and in many there was no semifinal round until the 1980s. The preliminary is generally considered a screening round to eliminate unqualified candidates. As a result, the committee is free to advance as many, or as few, as they wish. Candidates advanced from the semifinal round are generally considered “acceptable for hire” by the audition committee (which does not include the music director, a.k.a. conductor, until the finals). Again, this means that the committee can advance as many as it wishes. The final round generally results in a hire, but sometimes does not.

In blind auditions (or audition rounds) a screen is used to hide the identity of the player from the committee. The screens we have seen are either large pieces of heavy (but sound-porous) cloth, sometimes suspended from the ceiling of the symphony hall, or look like large room dividers. Some orchestras also roll out a carpet leading to center stage to muffle footsteps that could betray the sex of the candidate. Each candidate for a blind audition is given a number, and the

jury rates the candidate’s performance next to the number on a sheet of paper. Only the personnel manager knows the mapping from number to name and from name to other personal information. The names of the candidates are not revealed to the juries until after the last blind round.

Almost all preliminary rounds are now blind. The semifinal round, added as the number of applicants grew, may be blind. Finals are rarely blind and almost always involve the attendance and input of the music director. Although the music director still wields considerable power, the self-governance that swept orchestras in the 1970s has served to contain the conductor’s authoritarianism. The music director can ignore the audition committee’s advice, but does so at greater peril. Once an applicant is chosen to be a member of an orchestra, lifetime tenure is awarded after a brief probationary period. The basis for termination is limited and rarely used. The positions we are analyzing are choice jobs in the musical world. In 1995 the *minimum starting* base salary for musicians at the BSO was \$1,400 per week (for a 52-week year), not including recording con-

Table 1. Orchestra Audition Procedure Summary Table

| <i>Orchestra</i> | <i>Preliminaries</i> | <i>Semifinals</i> | <i>Finals</i> |
|------------------|---|---|---|
| A | Blind since 1973 | Blind (varies) since 1973 | Not blind |
| B | Blind since at least 1967 | Use of screen varies | Blind 1967–1969; since winter 1994 |
| C | Blind since at least 1979 (definitely after 1972) | Not blind: 1991–present Blind: 1984–1987 | Not blind |
| D | Blind since 1986 | Blind since 1986; varies until 1993 | 1st part blind since 1993; 2nd part not blind |
| E | Use of screen varies until 1981 | Use of screen varies | Not blind |
| F | Blind since at least 1972 | Blind since at least 1972 | Blind since at least 1972 |
| G | Blind since 1986 | Use of screen varies | Not blind |
| H | Blind since 1970 | Not blind | Not blind |
| I | Blind since 1979 | Blind since 1979 | Blind since fall 1983 |
| J | Blind since 1952 | Blind since 1952 | Not blind |
| K | Not blind | Not blind | Not blind |

Notes: The eleven orchestras (A through K) are those in the roster sample described in the text. A subset of eight form the audition sample (also described in the text). All orchestras in the sample are major big-city U.S. symphony orchestras and include the “Big Five.”

Sources: Orchestra union contracts (from orchestra personnel managers and libraries), personal conversations with orchestra personnel managers, and our mail survey of current orchestra members who were hired during the probable period of screen adoption.

tracts, soloist fees, overtime and extra service payments, bonuses, and per diem payments for tours and Tanglewood. . . .

The audition procedures of the 11 orchestras in the roster sample are summarized in Table 1.⁶ Although audition procedures are now part of union contracts, that was not the case in the more distant past and the procedures were not apparently recorded in any surviving documents. We gathered information on these procedures from various sources, including union contracts, interviews with personnel managers, archival documents on auditions, and a mail survey we conducted of orchestral musicians concerning the procedures employed during the audition that won them their current position.

An obvious question to ask is whether the adoption of the screen is endogenous. Of particular concern is that more meritocratic orchestras adopted blind auditions earlier, producing the spurious result that the screen

increased the likelihood that women were hired.⁷ We estimate a probit model of screen adoption by year, conditional on an orchestra’s not previously having adopted the screen (an orchestra exits the exercise once it adopts the screen). Two time-varying covariates are included to assess commonly held notions about screen adoption: the proportion female (lagged) in the orchestra, and a measure of tenure (lagged) of then-current orchestra members. Tenure is included because personnel managers maintain the screen was advocated more by younger players.

As the proportion female in an orchestra increases, so does the likelihood of screen adoption in the preliminary round, as can be seen in columns (1) and (2) in Table 2, although the effects are very small and far from statistically significant. We estimate a similar effect when we assess the role of female presence on the adoption of blind finals [see column (3)]. The impact of current tenure,

Table 2. Estimated Probit Models for the Use of a Screen

| | <i>Preliminaries blind</i> | | <i>Finals blind</i> |
|--|-------------------------------|-------------------------------|-------------------------------|
| | (1) | (2) | (3) |
| (Proportion female) _{<i>t</i>-1} | 2.744 (3.265) [0.006] | 3.120 (3.271) [0.004] | 0.490 (1.163) [0.011] |
| (Proportion of orchestra personnel with <6 years tenure) _{<i>t</i>-1} | -26.46 (7.314) [-0.058] | -28.13 (8.459) [-0.039] | -9.467 (2.787) [-0.207] |
| “Big Five” orchestra | | 0.367 (0.452) [0.001] | |
| pseudo <i>R</i> ² | 0.178 | 0.193 | 0.050 |
| Number of observations | 294 | 294 | 434 |

Notes: The dependent variable is 1 if the orchestra adopts a screen, 0 otherwise. Huber standard errors (with orchestra random effects) are in parentheses. All specifications include a constant. Changes in probabilities are in brackets. “Proportion female” refers to the entire orchestra. “Tenure” refers to years of employment in the current orchestra. “Big Five” includes Boston, Chicago, Cleveland, New York Philharmonic, and Philadelphia. The data begin in 1947 and an orchestra exits the exercise once it adopts the screen. The unit of observation is an orchestra-year.

Source: Eleven-orchestra roster sample. See text.

measured by the proportion with less than six years with the orchestra, is—contrary to general belief—negative and the results do not change controlling for whether the orchestra is one of the “Big Five.” In all, it appears that orchestra sex composition had little influence on screen adoption, although the stability of the personnel may have increased its likelihood.

The Role of Blind Auditions on the Audition and Hiring Process

A. Data and Methods

Audition Records. We use the actual audition records of eight major symphony orchestras obtained from orchestra personnel managers and the orchestra archives. The records are highly confidential and occasionally contain remarks (including those of the conductor) about musicians currently with the orchestra. To preserve the full confidentiality of the records, we have not revealed the names of the orchestras in our sample.

Although availability differs, taken together we obtained information on auditions dating from the late 1950s through 1995. Typically, the records are lists of the names of individuals who attended the auditions, with notation near the names of those advanced to the next round. For the preliminary round, this would indicate advancement to either the semifinal or final round. Another list would contain the names of the semifinalists or finalists with an indication of who won the audition. From these records, we recorded the instrument and position (e.g., section, principal, substitute) for which the audition was held. We also know whether the individual had an “automatic” placement in a semifinal or final round. Automatic placement occurs when a musician is already known to be above some quality cutoff and is invited to compete in a semifinal or final round. We also recorded whether the individual was advanced to the next round of the current audition.

We rely on the first name of the musicians to determine sex. For most names establish-

ing sex was straightforward. Sexing the Japanese and Korean names was equally straightforward, at least for our Japanese and Korean consultants. For more difficult cases, we checked the names in three baby books (Connie Lockhard Ellefson, 1990; Alfred J. Kolatch, 1990; Bruce Lansky, 1995). If the name was listed as male- or female-only, we considered the sex known. The gender-neutral names (e.g., Chris, Leslie, and Pat) and some Chinese names (for which sex is indeterminate in the absence of Chinese characters) remained ambiguous. Using these procedures, we were able to determine the sex of 96 percent of our audition sample.

In constructing our analysis sample, we exclude incomplete auditions, those in which there were no women (or only women) competing, rounds from which no one was advanced, and the second final round, if one exists, for which the candidates played with the orchestra. In addition, we generally consider each round of the audition separately. These sample restrictions exclude 294 rounds (199 contained no women) and 1,539 individuals. Our final analysis sample has 7,065 individuals and 588 audition rounds (from 309 separate auditions) resulting in 14,121 person-rounds and an average of 2.0 rounds per musician. . . .

Roster Data. Our second source of information comes from the final results of the audition process, the orchestra personnel rosters. We collected these data from the personnel page of concert programs, one each year for eleven major symphony orchestras. These records are in the public domain and thus we have used the orchestra names in the graphs containing those data alone. As opposed to the auditionees, we were able to confirm the sex of the players with the orchestra personnel managers and archivists. We considered a musician to be new to the orchestra in question if he or she had not previously been a regular member of that or-

chestra (i.e., we did not count returning members as new). We excluded, when possible, temporary and substitute musicians, as well as harpists and pianists. Our final sample for 1970 to 1996 has 1,128 new orchestra members. . . .

B. The Effect of the Screen on the Hiring of Women

Using the Audition Sample. We turn now to the effect of the screen on the actual hire and estimate the likelihood an individual is hired out of the initial audition pool.⁸ Whereas the use of the screen for each audition round was, more or less, an unambiguous concept, that for the entire process is not and we must define a blind audition. The definition we have chosen is that a blind audition contains all rounds that use the screen. In using this definition, we compare auditions that are completely blind with those that do not use the screen at all or use it for the early rounds only. We divide the sample into auditions that have a semifinal round and those that do not, because the earlier analysis suggested they might differ (Goldin and Rouse 2000).

The impact of completely blind auditions on the likelihood of a woman's being hired is given in Table 3, for which all results include individual fixed effects.⁹ The impact of the screen is positive and large in magnitude, but only when there is no semifinal round. Women are about 5 percentage points more likely to be hired than are men in a completely blind audition, although the effect is not statistically significant. The effect is nil, however, when there is a semifinal round. The impact for all rounds [columns (5) and (6)] is about 1 percentage point, although the standard errors are large and thus the effect is not statistically significant. Given that the probability of winning an audition is less than 3 percent, we would need more data than we currently have to estimate a statistically significant effect, and

Table 3. Linear Probability Estimates of the Effect of Blind Auditions on the Likelihood of Being Hired with Individual Fixed Effects

| | <i>Without semifinals</i> | | <i>With semifinals</i> | | <i>All</i> | |
|------------------------------------|---------------------------|------------------|------------------------|-------------------|------------------|------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Completely blind audition | -0.024 (0.028) | 0.047 (0.041) | 0.001 (0.009) | 0.006 (0.011) | 0.001 (0.008) | 0.005 (0.009) |
| Completely blind audition × female | 0.051 (0.046) | 0.036 (0.048) | 0.001 (0.016) | -0.004 (0.016) | 0.011 (0.013) | 0.006 (0.013) |
| Year effects? | No | Yes | No | Yes | No | Yes |
| Other covariates? | No | Yes | No | Yes | No | Yes |
| R ² | 0.855 | 0.868 | 0.692 | 0.707 | 0.678 | 0.691 |
| Number of observations | 4,108 | 4,108 | 5,883 | 5,883 | 9,991 | 9,991 |

Notes: The unit of observation is a person-round. The dependent variable is 1 if the individual is advanced (or hired) from the final round and 0 if not. Standard errors are in parentheses. All specifications include individual fixed effects, whether the sex is missing, and an interaction for sex being missing and a completely blind audition. “Other covariates” are the size of the audition, the proportion female at the audition, the number of individuals advanced (hired), whether a “Big Five” orchestra, the number of previous auditions, and whether the individual had an automatic semifinal or final.

Source: Eight-orchestra audition sample. See text.

even a 1-percentage-point increase is large, as we later demonstrate.

Using the Roster Data. The roster data afford us another way to evaluate the effect of the screen on the sex composition of orchestras. Using the rosters we know the sex of new hires each year for 11 orchestras, and we also have information (see Table 1) on the year the screen was adopted by each orchestra. We treat the orchestra position as the unit of observation and ask whether the screen affects the sex of the individual who fills the position. We model the likelihood that a female is hired in a particular year as a function of whether the orchestra’s audition procedure involved a screen, again relying on the variation over time within a particular orchestra. Thus, in all specifications, we include orchestra fixed effects and an orchestra-specific time trend.

The roster data extend further back in time than do the audition data and could

conceivably begin with the orchestra’s founding, although there is no obvious reason to include many years when none used the screen. We report, in Table 4, the effects of the screen on the hiring of women from 1970 to 1996 using a probit model. The screen is first defined to include any blind auditions [column (1)]. In column (2) we estimate separate effects for orchestras using blind preliminary (and semifinal) rounds but not blind finals and those with completely blind auditions.

To interpret the probit coefficient, we first predict a base probability, under the assumption that each orchestra does not use a screen. We then predict a new probability assuming the orchestra uses a screen. The mean difference in the probabilities is given in brackets.

The coefficient on blind in column (1) is positive, although not significant at any usual level of confidence. The estimates in column (2) are positive and equally large in

Table 4. Probit Estimates of the Effect of Blind Auditions on the Sex of New Members: 1970 to 1996

| | <i>Any blind auditions</i> | <i>Only blind preliminaries and/or semifinals vs. completely blind auditions</i> |
|---|-------------------------------|--|
| | (1) | (2) |
| Any blind auditions | 0.238 (0.183) [0.075] | |
| Only blind preliminaries and/or semifinals | | 0.232 (0.184) [0.074] |
| Completely blind auditions | | 0.361 (0.438) [0.127] |
| Section: | | |
| Woodwinds | -0.187 (0.114) [-0.058] | -0.188 (0.114) [-0.058] |
| Brass | -1.239 (0.157) [-0.284] | -1.237 (0.157) [-0.284] |
| Percussion | -1.162 (0.305) [-0.235] | -1.164 (0.305) [-0.235] |
| <i>p</i> -value of test: only blind preliminaries and/or semifinals = completely blind | | 0.756 |
| pseudo R ² | 0.106 | 0.106 |
| Number of observations | 1,128 | 1,128 |

Notes: The dependent variable is 1 if the individual is female and 0 if male. Standard errors are in parentheses. All specifications include orchestra fixed effects and orchestra-specific time trends. Changes in probabilities are in brackets; see text for an explanation of how they are calculated. New members are those who enter the orchestra for the first time. Returning members are not considered new. The omitted section is strings.

Source: Eleven-orchestra roster sample. See text.

magnitude to those in column (1). Further, these estimates show that the existence of any blind round makes a difference and that a completely blind process has a somewhat larger effect (albeit with a large standard error). According to the point estimates in column (1) of Table 4, blind auditions increase the likelihood a female will be hired by 7.5 percentage points. The magnitude of the effect must be judged relative to the overall average and, for the period under consideration, it was about 30 percent. Thus blind auditions increased the likelihood a female would be hired by 25 percent.¹⁰ . . .

Conclusion

The question is whether hard evidence can support an impact of discrimination on hiring. Our analysis of the audition and roster data indicates that it can, although we mention various caveats before we summarize the reasons. Even though our sample size is large, we identify the coefficients of interest from a much smaller sample. Some of our coefficients of interest, therefore, do not pass standard tests of statistical significance and there is, in addition, one persistent result that goes in the opposite direction. The weight of the evidence, however, is what we

find most persuasive and what we have emphasized. . . .

As in research in economics and other fields on double-blind refereeing (see, e.g., Blank, 1991), the impact of a blind procedure is toward impartiality and the costs to the journal (here to the orchestra) are relatively small. We conclude that the adoption of the screen and blind auditions served to help female musicians in their quest for orchestral positions.

NOTES

1. The data referred to, and used in Figures 1 to 3, are from orchestral rosters, described in more detail below.

2. Our roster sample also includes the Metropolitan Opera Orchestra and the St. Louis Symphony.

3. A centered five-year moving average is also used for this variable.

4. Seltzer (1989), p. 215.

5. Seltzer (1989), p. 215.

6. We identify the orchestras by letter, rather than by name, to preserve confidentiality of the audition sample.

7. Note, however, it is unlikely that the orchestras that sought to hire more women chose to adopt the screen earlier since the best way to increase the number of women in the orchestra is to have not-blind auditions (so that one could be sure to hire more women).

8. The original article (Goldin and Rouse 2000) includes much additional analysis examining the effect of the screen on the likelihood of being advanced through each stage of the audition process.—EDS

9. There are four auditions in which the committee could not choose between two players and therefore asked each to play with the orchestra. We consider both to be winners. The results are not sensitive to this classification. For this analysis we exclude auditions with no women, all women, or no winner; these exclusions do not change the results.

10. In Table 3 we are identified off of individuals who competed in auditions that were completely blind *and* those that were not completely blind (that is, *any* one round could not be blind). The unit of observation is the person-round and there are 92 ful-

filling this criterion for auditions without a semifinal [columns (1) and (2)]; on average these persons competed in 3.6 auditions in this sample. There are 625 person-rounds fulfilling this criterion that included a semifinal [columns (3) and (4)] and on average these persons competed in 3.5 auditions in this sample. Finally, there are 911 person-rounds fulfilling this criterion across all auditions [columns (5) and (6)] and on average these persons competed in 3.5 auditions in this sample. The sample off of which we are identified is larger for all auditions than for the sum of the other two because some individuals auditioned both with and without a semifinal round.

REFERENCES

- Arrow, Kenneth. "The Theory of Discrimination," in Orley Ashenfelter and Albert Rees, eds., *Discrimination in labor markets*. Princeton, NJ: Princeton University Press, 1973, pp. 3–33.
- Becker, Gary. *The economics of discrimination*, 2nd Ed. Chicago: University of Chicago Press, 1971 [orig. pub. 1957].
- Blank, Rebecca. "The Effects of Double-Blind versus Single-Blind Refereeing: Experimental Evidence from the *American Economic Review*." *American Economic Review*, December 1991, 81(5), pp. 1041–67.
- Ellefson, Connie Lockhard. *The melting pot book of baby names*, 2nd Ed. Cincinnati, OH: Better Way Books, 1990.
- Goldin, Claudia and Cecilia Rouse. 2000. "Orchestrating Impartiality: The Impact of 'Blind' Auditions on Female Musicians." *American Economic Review* 90 (September): 715–741.
- Kenney, Genevieve and Wissoker, Douglas A. "An Analysis of the Correlates of Discrimination Facing Young Hispanic Job-Seekers." *American Economic Review*, June 1994, 84(3), pp. 674–83.
- Kolatch, Alfred J. *The Jonathan David dictionary of first names*. New York: Perigee Books, 1990.
- Lansky, Bruce. *35,000+ baby names*. New York: Meadowbrook Press, 1995.
- Neumark, David (with Bank, Roy and Van Nort, Kyle D.). "Sex Discrimination in Restaurant Hiring: An Audit Study." *Quarterly Journal of Economics*, August 1996, 111(3), pp. 915–41.
- Seltzer, George. *Music matters: The performer and the American Federation of Musicians*. Metuchen, NJ: Scarecrow Press, 1989.

81 ■ Shelley J. Correll, Stephen Benard, and In Paik

Getting a Job

Is There a Motherhood Penalty?

Mothers experience disadvantages in the workplace in addition to those commonly associated with gender. Recent studies show employed mothers in the United States suffer a 5 percent per-child wage penalty on average after controlling for the usual human capital and occupational factors that affect wages (Budig and England 2001; Anderson, Binder, and Krause 2003). The pay gap between mothers and nonmothers under age 35 is larger than the pay gap between men and women (Crittenden 2001), and employed mothers now account for most of the “gender gap” in wages (Glass 2004).

The disadvantages are not limited to pay. Describing a consultant as a mother leads evaluators to rate her as less competent than when she is described as not having children (Cuddy, Fiske, and Glick 2004), and visibly pregnant managers are judged as less committed to their jobs, less dependable, less authoritative, but warmer, more emotional, and more irrational than otherwise equal women managers (Halpert, Wilson, and Hickman 1993; Corse 1990). While the pattern is clear, the underlying mechanism remains opaque. Why would being a parent lead to disadvantages in the workplace for women? And why might similar disadvantages not occur for men?

This paper presents a laboratory experiment and an audit study. The laboratory experiment evaluates the hypothesis that the

“motherhood penalty” occurs because cultural understandings of motherhood lead evaluators to, perhaps unconsciously, expect mothers to be less competent and less committed to their jobs (Blair-Loy 2003; Ridgeway and Correll 2004). As a result, we argue, employers will discriminate against mothers when making hiring, promotion, and salary decisions. We do not expect fathers to experience similar workplace disadvantages since being a good father is not seen in our culture as incompatible with being a good worker (Townsend 2002). By having participants rate job applicants, we expect that applicants presented as women with children will be viewed as less competent and less committed to work, will need to present evidence that they are more qualified for the job, will be rated as less promotable, and will be offered lower starting salaries compared with otherwise similar applicants presented as women without children. While the laboratory experiment isolates and examines the mechanism of discrimination, the audit study provides external validity by evaluating whether actual employers discriminate against mothers.

Wage Penalty for Motherhood

Explanations for the motherhood wage penalty generally can be classified as worker explanations, which seek to identify differences in the traits, skills, and behaviors between

mothers and nonmothers, and discrimination explanations, which rely on the differential preference for or treatment of mothers and nonmothers. Empirical evaluations of these explanations have largely focused on the former.

Budig and England (2001) find that interruptions from work, working part-time, and decreased seniority/experience collectively explain no more than about one-third of the motherhood penalty. In addition, “mother-friendly” job characteristics (i.e., differences in the type of jobs chosen) explain very little of the penalty. Similarly, Anderson, Binder, and Krause (2003) find that human capital and occupational and household resource variables (e.g., number of adults in household) collectively account for 24 percent of the total penalty for one child and 44 percent for women with two or more children. As Budig and England (2001) conclude, the remaining wage gap likely arises either because mothers are somehow less productive at work than nonmothers or because employers discriminate against mothers (or some combination of the two processes).

Productivity and Discrimination

To distinguish between discrimination and productivity explanations, ideally one would compare the outcomes of employed mothers and nonmothers who have equal levels of workplace productivity. If differences in pay or promotion rates were found between equally productive mothers and nonmothers, this would suggest that discrimination factors were at work. However, the datasets analyzed in the previous studies lack direct measures of worker productivity. One likely reason for this is that it is inherently problematic to specify what makes someone a productive employee. This difficulty leads to another: unexplained gaps in wages between employed mothers and nonmothers can always be attributed to unmeasured productivity differences between the two groups.

To address these problems, we experimentally held constant the workplace performances and other relevant characteristics of a pair of fictitious job applicants and varied only their parental status. By holding constant workplace-relevant characteristics, differences between the ratings of mothers and nonmothers cannot be attributed to productivity or skill differences. While this design cannot rule out the possibility that productivity differences account for part of the wage penalty found in previous studies, the laboratory study will isolate a potential status-based discrimination mechanism by evaluating whether being a parent disadvantages mothers in the workplace even when no productivity differences exist between them and women without children.

Performance Expectations and Evaluations of Workplace Competence

Status Characteristics Theory

The laboratory study evaluates the theoretical claim that motherhood is a “status characteristic” that, when salient, results in biased evaluations of competence and commitment, a stricter standard for evaluating the workplace performances of mothers, and a bias against mothers in hiring, promotion, and salary decisions. A status characteristic is a categorical distinction among people, such as race or occupational status, that has attached to it widely held cultural beliefs associating greater status worthiness and competence with one category of the distinction over others (Berger et al. 1977).

Theory and empirical research suggest that ability standards are stricter for those with lower performance expectations (Foschi 1989). The logic behind this prediction is that good performances are inconsistent with expectations for lower status actors; therefore when they perform well, their performances are critically scrutinized and judged by a stricter standard compared with higher status

actors. Thus, performances of low status actors—even when “objectively” equal to that of their high status counterparts—are less likely to be judged as demonstrating task ability or competence (for a comparison of status discrimination and economic theories of statistical discrimination, see Correll and Benard 2006).

Motherhood as a Status Characteristic

To understand how motherhood might function as a devalued status characteristic in workplace settings, it is helpful to broaden the conventional usage of “performance expectations.” While researchers typically focus on the anticipated relative *competence* of group members, cultural beliefs about the relative *effort* that social groups exert in task situations can also be the basis for forming differentiated performance expectations. While it is logically difficult to understand why taking on the motherhood role should affect a person’s underlying competence, there is considerable evidence that contemporary cultural beliefs assume that employed mothers are less committed to work than nonmothers and, consequently, put less *effort* into it (Ridgeway and Correll 2004). Motherhood affects perceptions of commitment because contradictory schemas govern conceptions of “family devotion” and “work devotion” (Blair-Loy 2003: 5). A cultural norm that mothers should always be on call for their children coexists in tension with another widely held normative belief in our society that the “ideal worker” be unencumbered by competing demands and “always there” for his or her employer (Acker 1990; Hays 1996; Williams 2001; Blair-Loy 2003). The tension between these two roles occurs at the level of normative cultural assumptions, and not necessarily at the level of mothers’ own commitment to work. Indeed, Bielby and Bielby (1984) found no differences in the workplace commitment of mothers and nonmothers. Instead it is the

perceived tension between these two roles that leads us to suggest that motherhood is a devalued status in workplace settings.

Therefore, we predict that mothers will be rated as less competent, less committed, less suitable for hire and promotion, and deserving of lower starting salaries compared with otherwise equal women who are not mothers. In addition, we expect mothers will be judged by a harsher standard. Since being a good father is not seen as culturally incompatible with being an ideal worker (Townsend 2002), we do not expect that fathers will experience lower workplace evaluations.

The Laboratory Experiment

Paid undergraduate volunteers (84 men and 108 women) rated a pair of equally qualified, same-gender (either male or female), same-race (either African American or white) fictitious job applicants, presented as real, who differed on parental status. Since there were very few significant effects of participant gender or applicant race, we do not discuss these results here.

The Use of Undergraduates

The laboratory setting ensures sufficient control over factors that would interfere with tests of our hypotheses, such as other people in the room to prime other status characteristics, telephones, or other distractions, and it allows us to collect detailed measures to fully test our argument. By necessity we rely on a sample of undergraduates. The theory presented here implies that to the extent that employers share the belief that mothers are less committed to or competent in workplace settings, they too will subtly discriminate against mothers. Qualitative and quantitative research provides some evidence that employers share this belief (Blair-Loy 2003; Crittenden 2001; Kennelly 1999; Cleveland and Berman 1987; Cleveland 1991; Olian and Schwab 1988). The audit study, described below, will provide

more direct evidence regarding the extent to which employers discriminate against mothers.

Procedure

Participants read a description of a company that was purportedly hiring for a mid-level marketing position and examined application materials for two equally qualified applicants who differed on parental status. To increase their task orientation, participants were told that their input would impact actual hiring decisions. They then inspected each applicant's file, containing three items: a short memo, a "fact sheet," and a resume. The memos contained notes purportedly from a company human resources staff member who conducted a screening interview with the applicant. The "fact sheet" summarized relevant information about the potential employee (e.g., college GPA) not presented on the resume. The fact sheets and the resumes established that the candidates were equally productive in their past jobs and had equivalent skills and backgrounds. Prior to the actual experiment, pretesting of the two versions of the materials confirmed that the resumes were perceived to be of equivalent quality.

Experimental Manipulations

The race and gender of applicants were manipulated by altering first names on the applicant files (Bertrand and Mullainathan 2003). Parental status was manipulated on the resume and on the human resources memo. The resume described one applicant as an officer in an elementary school Parent-Teacher Association, and the accompanying memo included the following phrase: "Mother/Father to Tom and Emily. Married to John/Karen." The nonparent was presented as married, but with no mention of children.

Dependent Measures

There are eight dependent measures: two that measure competence and commitment,

two that measure the ability standard participants used to judge the applicants, and four that serve as our key evaluation measures. The competence measure is a weighted average of participants' ratings of the applicants on seven-point scales ranging from "not at all" to "extremely" capable, efficient, skilled, intelligent, independent, self-confident, aggressive, and organized ($\alpha=.85$). The commitment measure comes from a single-item question that asked participants how committed they thought the applicant would be relative to other employees in similar positions at the company.

There are two ability standard items. Participants were asked: (1) what percentile the applicant would need to score on an exam diagnostic of management ability, and (2) "how many days could this applicant be late or leave early per month before you would no longer recommend him/her for management track?" We predict that mothers will be required to score in a *higher* percentile than nonmothers before being considered hireable and will be allowed *fewer* days of being late or leaving early.

There are four evaluation measures. Participants were asked: (1) to recommend a *salary* for each applicant, (2) to estimate the likelihood that an applicant would be subsequently *promoted* if hired, (3) to judge whether the applicant, if hired, should be recommended for a *management-training* course designed for those with strong advancement potential, and (4) to decide if they would recommend each applicant for *hire*. We predict that mothers will be offered lower starting salaries, will be rated as less promotable, will be less likely to be recommended for management, and will be less likely to be recommended for hire than nonmothers.

Laboratory Experiment Results

As predicted, mothers were judged significantly less competent and committed than women without children (see Table 1, left

Table 1. Means or proportions of status, standards and evaluation variables by gender and parental status of applicant. (Standard deviations in parentheses.)

| | <i>Female Applicants Mothers</i> | <i>Female Applicants Non-mothers</i> | <i>Male Applicants Fathers</i> | <i>Male Applicants Non-fathers</i> |
|--|--------------------------------------|--|------------------------------------|--|
| Competence | 5.19 (0.73)** | 5.75 (0.58) | 5.51 (0.68) | 5.44 (0.66) |
| Commitment | 67.0 (19.1)** | 79.2 (15.2) | 78.5** (16.3) | 74.2 (18.6) |
| Days allowed late | 3.16 (1.98)** | 3.73 (2.01) | 3.69** (2.55) | 3.16 (1.85) |
| Percent score required on exam | 72.4 (27.5)** | 67.9 (27.7) | 67.3 (32.7) | 67.1 (33.0) |
| Salary recommended | \$137,000** (21,000) | \$148,000 (25,000) | \$150,000** (23,000) | \$144,000 (20,700) |
| Proportion recommend for management | .691++ | .862 | .936+ | .851 |
| Likelihood of promotion | 2.74 (0.65)** | 3.42 (0.54) | 3.30* (0.62) | 3.11 (0.70) |
| Proportion recommend for hire | .468++ | .840 | .734+ | .617 |

* $p < .1$, test for difference in means between parent and non-parents

** $p < .05$, test for difference in means between parent and non-parents

+ $z < .1$, test for difference in proportion between parents and non-parents

++ $z < .05$, test for difference in proportion between parents and non-parents

Notes: 94 participants rated female applicants and 94 rated male applicants. For this table, the data for male and female subjects are pooled, as are the data by race of applicant.

side) and were held to harsher performance and punctuality standards. They were allowed significantly fewer times of being late to work and needed a significantly higher score on the management exam than nonmothers before being considered hireable. Similarly, the evaluation measures show significant and substantial penalties for motherhood. The recommended starting salary for mothers was \$11,000 (7.4 percent) less than that offered nonmothers, a significant difference. Mothers were also rated significantly less promotable and were less likely to be recommended for management. Finally, while participants recommended 84 percent of female nonmothers for hire, they recommended a significantly lower 47 percent of mothers.

Fathers were not disadvantaged, and in fact were advantaged on some of these measures. Relative to nonfathers, fathers were rated significantly more committed to their jobs, al-

lowed to be late to work significantly more times, and offered significantly higher salaries.

Multivariate Analysis

We now turn to multivariate models to evaluate the motherhood penalty hypothesis by estimating the effects of gender of applicant, parental status, and the interaction of gender of applicant with parental status on each of the eight dependent variables. We refer to the interaction term (gender of applicant x parental status) as the “motherhood penalty interaction.” Applicant race and participant gender are included in all models, and standard errors are clustered by participant ID to take into account the nonindependence of observations that results from asking participants to rate applicants in pairs. Linear regression models are used for the continuous dependent variables. Logistic regression models are estimated for the binary evaluation variables (recommend

Table 2. Estimated regression coefficients for the effects of gender, parental status and race on applicant's perceived competence and commitment. (Robust standard errors in parentheses, clustered by participant ID).

| <i>Independent Variables</i> | <i>Competence</i> | <i>Commitment</i> |
|-------------------------------------|-----------------------|---------------------|
| Parent | 0.089 (0.088) | 5.15 *** (1.73) |
| Female applicant | 0.376 *** (0.104) | 5.68 ** (2.51) |
| African American | -0.038 (0.090) | -2.01 (2.27) |
| Female participant | 0.060 (0.094) | -2.61 (2.26) |
| Motherhood interaction [^] | -0.750 *** (0.132) | -17.3 *** (2.32) |
| Intercept | 5.42 *** (0.100) | 75.8 *** (2.55) |

[^] Parent * Female applicant

* $p < .1$

** $p < .05$

*** $p < .001$

Note: N=188 participants

for management and recommend for hire). Ordered logistic regression, with the proportional odds specification, is used for the ordered categorical evaluation variable, likelihood of promotion. Parental status, gender of applicant, gender of participant, and race of applicant are dummy variables, with parents, females, and African Americans coded as 1.

The estimated regression coefficients are presented in Tables 2–4. For all eight dependent variables, the motherhood penalty interaction is significant and in the predicted direction. This result shows strong support for the main prediction that parental status negatively impacts ratings for female, but not male, applicants.

Confirming our prediction, mothers were viewed as less competent than nonmothers. The significant, negative motherhood penalty interaction indicates that being a parent lowers the competence ratings for women, but not men (see left column, Table 2). Participants also perceived mothers as less committed than other applicants: the motherhood penalty interaction is significant and negative

in the model predicting commitment ratings (see right column, Table 2). The positive and significant main effect for parental status implies that fathers are actually rated as more committed than nonfathers.

Consistent with the status-based discrimination argument, mothers were held to a stricter performance standard (see Table 3). The motherhood interaction is significant and positive in the model predicting the required test score, while the main effects of gender of applicant and parental status are insignificant, showing that participants require mothers (but not fathers) to score higher on a test of management ability than other applicants before considering them for a job. Mothers are also held to a higher standard of punctuality, being allowed fewer days of being late.

In Table 4, the motherhood penalty interaction is significant and negative across all four models, indicating that mothers, relative to other applicants, are believed to deserve lower salaries and to be less suitable for hiring, promoting, and training for management. In the model predicting likelihood of promotion, the

Table 3. Estimated regression coefficients for the effects of gender, parental status and race on ability standard variables. (Robust standard errors in parentheses, clustered by participant ID.)

| <i>Independent Variables</i> | <i>Days allowed late</i> | <i>Test score required (%)</i> |
|-------------------------------------|--------------------------|--------------------------------|
| Parent | 0.515 *** (0.137) | 1.03 (0.968) |
| Female applicant | 0.572 ** (0.294) | 1.25 (4.52) |
| African American | -0.361 (0.294) | -4.06 (4.38) |
| Female participant | 0.234 (0.289) | -9.44 ** (4.30) |
| Motherhood interaction [^] | -1.10 *** (0.213) | 3.56 *** (1.21) |
| Intercept | 3.22 *** (0.322) | 73.7 *** (4.27) |

[^] Parent * Female applicant

* p<.1

** p<.05

*** p<.001

Note: N=188 participants

Table 4. Estimated regression coefficients for the effects of gender, parental status and race on evaluation variables. (Robust standard errors in parentheses, clustered by participant ID.)

| <i>Independent Variables</i> | <i>Promotion likelihood</i> (ordered logistic estimates) | <i>Mgmt training?</i> (binary logistic estimates) | <i>Hire?</i> (binary logistic estimates) | <i>Recommended salary in thousands of dollars</i> (linear estimates) |
|-------------------------------------|---|--|---|---|
| Parent | 1.03 * (0.545) | 0.605 * (0.321) | 0.570 (0.366) | 4.47 *** (1.84) |
| Female applicant | 0.256 (0.425) | 1.009 *** (0.319) | 1.21 *** (0.365) | 2.56 (3.18) |
| African American | 0.309 (0.299) | -0.211 (0.218) | -0.163 (0.197) | -6.80 ** (2.94) |
| Female participant | 0.496 * (0.298) | 0.526 ** (0.226) | 0.606 *** (0.199) | 0.691 (2.82) |
| Motherhood interaction [^] | -2.14 *** (0.651) | -2.72 *** (0.426) | -2.38 *** (0.548) | -15.9 *** (2.42) |
| Intercept | ^{^^} | 4.56 *** (0.601) | 0.210 (0.266) | 148 (2.55) |

[^] Parent * Female applicant

^{^^} Since ordered logistic regression produces multiple intercepts, we do not present them here.

* p<.1

** p<.05

*** p<.001

Note: N=188 participants

main effect of parental status is marginally significant and positive, while the motherhood penalty interaction is significant and negative, indicating that the negative effect of parental status on perceptions of promotability accrues only to women.

Consistent with previous literature on the motherhood wage penalty, mothers are offered lower starting salaries than other types of applicants, as indicated by the significant, negative coefficient for the motherhood interaction term. Childless men were recommended an average salary of approximately \$148,000. Fathers were offered a significantly higher salary of approximately \$152,000. Women without children were offered approximately \$151,000, whereas mothers were recommended a significantly lower salary of about \$139,000, or about 7.9 percent less than otherwise equal childless women.

While the motherhood penalty interaction is significant and its sign is in the predicted direction for each model, to complete our argument, we need to give evidence that motherhood disadvantages job applicants *because* it is a status characteristic. If the theory is correct, then evaluations of competence and commitment should mediate the motherhood penalty.

When the competence and commitment ratings were added as independent variables to the models (see Table 5), the negative effect of motherhood status on workplace evaluations was significantly reduced, by a magnitude of 31–46 percent. Thus, mothers are rated as less hireable, less suitable for promotion and management training, and deserving of lower salaries in part because they are believed to be less competent and less committed to paid work. Having established support for the causal mechanism with the laboratory data, we turn to the audit study to assess whether actual employers discriminate against mothers.

The Audit Study

The audit methodology combines experimental design with real-life settings. As in

laboratory experiments, audit studies isolate a characteristic (e.g., race or gender) and test for discriminatory behavior. Distinct from most laboratory studies, audit study participants are the people who make important decisions about actual applicants, such as employers conducting new employee searches. While laboratory experiments permit closer investigation of social and cognitive processes, audit studies provide greater generalizability of results.

Resumes and cover letters from a pair of fictitious, equally qualified, same-gender applicants were sent to employers advertising for marketing and business job openings in a large, northeastern city newspaper over an 18-month period of time. The same-sex pair contained one parent and one nonparent. Job openings were randomly assigned to either the male or female condition. We manipulated parental status on the resume and on the cover letter. We did not manipulate race in this study.

We monitor whether gender and parental status impact the odds that an employer will call back an applicant. Based on the 5–8 percent callback rate found in an audit study of race in hiring (Bertrand and Mullainathan 2003), and to ensure that we had sufficient statistical power to evaluate the effect of parental status, we submitted 1,276 resumes and cover letters to 638 employers.

Results

The results suggest that real employers do discriminate against mothers (see Table 6). Childless women received 2.1 times as many callbacks as equally qualified mothers. This finding is similar to the laboratory experiment (see Table 1) in which childless women were recommended for hire 1.8 times more frequently than mothers. In the laboratory study, fathers were recommended for hire at a slightly *higher* rate, although the difference was only marginally significant; in the audit study, fathers were called back at a higher rate, although the difference was not significant.

Table 5. Estimated regression coefficients for the mediation of competence and commitment on the impact of parental status on workplace evaluations. (Robust standard errors in parentheses, clustered by participant ID.)

| <i>Independent Variables</i> | <i>Promotion likelihood</i> (ordered logistic estimates) | <i>Mgmt training?</i> (binary logistic estimates) | <i>Hire?</i> (binary logistic estimates) | <i>Recommended salary in thousands of dollars</i> (linear estimates) |
|---|---|--|---|---|
| Competence | 0.628 ** (0.295) | 1.263 *** (0.281) | 1.21 *** (0.258) | 7.00 *** (1.99) |
| Commitment | 0.237 *** (0.095) | 0.206 ** (0.099) | 0.308 *** (0.081) | 1.08 (0.762) |
| Parent | 0.901 * (0.558) | 0.508 (0.340) | 0.433 (0.426) | 3.23 * (1.78) |
| Female applicant | -0.140 (0.426) | 0.661 ** (0.332) | 0.755 * (0.410) | -0.817 (3.31) |
| African American | 0.374 (0.319) | -0.154 (0.237) | -0.092 (0.244) | -6.30 *** (2.86) |
| Female participant | 0.557 * (0.316) | 0.606 *** (0.236) | 0.755 *** (0.254) | 0.512 (2.81) |
| Motherhood interaction [^] | -1.34 ** (0.646) | -1.89 *** (0.437) | -1.39 ** (0.606) | -8.52 *** (2.66) |
| Intercept | ^^ | 3.64 *** (0.947) | -2.09 *** (0.702) | 140 (6.37) |
| Percent reduction of motherhood penalty | 37.4 % | 30.5 % | 41.6 % | 46.4 % |

[^] Parent * Female applicant

^{^^} Since ordered logistic regression produces multiple intercepts, we do not present them here.

* p<.1

** p<.05

*** p<.001

Note: N=188 participants

Table 6. Proportions of applicants receiving callbacks by gender and parental status.

| | <i>Callbacks / Total Jobs</i> | <i>Proportion Called Back</i> |
|-----------------|-------------------------------|-------------------------------|
| Mothers | 10 / 320 | .0313 |
| Childless women | 21 / 320 | .0656++ |
| Fathers | 16 / 318 | .0503 |
| Childless men | 9 / 318 | .0283 |

++ z <.05 Test for difference in proportions between parents and non-parents

Notes: Mothers and childless women applied to the same 320 jobs; fathers and childless men applied to the same 318 jobs.

We now consider a multivariate model for the effects of parental status, applicant gender, and the interaction of parental status and applicant gender on the odds that an applicant receives a callback from an employer. Table 7 shows the motherhood penalty interaction is significant and negative, while the main effect for parental status is insignificant, and the main effect for the female applicant variable is significant and positive. The significant negative motherhood penalty interaction term indicates that being a parent lowers the odds that a woman, but not a man, will receive a callback from employers. In sum, the audit data show that, compared with their equally qualified childless counterparts, mothers are disadvantaged when actual employers make hiring decisions.

Strengths and Limitations

While the audit study evaluates whether actual employers discriminate against mothers in the hiring process, it does not give us insight into the mechanism underlying discrimination, because it was not possible to collect employers' rankings of commitment, competence, performance standards, and other relevant variables. These limits mean that while the audit study establishes that actual employers discriminate against mothers, it cannot establish why.

By considering the results of these two companion studies simultaneously, however, we find support for the status-based discrimination mechanism using the laboratory data and see real world implications of the argument with data generated from the audit study. Further, these results are consistent with qualitative work showing that employers discriminate against mothers (Blair-Loy 2003; Crittenden 2001; Kennelly 1999) and with survey research that consistently finds a wage penalty for motherhood (Budig and England 2001; Anderson, Binder, and Krause 2003). Thus, across a wide range of methodological approaches—each of which has its unique strengths and weaknesses—we find evidence

Table 7. Estimated binary logistic regression coefficients for the effects of parental status and gender on the odds of receiving a callback. (Robust standard errors in parentheses, clustered by job.)

| <i>Independent Variables</i> | <i>Callback?</i> |
|-------------------------------------|--------------------|
| Parent | 0.598 (.433) |
| Female applicant | 0.887** (.407) |
| Motherhood interaction [^] | -1.38** (.590) |
| Intercept | -3.54*** (.338) |

[^] Parent *Female applicant

* p<.10

** p<.05

*** p<.001

Notes: Mothers and childless women applied to the same 320 jobs; fathers and childless men applied to the same 318 jobs, for a total of 1276 applications to 638 jobs.

that mothers experience disadvantages in workplace settings and that discrimination plays a role in producing these disadvantages.

Summary and Conclusions

This project makes two main contributions. First, it isolates and experimentally evaluates a status-based discrimination mechanism that explains some of the disadvantages mothers experience in the paid labor market. Second, it shows that real employers discriminate against mothers. The results of this study have implications for understanding some of the enduring patterns of gender inequality in paid work. Studies have documented the motherhood penalty in at least 15 countries (Harkness and Waldfogel 1999; Misra, Budig, and Moller 2005) and shown its stability over time (Avellar and Smock 2003). This study offers a partial explanation for the mechanism behind a widespread, durable phenomenon with implications for a broad segment of the population.

More generally, a gender gap in wages has persisted despite the vast movement of

women into paid labor in the United States since the early 1970s, and employed mothers account for most of this gap (Glass 2004). This study suggests that cultural beliefs about the tension between the motherhood and “ideal worker” roles may play a part in reproducing this pattern of inequality. A second enduring pattern of gender inequality is the so-called “glass ceiling,” a metaphor for the barriers that restrict women’s movement up the career ladder to the highest positions in organizations and firms. To the extent that employers view mothers as less committed to their jobs and less “promotable,” the glass ceiling women face could be, in part, a motherhood ceiling.

Writing for the National Center for Policy Analysis, Denise Venable (2002) reports that among people ages 27 to 33 who have never had children, women’s earnings approach 98 percent of men’s. She concludes, “When women behave in the workplace as men do, the wage gap between them is small.” Claims of unequal pay, she continues, “almost always involve comparing apples and oranges.” However, since most employed men and employed women have children at some point in their lives, the most illustrative “within fruit” comparison is not the comparison of childless men to childless women, but the comparison of men with children to women with children. As the two studies reported here show, when women “behave as men do,” by giving evidence of being a parent, they are discriminated against, while their male counterparts are often advantaged. Far from being an “apples to oranges” comparison, the male and female applicants who were evaluated in these studies were exactly equal by experimental design. That parental status disadvantaged only female applicants is strong evidence of discrimination.

REFERENCES

- Acker, Joan. 1990. “Hierarchies, jobs, and bodies: a theory of gendered organizations.” *Gender & Society* 4: 139–158.
- Anderson, Deborah J., Melissa Binder, and Kate Krause. 2003. “The motherhood wage penalty revisited: Experience, heterogeneity, work effort and work-schedule flexibility.” *Industrial and Labor Relations Review* 56: 273–294.
- Avellar, Sarah, and Pamela Smock. 2003. “Has the price of motherhood declined over time? A cross-cohort comparison of the motherhood wage penalty.” *Journal of Marriage and the Family* 65: 597–607.
- Berger, Joseph, Hamit Fisek, Robert Norman, and Morris Zelditch, Jr. 1977. *Status Characteristics and Social Interaction*. New York: Elsevier.
- Bertrand, Marianne, and Sendhil Mullainathan. 2003. “Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination.” *National Bureau of Economic Research Working Paper Series*, Working Paper no. 9873. Retrieved September 2003. (<http://www.nber.org/papers/w9873>).
- Bielby, Denise D., and William T. Bielby. 1984. “Work commitment, sex-role attitudes, and women’s employment.” *American Sociological Review* 49: 234–247.
- Blair-Loy, Mary. 2003. *Competing Devotions: Career and Family Among Women Executives*. Cambridge, MA: Harvard University Press.
- Budig, Michelle, and Paula England. 2001. “The wage penalty for motherhood.” *American Sociological Review* 66: 204–225.
- Cleveland, Jeanette N. 1991. “Using hypothetical and actual applicants in assessing person-organization fit: A methodological note.” *Journal of Applied Social Psychology* 21: 1004–1011.
- Cleveland, Jeanette N., and Andrew H. Berman. 1987. “Age perceptions of jobs: Agreement between samples of students and managers.” *Psychological Reports* 61: 565–566.
- Correll, Shelley J., and Stephen Benard. 2006. “Biased estimators? Comparing status and statistical theories of gender discrimination.” Pp. 89–116 in *Social Psychology of the Workplace (Advances in Group Process Volume 23)*, edited by Shane R. Thye and Edward J. Lawler. New York: Elsevier.
- Corse, Sara J. 1990. “Pregnant managers and their subordinates: The effects of gender expectations on hierarchical relationships.” *Journal of Applied Behavioral Science* 26: 25–48.
- Crittenden, Ann. 2001. *The Price of Motherhood: Why the Most Important Job in the World Is Still the Least Valued*. New York: Metropolitan Books.
- Cuddy, Amy J. C., Susan T. Fiske, and Peter Glick. 2004. “When professionals become mothers,

- warmth doesn't cut the ice." *Journal of Social Issues* 60: 701–718.
- Foschi, Martha. 1989. "Status characteristics, standards and attributions." Pp. 58–72 in *Sociological Theories in Progress: New Formulations*, edited by Joseph Berger, Morris Zelditch, Jr., and Bo Anderson. Boston: Houghton Mifflin.
- Glass, Jennifer. 2004. "Blessing or curse? Work-family policies and mother's wage growth over time." *Work and Occupations* 31: 367–394.
- Halpert, Jane A., Midge L. Wilson, and Julia Hickman. 1993. "Pregnancy as a source of bias in performance appraisals." *Journal of Organizational Behavior* 14: 649–663.
- Harkness, Susan, and Jane Waldfogel. 1999. "The family gap in pay: Evidence from seven industrialised countries." CASE paper 29: Centre for Analysis of Social Exclusion.
- Hays, Sharon. 1996. *The Cultural Contradictions of Motherhood*. New Haven, CT: Yale University Press.
- Kennelly, Ivy. 1999. "That single mother element: How white employers typify black women." *Gender & Society* 13: 168–192.
- Misra, Joya, Michelle Budig, and Stephanie Moller. 2005. "Employment, wages, and poverty: Reconciliation policies and gender equity." Paper presented at the Annual Meeting of the American Sociological Association, Philadelphia, PA.
- Olian, Judy D., and Donald P. Schwab. 1988. "The impact of applicant gender compared to qualifications on hiring recommendations: A meta-analysis of experimental studies." *Organizational Behavior and Human Decision Processes* 41: 180–195.
- Ridgeway, Cecilia, and Shelley J. Correll. 2004. "Motherhood as a status characteristic." *Journal of Social Issues* 60: 683–700.
- Townsend, Nicholas W. 2002. *The Package Deal: Marriage, Work and Fatherhood in Men's Lives*. Philadelphia, PA: Temple University Press.
- Venable, Denise. 2002. "The wage gap myth." National Center for Policy Analysis, April 12, 2002, Brief Analysis no. 392. Retrieved December 12, 2004. (<http://www.ncpa.org/pub/ba/ba392/>).
- Waldfogel, Jane. 1997. "The effect of children on women's wages." *American Sociological Review* 62: 209–217.
- Williams, Joan. 2001. *Unbending Gender: Why Work and Family Conflict and What to Do About It*. Oxford: Oxford University Press.

82 ■ Barbara F. Reskin

Rethinking Employment Discrimination and Its Remedies

The paradigmatic definition of discrimination is the differential treatment of persons because of status characteristics that are functionally irrelevant to the outcome in question (Merton 1972; see also Allport 1954, 51). Implicit in this definition is the proposition that individuals discriminate because of their negative feelings (antipathy, distaste, fear) toward or negative beliefs (social stereotypes) about

members of a status group (Allport 1954, 14; Becker 1957). Media accounts of whites' hostility toward blacks prior to and during the civil rights movement lent face validity to an antipathy-based theory of discrimination. Although the success of the civil rights movement depended partly on its ability to expose the antagonism that many whites harbored toward blacks, it simultaneously reinforced the

Barbara Reskin. "Rethinking Employment Discrimination and Its Remedies," in *The New Economic Sociology: Developments in an Emerging Field*, edited by Mauro F. Guillén, Randall Collins, Paula England, and Marshall Meyer, pp. 218–230, 232–233, 237, 239–244.

theory that discrimination resulted from the deliberate efforts of individual whites to harm blacks because of their hostility to them.

In emphasizing differential treatment *based on* or *because of* an irrelevant characteristic, social scientists construed discrimination as *intentional* behavior. So too did the major antidiscrimination law, Title VII of the 1964 Civil Rights Act. Although the statute did not explicitly define discrimination, it conceptualized it as an intended and consciously motivated act (Stryker 2001, 18). Thus, the law operationalized illegal employment discrimination as intentional.

Economists have theorized that employers may also intentionally discriminate against individuals on the rational rather than emotional ground that they belong to a race or sex thought to be less productive or more costly to employ, or that they will otherwise adversely affect business because of customers' or coworkers' prejudices (Phelps 1972; Bergmann and Darity 1981). Economists term this intentional, non-animus-based discrimination "statistical discrimination." The Equal Employment Opportunity Commission (EEOC) and the Supreme Court (*Los Angeles Department of Water and Power v. Manhart* 1978; 435 U.S. 702)—the federal agencies charged with enforcing and interpreting Title VII—have interpreted the law as prohibiting intentional, non-animus-based discrimination, including statistical (stereotype-based) discrimination. The law prohibits employers from *deliberately* using race and sex in personnel decisions, regardless of their motive (except to redress egregious discrimination of the past; see Reskin 1998). Thus, the major U.S. antidiscrimination law bans individuals—whether acting on their own behalf or as agents of their employer—from intentionally discriminating regardless of whether their motives derive from negative sentiments toward a group or from concerns about the economic ramifications of employing members of that group.

There is ample evidence that individuals continue to engage in intentional acts that con-

form to the paradigmatic definition of discrimination (Kirschenman and Neckerman 1991; Kasinitz and Rosenberg 1996; Browne and Kennelly 1999; U.S. Equal Employment Opportunity Commission 2001a-g). However, both the paradigmatic theory of intentional disparate treatment and the theory of statistical discrimination ignore another important source of employment discrimination: automatic, nonconscious cognitive processes.¹ In this chapter, I review research that implies that employment decisions are routinely biased as a result of normal cognitive processes. If nonconscious cognitive processes—unless checked by employment structures—regularly contribute to discrimination, our theoretical approach to discrimination needs to be modified.

In the years since the animus theory of discrimination gained paradigmatic status, social cognition researchers have established that automatic cognitive processes give rise to cognitive errors whose results are likely to be discriminatory. Largely outside individual control, these processes are marked by a cognitive efficiency that makes them generally adaptive. Because of their automaticity, neither intention nor antipathy is a necessary condition for discriminatory outcomes. The role of automatic cognitive processes in discrimination means that we also cannot continue to view discrimination as anomalous acts by a few biased individuals within an otherwise fair system of employment opportunities (Black 1989). Instead, we must recognize that discriminatory outcomes—many of them involving micro acts of discrimination—are pervasive, permitting members of favored groups to accumulate advantages while members of disfavored groups accumulate disadvantages.

Social Cognitive Processes and Employment Discrimination

Categorization

The core cognitive process that links gender and race to workplace outcomes is categorization. We automatically categorize the people

we encounter into ingroups and outgroups, into “we” and “they” (Rothbart and Lewis 1994; Brewer and Brown 1998). In 1908 Georg Simmel wrote that we necessarily think of others in terms of general categories (1971 [1908], 10–12). Almost half a century later Gordon Allport (1954, 19) echoed this insight, asserting that “the human mind must think with the aid of categories.” Recent research supports these claims: we understand the social world in terms of categorical distinctions and readily categorize others into ingroups and outgroups, on the basis of both highly visible as well as arbitrary and trivial characteristics (Brewer 1997, 200; Brewer and Brown 1998, 566; Fiske 1998, 364, 375).

A critical concomitant of categorization is the generalization of similarity and difference. Having categorized others on trait A, we infer both that ingroup members resemble us and that outgroup members differ from us on traits B, C, D, and so on. As a result, we systematically underestimate within-group differences and exaggerate between-group differences on a variety of characteristics (Brown 1995, 78). Categorization also affects how and how quickly we encode and recall information about others, thereby (unbeknownst to us) distorting our perceptions and shaping our preferences (Perdue et al. 1990).

It is important to note that we automatically prefer ingroup members to outgroup members. We are more comfortable with them, feel more obligated and loyal to them, impute to them positive attributes, trust them, remember their positive traits while forgetting their negative ones, are predisposed to cooperate with them, and favor them when distributing rewards (Perdue et al. 1990; Baron and Pfeffer 1994; Rothbart and Lewis 1994, 369; Brown 1995; Brewer 1997, 201; Brewer and Brown 1998, 567; Fiske 1998, 361).

Although we tend to distrust and depersonalize outgroup members and see them as competitors (Brewer 1997, 201), laboratory research indicates that we apparently do not

automatically see them in negative ways (Perdue et al. 1990, 482) or discriminate against them when allocating negative outcomes (Brewer and Brown 1998, 559). Together, this body of research suggests that automatic categorization gives rise to discrimination primarily through pervasive ingroup favoritism rather than pervasive outgroup antipathy.

Automatic categorization is not inevitably group-serving. We categorize others into ingroups and outgroups largely independently of our conscious feelings toward the groups or our desire to protect our own status (Fiske 1998, 364). Nonetheless, ingroup preferences reinforce dominant groups’ privileges. For example, group status affects whether people infer “illusory correlations” between positive traits and group membership. Members of high-status ingroups show more ingroup favoritism than do members of low-status ingroups (Brewer and Brown 1998, 570). Also, ingroup attachment is especially strong among people who think that group membership affects their opportunities or risks, both among competing groups and among ingroups in a numerical minority (Weber 1994).

Group membership is often based on ascribed characteristics. Ingroup favoritism cannot operate unless we can readily distinguish who is “us” and who is “them” (Brewer 1997, 205). Given its cultural salience and socially exaggerated visibility, we habitually use sex in assigning others to ingroups or outgroups (Brewer and Brown 1998). Because sex and race are strongly correlated with control over workplace opportunities, with white men monopolizing these roles, ingroup preference favors whites, men, and especially white men. Thus, ingroup favoritism produces status-group discrimination.

In sum, concomitant to our automatic sorting of people into categories are automatic biases in our feelings toward others based on whether we see them as “us” or “them.” The effects of such biases are discriminatory, although they are neither intended nor even known to us as we perpetrate them (Fiske

1998, 362). In addition, they apparently operate more through our automatic inclination to favor others like ourselves than to deprive or punish those who differ (Brewer and Brown 1998; DiTomaso 2000).

Stereotyping

Categorization leads to a second automatic cognitive process with potentially discriminatory effects: stereotyping. Stereotyping is an inferential process in which we attribute traits that we habitually associate with a group to individuals who belong to that group. Although *conscious* stereotypes can culminate in intentional statistical discrimination (see, for example, Wilson 1996, ch. 5; Moss and Tilly 1996, 264), here I address the *unintentional* discriminatory effects of *non-conscious* stereotypes. Almost invariably, noticing that a “target” belongs to a stereotyped group automatically brings to mind characteristics that are stereotypically linked to that group, even when we consciously reject those stereotypes as false (Bodenhausen, Macrae, and Garst 1998, 316). In fact, we automatically pursue, prefer, and remember “evidence” that supports our stereotypes (including untrue “evidence”) and ignore, discount, and forget facts that challenge them. Stereotypes have been described as “hypotheses in search of evidence” (Brown 1995) because we are predisposed toward behavioral interpretations that conform to our stereotypes, blind to stereotype-inconsistent interpretations, and cognitively better able to find stereotype-consistent than -inconsistent evidence (Fiske 1998, 367). This selectivity means that our nonconscious stereotypes include a confirmatory bias (Brown 1995).

Given our cognitive capacity to override evidence with stereotypes, it stands to reason that we invoke stereotypes when we lack complete information about others. For example, given the stereotype of blacks as lazy (Bobo 2001), an employer who is choosing between a black applicant and a white applicant and has limited information about the black is

likely to assume automatically that the black applicant is lazy. Stereotypes also distort our recollections of others. We more readily remember stereotypical than nonstereotypical descriptions—an attractive flight attendant, for example, is more memorable than an unattractive flight attendant (Brown 1995, 96). As discussed later, stereotypes also bias the attributions we automatically construct for others’ behavior; these attribution errors in turn distort our assessments of others and our expectations of their future behavior.

Automatic stereotypes often prevail even when we are consciously motivated not to stereotype. For instance, experimental subjects asked to judge the heights of women and men from photographs routinely underestimated women’s height relative to same-height men, although they had been instructed to judge each photograph as an individual and they knew that accurate judgments could yield a substantial payoff (Brown 1995, 92). Subjects in another study who were instructed to suppress their stereotypes were able to do so, but in a subsequent task they expressed stronger stereotypes than members of a control group who had not been told to avoid stereotyping (Bodenhausen et al. 1998, 326).

Automatic stereotyping is tenacious because it is cognitively efficient (Fiske 1998, 366; Fiske, Lin, and Neuberg 1999, 237, 244). We can process information that supports our stereotypes more quickly than inconsistent information. Time pressure, information overload, mental busyness, and pressure to make a decision—conditions that characterize the work lives of many decision-makers—increase our likelihood of stereotyping and exacerbate the effects of stereotypes on judgment and memory, presumably because these situations require cognitive efficiency (Bodenhausen et al. 1998: 319; Fiske 1998, 389; Goodwin, Operario, and Fiske 1998, 694). For example, subjects in a sentence-completion task could refrain from making sexist statements when they had

ample time; given time limits, however, their statements were more sexist than those of members of a control group (Bodenhausen et al. 1998, 326).

Any attribute that is stereotypically linked to a status characteristic can activate (or “prime”) group stereotypes (Heilman 1995; Fiske 1998, 366). For example, research subjects who were subliminally exposed to the words “black” and “white” were more likely to link negative traits to blacks and positive traits to whites than those who were not subliminally exposed (Dovidio et al. 1997), and men who were primed with stereotypic statements about women were more likely to ask a female job applicant sexist questions than were nonprimed men, and they took longer to recognize nonsexist words (Fiske et al. 1999). These findings imply that an idle remark can activate sex or race stereotypes that in turn affect how decisionmakers assess job or promotion candidates or other workers (Heilman 1995). Importantly, by priming race or sex stereotypes, diversity training may actually foster distorted impressions and differential treatment based on sex or race.

Status affects the propensity to automatic stereotyping. Subordinates are less likely to stereotype members of dominant groups than the reverse, presumably because they need to make accurate assessments of the people who control their work environment and rewards (Fiske et al. 1999, 241). In addition, dominant-group members appear to have greater confidence in their stereotyped assessments than do members of subordinate groups, probably both because their dominant-group status gives them confidence *and* because being wrong is less likely to have negative consequences.

In sum, social cognition research suggests that unconscious sex and race stereotyping can lead to discrimination by distorting our impressions of individuals based on group membership. This occurs, unless checked, when group-based stereotypes automatically are applied to individuals, replace missing in-

formation about individuals, distort what we recall of others’ behavior, bias our attributions of others’ success and failures, and influence our predictions of others’ future behavior.

Attribution Error

How we expect others to perform affects the meaning we assign to their behavior. When people’s performance conforms to our expectations, we attribute it to their stable, internal traits (for example, ability); when it contradicts our expectations, we attribute it to transient, external causes (for example, task difficulty or luck). Both ingroup versus outgroup membership and sex and race stereotypes influence our expectations about others’ performance. We expect members of socially preferred groups to succeed and members of devalued groups to fail. Thus, we tend to chalk up successes by ingroup members and white men and failures by outgroup members, minorities, and white women to their stable predispositions, while we attribute failures by ingroup members and white men and successes by outgroup members, minorities, and white women to external (and hence, unstable) factors (Crocker, Major, and Steele 1998, 539). For example, we credit a man’s talents for his success in a customarily male job, while we attribute a woman’s success to situational factors, such as help from others. Importantly, we do not see successes or failures that we have attributed to external causes as predicting future success (Swim and Sanna 1996; Brewer and Brown 1998, 560).

Moreover, our brains tend to encode behavior that conforms to our expectations in terms of abstract traits (intelligence, honesty, laziness) and encode unexpected behavior in concrete terms (his car would not start; she took the class twice). Thus, we attribute an ingroup member losing his temper to provocation and the same behavior by an outgroup member to personality. Attributions in abstract terms are more likely than ones in concrete terms to affect our global evaluation of others as well as our predictions of their fu-

ture performance, and they are more resistant to challenge by counterevidence. As a result, these cognitive propensities give ingroup members and persons from socially valued groups the benefit of the doubt, while preventing new data on outgroup members and persons in socially subordinate status groups from dispelling negative attributions.

Once again, while these propensities are universal, perceptions by members of majority groups are more prone to attribution error than those by minority-group members (Brown 1995, 101–2).

Micro-Macro Links

Automatic categorization, ingroup preference, and stereotyping are cognitively efficient for individuals. Their efficiency frees up cognitive resources for other purposes, making these propensities cognitively adaptive to individuals, especially those who are juggling multiple demands (Brown 1995, 95). Given what seems to be the near-impossibility of suppressing these cognitive propensities, the resulting biases are pervasive and predictable. Because they link employment outcomes to functionally irrelevant group membership, they are also discriminatory.

In addition, although many of the daily discriminatory events stemming from cognitive biases such as excluding someone from a project or denying them credit for a success may seem trivial (and certainly not legally actionable), over time members of preferred groups accumulate advantages and members of disparaged groups accumulate disadvantages. Eventually group differences emerge that appear to justify unequal career outcomes (Krieger 1995). Thus, unless employers implement structures to check the biasing effects of these microlevel processes, their long-term consequences create or exacerbate macrolevel disparities across race and sex groups in their economic and social fates.

In view of the automaticity of these cognitive processes, the proximate cause of employment discrimination becomes whether

employers' personnel practices prevent or permit these cognitive propensities from linking workers' status characteristics to their employment outcomes (Bielby 2000; Reskin 2000). In the next section, I discuss employment structures and practices that can reduce their discriminatory impact.

Organizational Practices, Cognitive Biases, and Discriminatory Outcomes

Work organizations cannot stamp out our propensities to categorize others automatically or to treat ingroup members more favorably than outgroup members. As we have seen, attempts to repress stereotyping can backfire. But employment structures can curb the *biasing effects* of these cognitive processes, thus inhibiting the discrimination that would otherwise result (Fiske 1998, 375). The impact of audition practices on women's share of positions in major symphony orchestras illustrates my point. With the adoption in the 1970s and 1980s of blind auditions that physically screened the musician from the auditors, women's odds of being selected rose sharply (Goldin and Rouse 2000). Screens that concealed the sex of the candidates circumvented intentional discrimination and obviated the discriminatory consequences of nonconscious sex stereotyping and ingroup favoritism. Social cognition and organizational research points to employment practices or structures that can arrest—or activate—the biasing effects of cognitive processes. Based on that scholarship, this section proposes organizational practices that mitigate the biasing effects of automatic cognitive processes.

Organizational Practices and Categorization

The automaticity of ingroup preference probably plays an important role in producing disparate employment outcomes for different race or sex groups. One way to minimize its discriminatory effects is to dissociate ingroup membership from ascribed statuses.

As Marilyn Brewer (1997, 205) observes, a precursor of ingroup favoritism is the identification of who is “us” and who is “them.” Employers can take advantage of this fact by discouraging ingroups based on race or sex through *decategorization* and *recategorization*.

Decategorization conveys individuating information about workers to decisionmakers (which in turn reduces stereotyping). Recategorization recognizes that people have many characteristics that compete to serve as the basis of categorization. According to Galen Bodenhausen and her colleagues (1998, 316), “the stereotypes associated with the category that ‘wins’ the competition” are activated, while we suppress the “losing” stereotype. Consider, for example, an experiment in which exposure to a photograph of a Chinese woman activated subjects’ stereotypes about women, while inhibiting their stereotypes about Asians, but exposure to a picture of the same woman eating noodles with chopsticks triggered stereotypes about Asians, while suppressing stereotypes about women (Bodenhausen et al. 1998, 317). In general, recategorization exploits people’s readiness to categorize others on the most trivial criteria by encouraging the categorization of workers on characteristics that are independent of their sex and race (Brewer and Brown 1998, 583). Employers can encourage recategorization based on work-related functions, for example. Function-based categories such as teams, projects, and divisions are cognitively available and exploitable by employers through work-group culture and structured competition. Firms can also emphasize recategorization to a “superordinate category” such as an officer of the law or a UPS employee (Brewer 1997, 202). . . .

Organizational Practices and Stereotyping

Although the propensity to automatic stereotyping is universal, whether we stereotype others varies, and whether our stereotypes give rise to biased cognitions depends on several factors, such as the type and quality of information

that is available to decisionmakers about “targets”; whether the decisionmaking process highlights status-group membership and hence primes stereotyping; decisionmakers’ awareness of potential biases and their motivation to perceive others accurately; and whether decisionmakers have the cognitive resources to consciously minimize stereotype-based distortions (Bodenhausen et al. 1998, 330).

Making available accurate individuating information about persons from negatively stereotyped groups can prevent stereotyping. (But feeling informed in the absence of objective information may promote stereotyping; see Fiske 1998, 387.) When we lack information that we need to make employment decisions, our descriptive stereotypes automatically provide it. Because stereotypes about socially devalued groups tend to be negative and those about socially favored groups to be positive, the effect is usually discriminatory. In addition, because multiple bases for categorizing a person compete for activation, providing relevant information (Is the applicant a college graduate? An experienced manager?) about members of stereotyped groups can prevent their being stereotyped by sex, race, and other ascribed characteristics (Bodenhausen et al. 1998, 316). Thus, employment practices that increase the amount of unambiguous, relevant information available to decisionmakers may suppress stereotyping (Heilman 1995, 12). By compiling and disseminating to decisionmakers standardized information for all candidates on all the criteria relevant to a personnel decision, organizations can reduce stereotyping. . . .

Decisionmakers’ Discretion

Discretion is an important prerequisite for managers. It not only permits autonomy but allows decisionmakers to control the amount of effort they must expend. Although organizational members tend to assume that managers’ decisions are rational (Baron 1984, 56), discretion invites unconscious and conscious biases to influence their decisions, thus spawning discrim-

inatory treatment. Employment decisions that are based on unstructured observations are especially vulnerable to cognitive biases (Fiske et al. 1991). The key to preventing such automatic biases from giving rise to unintended micro acts of discrimination lies in curbing decisionmakers' discretion by requiring specific procedures for distributing opportunities. . . .

Conclusion

In keeping with the paradigmatic conception of discrimination as conscious invidious treatment motivated by antipathy, Congress outlawed intentional discrimination in Title VII of the 1964 Civil Rights Act. Although the enforcement of Title VII has reduced employment discrimination (Burstein 1989; Burstein and Edwards 1994), decades after the enactment of Title VII the careers of thousands of Americans continue to be stunted by employment discrimination (Blumrosen et al. 1998; Reskin 2001). Discrimination persists partly because antidiscrimination remedies misconstrue discrimination as based in antipathy toward a group. Despite the applicability of the paradigmatic construction of discrimination in many instances, that construction ignores an even more important reason individuals' sex and race are routinely and illegitimately linked to employment rewards: automatic nonconscious cognitive processes that distort our perceptions and treatment of others.

Categorization, ingroup preference, stereotyping, attribution error—all part of the normal information processing that occurs largely outside our conscious control—bias our perceptions, evaluations, and treatment of people because of their sex, color, accent, and other discernible characteristics that signal membership in more or less valued groups. They engender biases because we automatically categorize others into ingroups and outgroups—often based on these same characteristics. Moreover, we prefer ingroup members to outgroup members, evaluate them more favorably, treat as

factual impressions that are based on stereotypes, and make erroneous inferences about the successes or failures of ingroup and outgroup members.

These cognitive biases occur during interpersonal relations, and interpersonal relations are an intrinsic part of most jobs. Thus, ordinary, everyday cross-sex and cross-race interactions elicit automatic ingroup favoritism, stereotyping, and attribution biases. These cognitive distortions lead on a daily basis to micro acts of discrimination: omitting a qualified but female inside candidate from the short list for a top position because no one thinks of her; not recognizing the contribution that a Japanese woman makes to a project; assuming that women do not want challenging assignments; not inviting a black man to a meeting of his specialty group; being confused about which black worker made a suggestion; giving a white male a tip on a job opening (Cose 1993; Fiske 1998, 371–72; Catalyst 1999; Barrett 1999; DiTomaso 2000; Reskin and McBrier 2000, 224; Sturm 2001, 113, 136).

At the individual level, these “countless small acts by a changing cast of characters . . . *incrementally* and consistently limit the employment prospects of one group of workers compared with those of another” (Nelson and Bridges 1999, 241–43, emphasis added; see also Feagin 2000, 139; Sturm 2001, 113). As a result, workers who regularly benefit from ingroup preference accumulate advantages, and members of disparaged groups who suffer the negative discriminatory effects of cognitive biases accumulate disadvantages. Over time, career disparities between sex- and race-based groups also accumulate. In the aggregate, the unchecked consequences of automatic cognitive biases help to preserve job segregation, promotion differences, and earnings disparities by race and sex. That they result from nonconscious cognitive processes rather than conscious antipathy does not mitigate their impact or reduce the need for more effective policy remedies. The results are the same.

NOTE

1. According to Shiffrin and Schneider (1977), automatic cognitive processes are those that are not effortful, intentional, or consciously controlled.

REFERENCES

- Allport, Gordon W. 1954. *The Nature of Prejudice*. Reading, Mass.: Addison-Wesley.
- Baron, James N. 1984. "Organizational Perspectives on Stratification." *Annual Review of Sociology* 10: 37–69.
- Baron, James N., and Jeffrey Pfeffer. 1994. "The Social Psychology of Organizations and Inequality." *Social Psychology Quarterly* 57: 190–209.
- Barrett, Paul M. 1999. *The Good Black: A True Story of Race in America*. New York: Dutton.
- Becker, Gary. 1957. *The Economics of Discrimination*. Chicago: University of Chicago Press.
- Bergmann, Barbara R., and William Darity Jr. 1981. "Social Relations, Productivity, and Employer Discrimination." *Monthly Labor Review* 104(April): 47–49.
- Bielby, William T. 2000. "Minimizing Workplace Gender and Racial Bias." *Contemporary Sociology* 29(1): 120–29.
- Black, Donald. 1989. *Sociological Justice*. New York: Oxford University Press.
- Blumrosen, Alfred W., Mark Bendick, J. J. Miller, and Ruth G. Blumrosen. 1998. "Employment Discrimination Against Women in Washington State, 1997." Employment Discrimination Project Report 3. Newark, N.J.: Rutgers University School of Law.
- Bobo, Lawrence D. 2001. "Racial Attitudes and Relations at the Close of the Twentieth Century." In *America Becoming: Racial Trends and Their Consequences*, vol. 1, edited by Neil J. Smelser, William J. Wilson, and Faith N. Mitchell. Washington: National Academy Press.
- Bodenhausen, Galen V., C. Neil Macrae, and Jennifer Garst. 1998. "Stereotypes in Thought and Deed: Social Cognitive Origins of Intergroup Discrimination." In *Intergroup Cognition and Intergroup Behaviors*, edited by Constantine Sedikides, John Schopler, and Chester A. Insko. Mahwah, N.J.: Erlbaum.
- Brewer, Marilynn B. 1997. "The Social Psychology of Intergroup Relations: Can Research Inform Practice?" *Journal of Social Issues* 53(1): 197–211.
- Brewer, Marilynn B., and Rupert J. Brown. 1998. "Intergroup Relations." In *Handbook of Social Psychology*, edited by Daniel T. Gilbert, Susan T. Fiske, and Gardner Lindzey. New York: McGraw-Hill.
- Brown, Rupert. 1995. *Prejudice*. Oxford: Blackwell.
- Browne, Irene, and Ivy Kennelly. 1999. "Stereotypes and Realities: Images of Black Women in the Labor Market." In *Latinas and African American Women at Work*, edited by Irene Browne. New York: Russell Sage Foundation.
- Burstein, Paul. 1989. "Attacking Sex Discrimination in the Labor Market: A Study in Law and Politics." *Social Forces* 67: 641–65.
- Burstein, Paul, and Mark E. Edwards. 1994. "The Impact of Employment Discrimination Litigation on Racial Disparity in Earnings: Evidence and Unresolved Issues." *Law and Society Review* 28: 79–111.
- Catalyst. 1999. "Women of Color Report a 'Concrete Ceiling' Barring Their Advancement in Corporate America." *Catalyst Press Room*, July 13. Available at: www.catalystwomen.org/press/mediakit/release071399woc.html (May 20, 2001).
- Cose, Ellis. 1993. *Rage of the Privileged Class*. New York: HarperCollins.
- Crocker, Jennifer, Brenda Major, and Claude Steele. 1998. "Social Stigma." In *Handbook of Social Psychology*, edited by Daniel T. Gilbert, Susan T. Fiske, and Gardner Lindzey. New York: McGraw-Hill.
- DiTomaso, Nancy. 2000. "Why Antidiscrimination Policies Are Not Enough: The Legacies and Consequences of Affirmative Inclusion—for Whites." Rutgers University, Newark. Unpublished manuscript.
- Dovidio, John F., Kerry Kawakami, Craig Johnson, Brenda Johnson, and Adaiah Howard. 1997. "On the Nature of Prejudice: Automatic and Controlled Processes." *Journal of Experimental Social Psychology* 33: 510–40.
- Feagin, Joe R. 2000. *Racist America: Roots, Current Realities, and Future Reparations*. New York: Routledge.
- Fiske, Susan T. 1998. "Stereotyping, Prejudice, and Discrimination." In *Handbook of Social Psychology*, edited by Daniel T. Gilbert, Susan T. Fiske, and Gardner Lindzey. New York: McGraw-Hill.
- Fiske, Susan T., Donald N. Bersoff, Eugene Borgida, Kay Deaux, and Madeline E. Heilman. 1991. "Social Science Research on Trial: Use of Sex Stereotyping Research in *Price Waterhouse v. Hopkins*." *American Psychologist* 46: 1049–60.
- Fiske, Susan T., Monica Lin, and Steven L. Neuberg. 1999. "The Continuum Model: Ten Years Later." In *Dual Process Theories in Social Psychology*, edited by Shelly Chaiken and Yaacov Trope. New York: Guilford Press.
- Goldin, Claudia, and Cecilia Rouse. 2000. "Orchestrating Impartiality: The Impact of 'Blind' Auditions on Female Musicians." *American Economic Review* 90: 715–41.
- Goodwin, Stephanie A., Don Operario, and Susan T. Fiske. 1998. "Situational Power and Interper-

- sonal Dominance Facilitate Bias and Inequality.” *Journal of Social Issues* 54: 677–98.
- Heilman, Madeline E. 1995. “Sex Stereotypes and Their Effects in the Workplace: What We Know and What We Don’t Know.” *Journal of Social Issues* 10: 3–26.
- Kasinitz, Philip S., and Jan Rosenberg. 1996. “Missing the Connection: Social Isolation and Employment on the Brooklyn Waterfront.” *Social Problems* 43: 180–96.
- Kirschenman, Joleen, and Kathryn M. Neckerman. 1991. “We’d Love to Hire Them But . . .’: The Meaning of Race for Employers.” In *The Urban Underclass*, edited by Christopher Jencks and Paul Peterson. Washington, D.C.: Brookings Institution.
- Krieger, Laura H. 1995. “The Contents of Our Categories: A Cognitive Bias Approach to Discrimination and Equal Employment Opportunity.” *Stanford Law Review* 47: 1161–1248.
- Merton, Robert K. 1972. “Insiders and Outsiders.” *American Journal of Sociology* 78: 9–47.
- Moss, Philip, and Chris Tilly. 1996. “Soft Skills and Race: An Investigation of Black Men’s Employment Problems.” *Work and Occupations* 23: 252–76.
- Nelson, Robert L., and William S. Bridges. 1999. *Legalizing Gender Inequality*. New York: Oxford University Press.
- Perdue, Charles, John F. Dovidio, Michael B. Gutman, and Richard B. Tyler. 1990. “Us and Them: Social Categorization and the Process of Intergroup Bias.” *Journal of Personality and Social Psychology* 59: 475–86.
- Phelps, Edmund S. 1972. “The Statistical Theory of Racism and Sexism.” *American Economic Review* 62: 659–61.
- Reskin, Barbara F. 1998. *The Realities of Affirmative Action*. Washington, D.C.: American Sociological Association.
- . 2000. “The Proximate Causes of Discrimination: Research Agenda for the Twenty-First Century.” *Contemporary Sociology* 29: 319–29.
- . 2001. “Employment Discrimination and Its Remedies.” In *Sourcebook on Labor Market Research: Evolving Structures and Processes*, edited by Ivar Berg and Arne Kalleberg. New York: Plenum.
- Reskin, Barbara F., and Debra B. McBrier. 2000. “Why Not Ascription?: Organizations’ Employment of Male and Female Managers.” *American Sociological Review* 65: 210–33.
- Rothbart, Myron, and Susan Lewis. 1994. “Cognitive Processes and Intergroup Relations: A Historical Perspective.” In *Social Cognition: Impact on Social Psychology*, edited by Patricia G. Devine and David L. Hamilton. San Diego: Academic Press.
- Simmel, Georg. 1971 [1908]. *On Individuality and Social Forms: Selected Writings*. Edited and with an introduction by Donald N. Levine. Chicago: University of Chicago Press.
- Stryker, Robin. 2001. “Disparate Impact and the Quota Debates: Law, Labor Market Sociology, and Equal Employment Policies.” *Sociological Quarterly* 42(1): 13–46.
- Sturm, Susan. 2001. “Second Generation Employment Discrimination: A Structural Approach.” *Columbia Law Review* 101(3): 101–207.
- Swim, Janet K., and Lawrence J. Sanna. 1996. “He’s Skilled, She’s Lucky: A Meta-analysis of Observers’ Attributions for Women’s and Men’s Successes and Failures.” *Personality and Social Psychology Bulletin* 22: 507–19.
- U.S. Equal Employment Opportunity Commission. 2001a. “EEOC and Private Plaintiffs Settle Harassment Suit for \$485,000 Against Chicken Processing Plant.” *EEOC Press Releases* (April 10). Available at: www.eeoc.gov/press/4-10-01.html (June 20, 2001).
- . 2001b. “EEOC Responds to Final Report of Mitsubishi Consent Decree Monitors.” *EEOC Press Releases* (May 23). Available at: www.eeoc.gov/press/5-23-01.html (June 20, 2001).
- . 2001c. “EEOC Seeks to Join Nationwide Sex Discrimination Suit Against Rent-A-Center.” *EEOC Press Releases* (March 12). Available at: www.eeoc.gov/press/3-12-01.html (June 20, 2001).
- . 2001d. “EEOC Settles Bias Suit for \$2.6 Million Against TWA.” *EEOC Press Releases* (May 24). Available at: www.eeoc.gov/press/5-24-01.html (June 20, 2001).
- . 2001e. “EEOC Settles Racial Harassment Suit Against Georgia-Pacific Corporation.” *EEOC Press Releases* (April 3). Available at: www.eeoc.gov/press/4-3-01.html (June 20, 2001).
- . 2001f. “EEOC Sues Two Indiana Employers for Race Harassment.” *EEOC Press Releases* (May 22). Available at: www.eeoc.gov/press/5-22-01.html (June 20, 2001).
- . 2001g. “Joe’s Stone Crab Liable for Intentional Discrimination Court Rules in Sex Bias Suit Brought by EEOC.” *EEOC Press Releases* (March 28). Available at: www.eeoc.gov/press/3-28-01.html (June 20, 2001).
- Weber, Joseph G. 1994. “The Nature of Ethnocentric Attribution Bias: Ingroup Protection or Enhancement?” *Journal of Experimental Social Psychology* 20: 177–94.
- Wilson, William J. 1996. *When Work Disappears*. Chicago: University of Chicago Press.

83 ■ Trond Petersen

Discrimination**Conscious or Nonconscious?****Introduction**

The struggle for equality between men and women in the workplace resulted in one of the most significant achievements of the twentieth century and probably the single most lasting legacy of the 1960s. It changed in major ways how organizations operate and how men and women live their lives.

The movement was prefaced on the idea that employers were a major culprit for gender inequality and that they hence need to be regulated through legislation and other means. Underlying the presumption was an implicit and often explicit purposive-action framework. Employers were seen as consciously treating men and women differently, either due to animus or to cost considerations.

A remarkable and perhaps unexpected challenge to such reasoning has arisen with great strength over the last decade. It originated in cognitive psychology, recently bolstered by neuroscience, and has spread to sociology and parts of legal and economic scholarship. It has furthermore entered the consciousness of journalists and human resource managers in the United States and Europe. It calls for abandoning the standard purposive-action framework for understanding female disadvantage in the workplace.

The claim behind the challenge is this. Discrimination is no longer intended and conscious. It results today primarily from cognitive processes that are nonconscious.

The mechanisms behind discrimination are hardwired by biology into our brains and are not accessible to human decision makers.

Reskin (2000, p. 327), a sociologist, writes about the sources of discrimination: “I have argued that much of it results from nonconscious cognitive processes.” Valian (1998, pp. 2, 22), a psychologist, refers to the sources as “a set of implicit, or nonconscious hypotheses” or as “the unexpressed and nonconscious nature of gender schemas and their subterranean mode of action.”

This view extends beyond researchers. In a significant U.S. media story, one commentator stated in a front-page newspaper article: “Everybody has nonconscious gender bias. It shows up in every study” (Rimer 2005, p. A15).

In this chapter I discuss the challenge posed by this literature. I first outline the standard view stemming from a purposive-action framework where agents are viewed as being conscious about the motivations for their actions, then outline the challenge from cognitive psychology, followed by an evaluation, and ending with the implications for jurisprudence, our understanding of gender inequality, and for its solutions. The focus is on gender inequality in the workplace, such as hiring, wages, and promotions. I abstract from the separate and different issues that arise in race inequality, where in addition to discrimination, disadvantages stemming from class play a large role.

Standard View: Conscious Action

In the standard view there are two main motives of discriminatory behaviors from employers. One motive arises from animus, where an employer dislikes or harbors a grudge against women and/or ethnic minorities. The other motive arises from statistical considerations, where women on average are viewed to be less productive, referred to as statistical discrimination.

Discrimination, from animus or statistical considerations, is rational in the sense that (a) it reflects preference, (b) it is intended, and (c) the employer is conscious about (a) and (b). Action is conscious and directed toward achieving goals that the agent values.

In case of animus discrimination, someone who acts on it maximizes his preferences. An employer is willing to lose money in order to avoid hiring women and/or minorities. The preference itself, the animus against women or minorities, is widely seen to be irrational, and certainly despicable. But once in place, acting upon it, as such, is not irrational. Economists refer to this as taste discrimination (Becker 1971).

Discrimination from statistical considerations has the same structure. An employer may harbor no animus against groups, but treating them differently may maximize preferences for higher profits. The underlying reason is that groups may differ in how productive they are. For example, average productivity may be higher for men in the construction industry and higher for women in electronics assembly. It may, however, be costly to assess individual productivity prior to hire or once hired. Hiring primarily from the more productive group may then save on hiring costs and increase productivity and profits.

Legislation, judicial practice, and much public policy were premised on the idea that discrimination is intentional, and hence conscious, and also rational in the sense that the agent is trying to satisfy his preferences, from either animus or desire for increased profits.

Legislation, at a minimum, made discriminatory choices costlier. For those acting on animus, it would become riskier to satisfy these. For those engaging in statistical discrimination, the law would erode the expected profit advantage. Legislation hence shifts the trade-offs, and rational agents should be expected to adjust accordingly.

While the standard view of conscious action undoubtedly applies to disparate treatment discrimination, it needs to be modified for analyzing disparate impact discrimination. The latter refers to the situation where two groups are treated equally within the given set of organizational rules and procedures, but where the procedures advantage one group over another and at the same time are irrelevant for job performance. An example in hiring would be a physical height requirement unrelated to job performance, which would favor men over women. The hiring agent or institution may be unaware that the procedures are discriminatory, and there is thus no discriminatory motive, conscious or nonconscious.

The Challenge: Nonconscious Processes

What is then the challenge from cognitive psychology? It proceeds from exactly the opposite assumptions of the purposive-action framework. As an explanation of discrimination it claims that there is (a) no necessary bias against women, (b) no necessary intent to treat the sexes differently, and (c) discriminatory agents are driven by nonconscious preferences or biases.

The reasons for discrimination are now cognitive structures unknown to the agent. These structures are part and parcel of how we think and function as human beings. They involve processes of categorization (e.g., by sex), the use of simplified schemas, and errors in information processing and causal attribution. Such nonconscious cognitive processes can lead to biases against women and minorities, without the agent

ever being aware of it. The biases could reflect broader societal stereotypes, or may just reflect hardwiring of the brain that compels humans to make categorizations.

The forceful claim from cognitive psychology is that the nonconscious schemas we use, together with the nonconscious biases, can crowd out well-meaning conscious preferences and intentions. They may result in a total disconnection between conscious intent and actual action. A person with no conscious animus or stereotype, with no conscious intent to discriminate, nevertheless, due to entirely automatic nonconscious cognitive processes, may end up making biased evaluations, may discriminate, and, to top it off, will be totally unaware that the behavior was discriminatory.

This change in theoretical orientation from conscious to nonconscious processes was in part prompted by the realization among many researchers that conscious prejudice and overt discrimination no longer are central sources of disadvantage for women and minorities. There are, however, still substantial inequalities between groups. The purported central cause is now nonconscious cognitive biases.

Evaluation

The claim that discrimination primarily is nonconscious rests on the view that most discrimination arises from bias, giving less role to statistical discrimination. Before discussing this claim, the evidence for it, and its implications, it is useful to assess the presumption that most discrimination arises from nonconscious bias. I have no empirical information to evaluate this claim, nor do I think any other researcher has. But I think it needs one important qualification.

I would like to suggest that in the instances where discrimination has occurred, it is prudent to reserve judgment on the prevalence of nonconscious versus conscious practices. To the extent that it arises from

bias, I think most discrimination is nonconscious. As argued by cognitive psychologists, one would have to search long to find many card-carrying sexists proudly thinking that “I will do my part for the patriarchal project to keep women from getting hired and promoted.”

But some employers may still consciously act upon correct, or what they perceive to be correct, statistical information. This is entirely conscious. And it is entirely rational. It may minimize hiring costs and maximize profits. It is still illegal: A woman should not be treated as a representative member of her sex; qualifications should be evaluated separately for each person. This could well be the most prevalent source of employment discrimination.

There are also instances where employers act on incorrect statistical information, though believing it to be correct, and then engage in what has been called error discrimination (England 1992, chap. 3). Though based on unfounded stereotypes, positive or negative, such action would not be nonconscious. The stereotypes will in all likelihood be eradicated as better information becomes available.

Exactly what evidence then exists for the claim of nonconscious discrimination? Most of the evidence comes from laboratory experiments on undergraduate students at U.S. and European universities and now increasingly from computer experiments conducted on large samples (Dasgupta 2004).

The main procedure currently used for detecting nonconscious bias is the Implicit Association Test (see Greenwald, McGhee, and Schwartz 1998). Subjects are asked to compare words at a brisk pace while sitting in front of a computer screen. Since the interval between receiving a stimulus and making a response is short, often a third or less of a second, the claim is that the response is beyond the subject’s conscious control and hence must be caused by nonconscious processes. The tests document widespread biases.

But do these tests and the conclusions from them identify the central problems women face in the workplace today?

One need be careful with respect to drawing wide-ranging conclusions from laboratory and computer experiments. Besides the question of whether the results reflect animus versus stereotypes or statistical generalizations (Arkes and Tetlock 2004), there is a broader issue that needs addressing. How undergraduates, but also older people, react in experimental settings may be entirely different from how decisions are made by experienced people in real-life settings where consequences may be far-reaching economically, socially (workplace culture), and legally. Employment decisions are often made with considerable deliberation, not on the spot, and often while carefully monitoring their potential adverse impacts on women and minorities. This will be the case in most large organizations. Parts of employment processes will clearly involve on-the-spot decisions, such as how to greet interviewees, and may hence be less deliberate. But much happens both before and after that may counterbalance split-second biases. And decisions are not only made with care, in midsized to large organizations they are typically made by multiple people, including equity officers, and create an extensive paper trail of job resumes, reports from meetings, and written justifications. The decisions are far removed from those made in the laboratory.

The difference between the two decision-making situations is hence stark: mostly spot and split-second in experiments with minimal consequences, versus deliberate and slow in the workplace with almost invariably significant consequences.

The question is not one of truth: cognitive science has almost certainly delivered real findings. The question is one of applicability, whether the findings extend to decisions in the workplace, and moreover one of balance—what is the relative weight of nonconscious processes in creating differential treatment and outcomes.

Implication 1: Jurisprudence

Implications for legal practice and for how organizations construct their procedures to avoid litigation may vary according to whether discrimination is conscious, and often rational, versus nonconscious and probably irrational. In the latter case there is no discriminatory intent, which may lead to difficulties in the courts.

White and Krieger (2001, p. 510) argue that proving nonconscious discrimination should rely on same types of evidence as proving covert but conscious discrimination; cases where the employer hides his discriminatory motives. The idea is that in such cases it should be sufficient to show that a disparate outcome by sex or race cannot be justified. One infers from the outcome that the employer acted from bias, in spite of no direct evidence of it. The same idea can be used for nonconscious discrimination. In either case, the courts would in principle not have to rule on whether the motive was conscious, but not revealed, or nonconscious.

Implication 2: Inequality

For understanding how inequality is created, and for public policy, the prevalence and importance of a mechanism matters. We need to assess which are the central forces in generating inequality: discrimination, conscious and nonconscious, educational choices, family adaptations, and more.

Krieger (1995, p. 1164) argues with respect to nonconscious discrimination that “these subtle forms of bias, I suggest, represent today’s most prevalent type of discrimination.” And concerning its scope, White and Krieger (2001, p. 509) state: “that ‘non-conscious’ discrimination frequently occurs is well documented.”

But how can these processes, which are nonconscious and often hardly noticeable, add up to a major gender gap in employment? Valian (1998, p. 18) stresses that “a succession

of small events, such as not getting a good assignment, result in large discrepancies in advancement and achievement.” They are like accrued interest, which in any one year makes little difference, but over time adds up.

There is no question that subtle and non-conscious mechanisms may operate. The question is whether they are the central causes of the major differences we still observe between men and women. I would like to suggest that other factors are more important and overpower the small causes, namely the educational choices and family adaptations men and women make. These are observable, not at all subtle, and undoubtedly have major impacts.

Implication 3: Solutions

What are the implications for public policy and organizational remedies?

A purposive-action framework would argue that to the extent that discrimination still is a major problem, then penalties and enforcement need to be strengthened. This will shift incentives. There will also be shaming of violators, an added cost. For statistical discrimination this should be sufficient. For nonconscious discrimination, the suggested remedies would depend on the kinds of evidence for bias that are acceptable. If the nonconscious motive must be shown, there is a difficulty. But if all that is required is proof of a disparate outcome that appears unjustified, strengthened enforcement and penalties should again suffice. Many employers are aware that non-conscious bias may exist, understand the difficulties of proving or disproving this, and may take precautions to guard against accusations and litigation. A purposive-action framework serves well to analyze such adaptive behavior. Although only a small portion of equality issues can be brought to the courts, legislation and legal practice can set important frameworks for how employers will act, on the assumption that the overwhelming majority of employers seek to comply with the law.

Scholars taking their view from cognitive science offer several kinds of advice for fighting discrimination arising from nonconscious biases.

Some scholars see the need for broader cultural changes. Jolls and Sunstein (2006) argue for the role of government information campaigns aided by the law. These may sensitize decision makers to nonconscious biases and induce them to safeguard against their operation. In a similar vein, Krieger (1995, pp. 1245-1247) sees the biases as part of a broad cultural problem, and concludes that to alter the cognitive schemas we use thus requires a broad-based cultural solution.

Other scholars focus on workplace processes. Bielby (2000) argues that organizations need to be more careful in their hiring and promotion procedures, and that procedures should be made sensitive to the operation of nonconscious biases. This can minimize their impact. Krieger (1995, p. 1216) also sees a role for policies at the workplace level, but focuses more on changing the psychology of decision makers than on procedures.

These proposals can be justified without recourse to a theory of nonconscious biases. For example, broad-based cultural solutions can change the nature of conscious preferences. They may root out animus against specific groups. This has already occurred. There used to be widespread conscious animus. There no longer is. What is left, according to cognitive psychologists, is widespread nonconscious bias. The puzzle is that the successful campaign to eradicate conscious bias may have failed in also changing nonconscious cognitive categories.

Conclusions

Summary

Legislation and theorizing on gender and race inequality originally proceeded on the assumption that employers consciously and rationally discriminated against women and minorities. A challenge to this view has

emerged from cognitive psychology. The claim is that discrimination today no longer is conscious, but arises rather from nonconscious schemas that underlie all decision making. This change in emphasis reflects in part the realization that conscious prejudice and overt discrimination no longer are central forces behind disadvantage. In part it reflects a broader movement in the social sciences that draws on cognitive psychology, neuroscience, and biology.

For differential treatment discrimination that has its genesis in animus or unfounded stereotypes, I think cognitive psychology is on target: It is mostly nonconscious. I just doubt that nonconscious discrimination is a major source of female disadvantage in the workplace. Statistical discrimination may still be important, and it is entirely conscious and economically rational, though illegal.

Cognition Extended

There is no question that our conscious cognitive structures have changed dramatically during the last 40 years. The focus on equality issues in legislation, in public policy, and in how organizations operate no doubt has brought along a major cultural change in most western societies. The framework for debate and decision making is entirely different in year 2000 than in 1960. Our conscious processes are hardly recognizable. But perhaps our nonconscious categories have changed as well? If they can be influenced at all by broader cultural changes, as White and Krieger (1995, p. 510) argue, then there must have been significant responses to the very visible material and cultural transformations that have occurred. Krieger (1995, p. 1202, n. 179) thinks this has yet to happen. I would think that no change seems unlikely, and major change appears likely.

There is simply reason to think that prolonged attention to a particular problem, in this case gender inequality, may have consequences not only for our procedures in organizations, as they have had, not only for the

content of our conscious thought, as they also have had due to major cultural changes, but also for the content of our nonconscious cognitive categories.

REFERENCES

- Arkes, Hal R., and Philip Tetlock. 2004. "Attributions of Implicit Prejudice, or 'Would Jesse Jackson Fail' the Implicit Association Test?" *Psychological Inquiry* 15(4): 257–278.
- Becker, Gary S. 1971. *The Economics of Discrimination*. Chicago: University of Chicago Press.
- Bielby, William T. 2000. "Minimizing Workplace Gender and Racial Bias." *Contemporary Sociology* 29(1): 120–129.
- Dasgupta, Nilanjana. 2004. "Implicit Ingroup Favoritism, Outgroup Favoritism, and Their Behavioral Manifestations." *Social Justice Research* 17(2): 143–168.
- England, Paula. 1992. *Comparable Worth: Theories and Evidence*. Hawthorne, NY: Aldine de Gruyter.
- Greenwald, Anthony G., Debbie E. McGhee, and Jordan L. K. Schwartz. 1998. "Measuring Individual Differences in Implicit Cognition: The Implicit Association Test." *Journal of Personality and Social Psychology* 74(6): 1464–1480.
- Jolls, Christine, and Cass R. Sunstein. 2006. "The Law of Implicit Bias." *California Law Review* 94(4): 969–996.
- Krieger, Linda Hamilton. 1995. "The Content of Our Categories: A Cognitive Bias Approach to Discrimination and Equal Employment Opportunity." *Stanford Law Review* 47(6): 1161–1248.
- Petersen, Trond. 2006. "Motive and cognition: Conscious and unconscious processes in employment discrimination." Chap. 14 (pp. 225–248) in Jon Elster, Olav Gjelsvik, Aanund Hylland, and Karl Moene (Eds.), *Understanding Choice, Explaining Behavior*. Oslo, Norway: Unipub Forlag/Oslo Academic Press.
- Reskin, Barbara F. 2000. "The Proximate Causes of Employment Discrimination." *Contemporary Sociology* 29(2): 319–328.
- Rimer, Sandra. 2005. "For Women in Sciences, the Pace of Progress in Academia Is Slow." *New York Times*, April 15, 2005, pp. A1, A15.
- Valian, Virginia. 1998. *Why So Slow? The Advancement of Women*. Cambridge, MA: MIT Press.
- White, Rebecca Hanner, and Linda Hamilton Krieger. 2001. "Whose Motive Matters?: Discrimination in Multi-Actor Employment Decision Making." *Louisiana Law Review* 61(3): 495–538.

Sex Segregation

84 ■ William T. Bielby

The Structure and Process of Sex Segregation

My work seeks to understand the persistence of occupational segregation by sex and the sex gap in earnings. Neoclassical economists, of course, have an elegant, coherent theory of the sources of both occupational segregation by sex and earnings differentials: they come from the rational, utility-maximizing behavior of men and women within households and labor markets.

Sociologists view the situation differently. Indeed, depending on whom you ask, a sociologist might offer any of a number of explanations for sex differences in labor-market outcomes. Social psychologists would stress sex-role socialization. Feminist scholars would emphasize the “patriarchal” interests of male employers and workers. Marxists would tell a story about capitalists creating divisions in the work force in order to control labor and boost profits. Organizational theorists might talk about the unintended consequences of bureaucratic rules and procedures. Not surprisingly, sociologists from different camps often talk past one another.

However, sociologists are increasingly taking the work of neoclassical economists seriously. The language of “human capital,” “statistical discrimination,” and the like is increasingly part of the sociological discourse

on gender and work. However, sociologists do not accept the neoclassical account uncritically. In this chapter, I first summarize how my research with James Baron and Denise Bielby challenges neoclassical explanations and then raise some issues concerning the promise and limitations of recent interdisciplinary approaches to the study of job segregation by sex and other forms of discrimination.

I. Neoclassical Views

In a sense, it all starts with the household division of labor. Women prefer, get stuck with, or have a comparative advantage caring for children and the home. Because of that, they invest less in market human capital. For example, intermittent labor-force participation leads to less work experience and thus to lower productivity and flatter age-earnings profiles. Women rationally *choose* occupations that are easy to leave and reenter, where skills do not atrophy, and so on. These choices lead to occupational segregation and wage differentials (see Polachek and Kao [1991]; Polachek [1979]).

Becker [1985] has offered a clever twist on this explanation. People do not simply allo-

cate time between household and market activities. They also allocate *effort*. Since women allocate more effort to household activities, they have less effort to allocate to work outside the home. So, even when men and women have the same *amount* of work experience, training, etc., women are less productive. An additional hour of women's work is less effort-intensive than one of men's work. As a result, women are worth less to employers and accumulate less market human capital with each additional hour of work. This leads to an explanation for both occupational segregation by sex and earnings differentials within occupations. Women choose less effort-intensive jobs, and they are less productive than men working on the same jobs.

There is, of course, another view of occupational segregation that admits to a certain kind of "discrimination," but only a *rational* kind of "statistical discrimination." In a world filled with uncertainty, employers cannot always know the relevant traits of job applicants. Suppose that there are important traits that are difficult to measure, but we know that, on average, women exhibit more (or less) of them than men. One example often cited is the likelihood of job turnover. Suppose also that there are two kinds of jobs—those for which it is costly to replace a worker who quits and those for which it is not. Not knowing actual quit propensities, rational, profit-maximizing employers will reserve the former jobs for men and fill the latter with women, even if there is considerable overlap in the underlying distributions of quit propensities of men and women (Phelps [1972]; Aigner and Cain [1977]).

My research in collaboration with Denise Bielby (Bielby and Bielby [1988]) and James Baron (Bielby and Baron [1986]) suggests that despite the elegance of the neoclassical accounts, the world just does not work that way. Men and women employees do not act in a way that is consistent with neoclassical accounts, nor do employers.

II. Empirical Challenges to Neoclassical Views: The Allocation of Effort

Denise Bielby and I, using the 1973 and 1977 Quality of Employment Surveys, attempted to operationalize a human-capital model of the allocation of work effort (Bielby and Bielby [1988]). Here I summarize the findings relevant to the issue of job segregation and earnings disparities.

Contrary to the clear implication from Becker's model, women do not exert less effort than men in the workplace in order to conserve effort for household activities. Compared to men with similar household responsibilities, market human capital, earnings, promotion opportunities, and job responsibilities, women allocate substantially more effort to work activities; the net sex difference is nearly one-half a standard deviation on the scale we use. To the extent that women do allocate effort away from the workplace in order to meet family demands, these trade-offs bring their work effort back to the level of the typical male with no such family responsibilities.

How could this be? We believe that the neoclassical model of the allocation of work effort is seriously flawed in several ways. First, it may be inappropriate to assume that the "stock" of effort to be allocated by an individual is fixed, analogous to a fixed stock of physical capital. For women to work just as hard as men, if not harder, despite their greater household responsibilities, women must be able to draw on a reserve of energy that is either not available to the typical male, or more realistically, that men choose not to draw upon. Stephen Marks, a sociologist who has written on this topic [1977: 927], cites physiological evidence in support of his claim that "individuals have abundant and perpetually renewing [energy] resources." He suggests that individuals generate the energy needed to participate in activities to which they are

committed and that they often feel more “energetic” after engaging in them. Our findings imply that as women add work roles to their family roles, they generate the energy necessary to fulfill their commitments to the two sets of activities.

A second factor ignored by human capital models is what social psychologists call “entitlement norms.” The allocation to effort at work is influenced by internalized norms of a “fair day’s work,” which may differ by sex. Determinations of fairness are based on perceptions of contributions relative to rewards. Experiments show that men and women differ in how they invoke equity considerations in allocating effort and rewards. On average, women pay themselves less than men performing the same task, and women tend to undervalue their efforts relative to men (Lenney [1977]; Callahan-Levy and Messe [1979]; Major, McFarlin, and Gagnon [1984]). It appears that, on average, women have lower internal standards of “personal entitlement,” and, in the absence of salient external comparison standards, they make fairness judgments based on application of same-sex norms about appropriate rewards (Berger et al. [1972]; Crosby [1982]).

In a laboratory experiment designed specifically to examine sex differences in effort and standards of personal entitlement, Major, McFarlin, and Gagnon [1984] asked men and women to do as much work as they thought was fair for a fixed amount of money. They collected objective measures of the accuracy and efficiency of performance on the task as well as information on each subject’s perceptions of his or her level of performance. They discovered that, on average, “women worked longer, did more work, completed more correct work, and were more efficient than men” (1409). At the same time, men and women did *not* differ in their self-evaluations of their performance, despite women’s superior objective performance. Nor did men and women differ in their reports of satisfaction from the task. In

short, laboratory research on sex differences in internal standards of personal entitlement suggests that, all else constant, women can be expected to allocate more effort than men to work activities.

Third, structural features of workplaces might reinforce sex differences in entitlement norms. For example, research on sex segregation in the workplace suggests why same-sex norms about rewards for performance persist. Most men and women work in sex-segregated jobs (Bielby and Baron [1984], [1986]; Reskin and Hartmann [1986]). Consequently, many women lack information on both the numerator and the denominator of the “equity ratio” of reward to effort for men’s work. According to social psychologists, it is precisely when external bases of social comparison are unavailable that individuals rely on internal, same-sex norms (Berger, Zelditch, Anderson, and Cohen [1972]; Berger, Rosenholtz, and Zelditch [1980]; Austin [1977]). Major, McFarlin, and Gagnon found that when external comparison information was available, neither sex relied on same-sex norms to make judgments of fair pay for a given task. Thus equity research suggests why employers might have a stake in a sex-segregated workforce: as long as women lack information about the reward structure for men, they will be willing to work for less pay. (Moreover, husbands of working women benefit by avoiding responsibility for household activities, despite the fact that their wives expend as much or more energy as they do at work.)

Fourth, another aspect of our analysis calls into question the extent to which decisions about work and family are made in a deliberately self-conscious, rational manner. Bielby and Bielby [1988] found that men’s work effort is greatly influenced by family demands, while the effects of “human-capital” investments such as schooling and labor-market experience are relatively weak. In contrast, compared to men, women’s work effort is less strongly influenced by family demands, but is

more strongly affected by investments in market human capital. For example, a man with child-care responsibilities reduces his work effort, but a woman in the same situation does not. At the same time, a man with meager investments in “market human capital”—education and labor-market experience—does not decrease his work effort, but a woman with comparable investments does decrease her work effort. This suggests that individuals take sex-specific culturally prescribed roles for granted. Women do not adjust their effort allocations in response to household responsibilities, and men do not adjust their effort allocations in response to workplace investments.

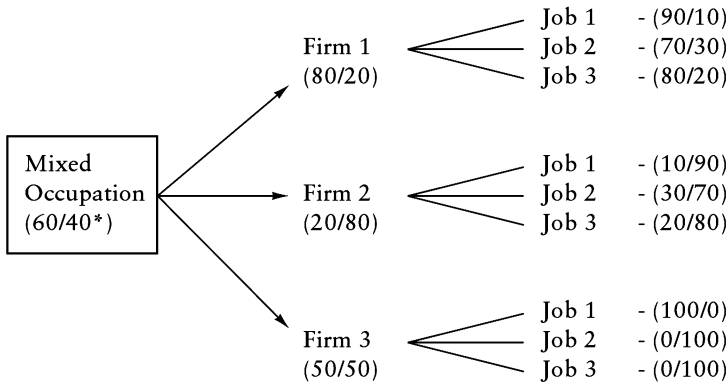
It appears that individuals *do* rationally reflect upon their efforts when they cannot fall back upon culturally prescribed roles. Men confronted with household responsibilities reduce their work efforts, as do women with meager investments in market human capital. For men, some level of work effort may be culturally prescribed and “taken for granted” regardless of investments in education and training. For women, a certain level of work effort may be culturally prescribed and “taken for granted” regardless of the level of household responsibilities. Moreover, the level of effort allocated to work by a woman with heavy household responsibilities is at least as large as that allocated by men with no such responsibilities at home.

In sum, Becker’s human-capital model of the allocation of work effort provides a plausible explanation for both sex segregation and wage differentials by sex. However, our research and other social psychological studies call into question some basic premises of this model. The assumption of a fixed stock of effort may be unrealistic, the social psychological processes may be more complex than what is typically assumed about “economic man,” and segregation of men’s and women’s jobs may reinforce sex differences in social psychological processes.

III. Empirical Challenges to Neoclassical Views: Noneconomic Bases of Statistical Discrimination

James Baron and I set out to analyze how men and women in the same general line of work are sorted into different firms and into different jobs within firms (Bielby and Baron [1984], [1986], [1987]; Baron and Bielby [1985]). Figure 1 shows the issue schematically. Consider an occupation like “electronics assembler” that is 40 percent female. Suppose that three firms in an industry employ electronics assemblers and that each of these firms uses assemblers in three of its job titles. In the hypothetical example, firm 1 relies disproportionately on male workers, firm 2 relies disproportionately on female workers, and firm 3 relies on a balanced mix of male and female workers. In firm 1, each of the jobs is staffed disproportionately by men, and in firm 2, the opposite is true. In firm 3, however, the first job is staffed exclusively by men, and the second and third jobs exclusively by women. We can imagine other patterns besides the one depicted in figure 1 for an occupation that is 40 percent female. For example, in a world without statistical discrimination, we might find the sex composition of specific jobs normally distributed around 40 percent female. In any case, according to neoclassical models of the labor market, knowledge of differences between firms and jobs in work requirements and differences between the sexes in job traits should explain how men and women in mixed occupations get distributed across firms and jobs.

Baron and I assembled data like this for several hundred California firms studied by the U.S. Employment Service in the late 1960s and 1970s. That is, we looked at how firms hiring from a sex-mixed labor pool staffed their jobs as a function of firm and job characteristics. Neoclassical models suggested the following hypotheses: scale, bureaucratization, and job-specific training are



*Ratio of Male Workers to Female Workers

Figure 1. Hypothetical distribution of workers in a mixed occupation across firms and jobs

related to turnover costs, so if women are more likely to quit, they should be underrepresented in firms and jobs where these costs are high. If employers statistically discriminate based on turnover costs, women should be excluded altogether from jobs where these costs are high. The remaining job traits we looked at are job-relevant skills. Women have been shown to perform better, on average, on traits such as verbal skills, motor coordination, finger dexterity, interpersonal skills, and stamina under repetitive, monotonous working conditions. Men have been shown to perform better, on average, on numerical skills, spatial skills, and upper body strength. Thus the distribution of men and women across jobs should reflect differences in jobs in the extent to which they require these sex-linked traits.

Our most startling finding was that California firms in the late 1960s and 1970s rarely employed men and women in the same job titles. Even when drawing from a sex-mixed labor pool for a work role like "electronics assembler," firms reserved some job titles exclusively for men and others exclusively for women. That is, firms either employed male workers exclusively or female workers exclusively, or, if they employed both, men and women were com-

pletely segregated by job title, as in firm 3 in Figure 1.

In addition, we found that the relationships between firm and job traits and the sex composition of work were exactly as hypothesized. "Male" traits were positively related to the likelihood of excluding women from a job and the effects of "female" traits were in the opposite direction. Our research indicates that employers exclude women from some jobs and men from others (in the same general line of work) based on their perceptions of group differences between men and women in job-related traits. That is, employers statistically discriminate.

However, our findings did not support the neoclassical version of statistical discrimination. Statistical discrimination is efficient only when it is costly to measure the actual job-relevant traits of individual workers. The factors that made the greatest difference in whether men or women were employed were not the subtle and difficult-to-measure factors like turnover costs. They were physical demands and finger dexterity. Men are assigned the jobs perceived to involve heavy work and women—in the same occupation—are assigned the "light" work. Women are assigned jobs defined as "detail" work, and men are not.

According to neoclassical economists, employers are sacrificing profits in order to sustain the type of sex segregation we observed. Our measure of physical demands was whether a job required lifting more than twenty-five pounds. In most of the firms we studied, no woman would be assigned to these jobs, regardless of that woman's own physical strength. Yet it is simple and inexpensive to devise a test to learn whether a female (or male) job applicant is capable of lifting twenty-five pounds. Similarly, it should be easy to devise a simple test that would allow an employer to observe whether a male (or female) job applicant could successfully complete a task requiring finger dexterity. Given that they were drawing from sex-mixed labor pools, why did employers persist in sustaining almost complete job segregation by sex? We had access to narrative material about the employment practices in these California firms, and analyses of these qualitative materials suggested several important factors.

Policies. In many of the firms studied by the Employment Service, it was a matter of company policy not to employ women in certain jobs and to favor employing women in others. Jobs closed to women were typically perceived to be physically demanding and those favoring women usually involved "detail" work. Curiously, while we often came across policies specifying that women were preferred for certain jobs, we rarely encountered official *prohibitions* against men being employed in these jobs. Yet in practice, men rarely were employed in predominantly female jobs unless they were supervisors and in that case, they would almost always have distinct job titles.

There are at least two reasons for the lack of any explicit prohibition against using men in specific jobs, despite the nearly perfect segregation in practice. First, if female labor is less expensive, employers have no incentive to employ men in the same job categories as

women, even if the company has no policy against doing so. Second, prohibitions against women working in certain jobs often originated in protective legislation (even if these regulations no longer applied), and thus complying with government regulations was often the original rationale for having a policy on women's work. No similar regulations mandated policies on men's work.

Organizational Inertia. Once a sex-based division of labor is established, it becomes "taken for granted," sustained over the years unless some deliberate effort is taken to undo it. This was especially evident in the exclusion of women from jobs perceived to be physically demanding. California laws prohibited the employment of women in jobs requiring heavy lifting until these laws were struck down in 1971. Yet our data showed that the exclusion of women from so-called physically demanding jobs was just as widespread through the late 1970s—when the practice was illegal—as it was during the period before 1971. We also saw this inertia in data from firms analyzed by the Employment Service at two points in time, typically six or seven years apart. The sex-based division of labor rarely changed over that six- or seven-year period, even when total employment fluctuated substantially, *unless* there was some deliberate effort taken to undo sex segregation.

Cultural Stereotypes. Cognitive psychologists have demonstrated that we all use stereotypes, or cognitive schema, as a cognitive shorthand we invoke to achieve economy in perception (Ashmore and Del Boca [1981]). Of course, the particular stereotypes we hold about sex roles are widely shared and continually reinforced in our culture. Despite what is assumed by neoclassical models of discrimination, the profit motive is not necessarily strong enough to overcome the cognitive processes that lead to the use of stereotypes. Thus it may be that employers make hiring

decisions based on widely shared *misperceptions* of differences in the average traits of men and women. To complicate matters further, statistical discrimination can lead to “self-fulfilling prophecies,” or what social psychologists call “expectancy confirmation sequences” (Darley and Fazio [1980]). That is, employers expect certain behaviors from women and assign them to routine tasks and dead-end jobs. Women respond by exhibiting the very behaviors employers expect, reinforcing the stereotype. Moreover, social psychological research suggests that employers are more likely to attend to and retain information that confirms stereotypes, ignoring behavior by women that does not fit their expectations.

Given the kind of data available to Baron and me, we had no direct evidence of cultural stereotypes held by specific employers or managers about men’s work and women’s work. However, such stereotypes were invoked quite often in policies regarding the employment of women in “heavy” work and in “detail” work. Indeed, it was not uncommon to encounter a policy prohibiting employment of women in certain jobs because of heavy lifting requirements and to then find in detailed job analyses by the Employment Service that these jobs required no heavy physical labor. In other words, the work is often labelled as “heavy work” because it is done by men.

Interests. Finally, the extreme sex segregation we discovered within mixed occupations could not have been sustained without someone’s interests being served. It is not difficult to see how employers might have an interest in sex segregation. For example, our research on the allocation of effort suggests that segregation sustains sex-specific norms of entitlement. Male employees who view integration as a threat to their wages may resist the assignment of women to the job titles they hold. However, not at all clear from existing research are the circumstances under which

male workers have the capacity to act on these interests, given employers’ economic incentive to substitute cheap female labor for expensive male labor. Although my research with Baron on statistical discrimination did not get at the interests issue directly, other work we have done suggests a slight tendency for greater segregation in settings where male workers have substantial market power (Bielby and Baron [1984]).

IV. Conclusion: Reconciling Alternative Perspectives on the Structure and Process of Sex Segregation

We can identify six more or less distinct kinds of explanations for a single phenomenon: sex segregation in the workplace. These explanations are (1) neoclassical, based on models of statistical discrimination and investments in human capital; (2) social psychological, emphasizing socialization and internalized norms; (3) institutional, emphasizing the intended and unintended consequences and inertia of organizational arrangements; (4) cultural, emphasizing taken-for-granted notions of men’s work and women’s work, often shared by both men and women; (5) political, stressing the different interests of male employers and employees with respect to maintaining the status quo; and (6) patriarchal, emphasizing the common interests of male workers and employers in maintaining a sex-based division of labor.

The research my colleagues and I have pursued has sought to test neoclassical explanations of gender differences in the workplace, not to offer fully developed alternatives. In posing tentative accounting for our findings, we have drawn on both institutional approaches to the structure of work and social psychological theories of entitlement norms. Of course, these areas are no longer beyond the purview of neoclassical economics. For example, Williamson ([1975], [1985]) explains variation in organizational structures

with the concept of economizing on transaction costs, and “efficiency wage” theorists invoke entitlement norms to explain interindustry wage differentials (Krueger and Summers [1987]; Akerlof and Yellen [1988]).

It is customary to applaud cross-disciplinary efforts, and the blurring of the boundaries between sociology and neoclassical economics may eventually lead to a comprehensive explanation of job segregation by sex that is consistent with the empirical evidence on the subject. By incorporating “sociological factors” (e.g., institutional barriers, entitlement norms), neoclassical models may be able to more faithfully represent the factors that slow the rate at which market forces erode discrimination in pay and job assignment. Moving in that direction requires continued commitment to the neoclassical assumption that were it not for various “drags” on the system, market forces would inevitably discount ascriptive factors like race and gender in valuing otherwise comparable labor inputs.

Adapting the neoclassical model in this way has appeal due to the power of the model’s formal analytical properties. However, the pervasiveness and persistence of job segregation by sex seems too far at odds with the imagery of even the most extreme adaptations of the model (for a similar view, see Darity [1991]). For example, it is difficult to imagine what kinds of “transaction costs” could explain the nearly universal job segregation of men and women *within the same occupation* that Baron and I found across a wide variety of organizational settings. That such segregation would persist is especially problematic from a neoclassical perspective, since the rationale was typically based on stereotypical perceptions of men’s and women’s job aptitudes, often with little basis in the actual duties men and women performed on the job. Indeed, on more than one occasion a labor economist has suggested to me that the kind of uniform sex segregation within occupations that Baron and I detected simply cannot exist in the real world!

Radical and Marxist economists who study discrimination have always been more open than their more orthodox colleagues to incorporating concepts and findings of sociologists. Furthermore, the kind of persistent job segregation by sex detected in my research with Baron is fully consistent with their models. Radical and Marxist models are built upon an imagery of conflict and struggle over control of the workplace among contending groups defined by race, class, and gender. In attempting to account for the persistence of the racial gap in unemployment, Shulman [1991] moves this line of theorizing forward by addressing the economic and social costs of dismantling discrimination. According to this perspective, job segregation and other forms of discrimination are likely to persist in the absence of any exogenous shock or direct intervention that changes the social relations among contending groups.

However, while radical and Marxist approaches appear more consistent with the findings of our research, I am not yet prepared to accept their models. My reservation is based on the absence of gender as a cultural or ideological phenomenon in their explanations of segregation and other forms of discrimination. Instead, their models are driven by the logic of systemic properties like concentration and monopolization and the self-interested, rational action of social groups (e.g., employers, white male workers) intent on preserving their privileged positions in the division of labor (Darity and Williams [1985]). Based on my research and my reading of the work of other sociologists, I have come to believe that gender is more than a dimension along which the labor force is segmented and that gender ideologies are a strong, semiautonomous force shaping segregation and other manifestations of socioeconomic inequality. I have yet to see this notion of gender and ideology incorporated into formal models of discrimination, but I can illustrate the empirical reality of the issue with two examples.

The first concerns the circumstances under which women were demobilized from factory jobs in the automobile industry after World War II. In the most definitive study to date on the topic, Milkman [1987] documented that the women who held these jobs performed at least as capably as the men they replaced. She also showed that many of the women mobilized by the automobile industry during the war wanted to keep their jobs. Moreover, at the war's end, the automobile manufacturers could no longer rely on "cost-plus" government contracts and therefore should have been more sensitive to labor costs after the war than during the war. Furthermore, Milkman showed that the unions that represented the interests of white male workers had virtually no power to influence management policy regarding hiring preferences and job assignments. In short, management appeared to have both the incentive and the ability to rely disproportionately on women during the tremendous expansion in employment during the years following the war.

Such behavior by employers would be consistent with both neoclassical and neo-Marxist approaches to the labor market. Instead, Milkman showed that management acted on a strong preference for young, male veterans and went to great lengths to purge women from all but the traditionally female job classifications. Perhaps management feared that female substitution in the automobile factories would destabilize labor relations and were taking into account one of the costs of dismantling discrimination emphasized in Shulman [1991]. However, there is little evidence that the very decisive actions taken by management to demobilize women were based on an informed view of the costs of possible resistance by male workers. Instead, for reasons that are difficult to explain from any perspective emphasizing economic self-interest (or class interest), employers concluded that women ought not to hold these jobs once men returned from war.

My second example comes from research I have done in collaboration with Denise Bielby (Bielby and Bielby [1991]) empirically testing Jacob Mincer's [1978] neoclassical model of family migration decisions among dual-earner couples. According to the model, couples maximize family well-being. When a husband or wife is confronted with a job advancement requiring geographic relocation, the couple migrates when that person's gain is greater than the cost of disrupting his or her spouse's career. Otherwise, there is no net gain to the family, and family well-being is maximized by forgoing the job opportunity at the new location.

Presumably, the risk to joint family earnings due to a relocation increases with the level of earnings of the spouse whose career will be disrupted. If so, then the neoclassical model predicts that reluctance to relocate for a new job should increase directly with the level of the spouse's earnings. Accordingly, it was not surprising when data from the 1977 Quality of Employment Survey showed spouse's earnings to be the strongest predictor of a reluctance to relocate for a much better job in a new location. What was surprising (from the neoclassical perspective) was that the net effect of spouse's earnings was highly contingent upon both gender and gender-role beliefs. As Figure 2 illustrates, traditional males—those who believed in the primacy of the husband's role as provider and who disapproved of working mothers—were not influenced at all by their wives' earnings. In contrast, traditional females were extremely sensitive to their husbands' earnings. In effect, such women expressed reluctance to relocate for job advancement unless their husbands earned very little. Holding nontraditional gender-role beliefs had opposite effects on men and women. Compared to traditional women, nontraditional women were more inclined to pursue job advancement at a new location, while the opposite was true for nontraditional men compared to traditional men. In other words, our results

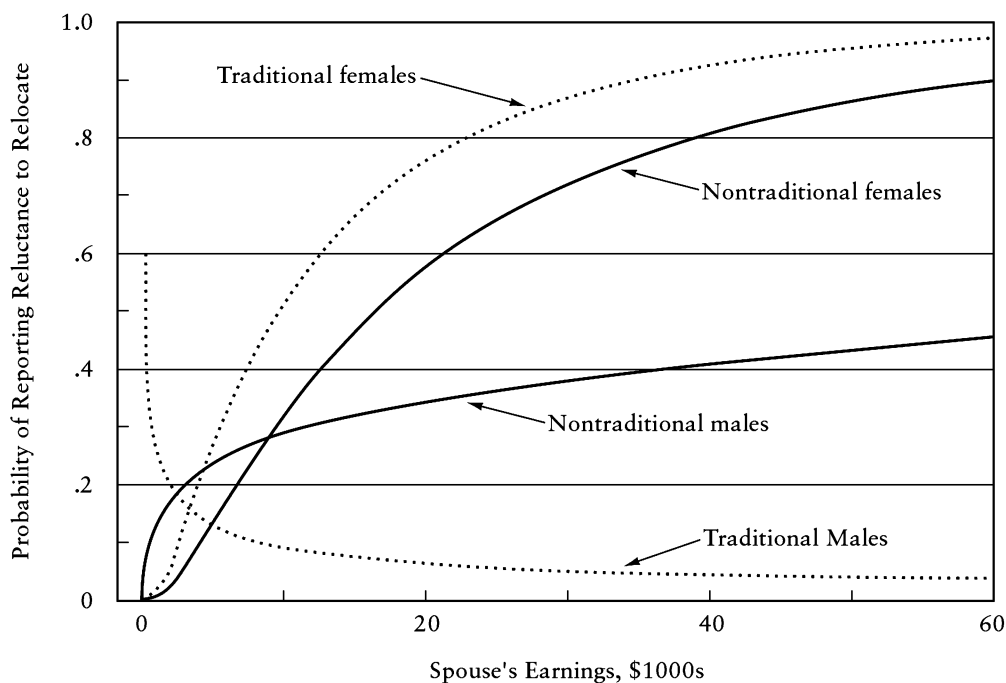


Figure 2. Reluctance to relocate due to family considerations, by sex, spouse's earnings, and gender-role beliefs

showed that among those holding traditional beliefs, men were likely to put their private career interests ahead of the interest of overall family economic well-being (contrary to the assumption of the neoclassical model), while women behaved altruistically toward the family, maximizing family well-being instead of personal well-being (consistent with Mincer's assumption). However, the relationship was sharply mediated by gender-role beliefs, and the sex differences in orientation toward self and family were substantially attenuated among those holding nontraditional beliefs.

Thus our research suggests that traditional gender-role ideologies sustain women's disadvantages in both workplace and family dynamics. Traditional notions of women's role provide a rationale for job segregation and statistical discrimination in the workplace, and those same beliefs make it more difficult for women than men to pursue job opportunities (which in turn can reinforce stereo-

types and provide a rationale for statistical discrimination). In short, gender-role ideologies appear to directly shape the extent to which husbands and wives define and act upon their interests.

In sum, in an important sense, the economic theories of segregation and discrimination—neoclassical supply-side, demand-side, radical, and neo-Marxist—are neutral with respect to gender (and race). In these theories, whichever group happens to have fewer resources or less predictable traits or arrived last when either capital or "good jobs" were being allocated happens to end up disadvantaged. The ideological content of cultural idioms of gender and race have no causal impact in any of these theories. Curiously, this is true even when these theories are borrowed by sociologists, who, in recent years, have been captivated by theories that purport to explain social behavior as the outcome of "rational action" (Hechter [1983]; Coleman [1990]). Yet the examples I have

cited suggest that culture and ideology might have considerable explanatory power in more fully accounting for both persistence and change in job segregation and other forms of discrimination. At least at a descriptive level, sociocultural factors are impossible to ignore in any close empirical investigation of these phenomena. However, it is quite clear that we have a long way to go before such factors are fully incorporated into our theoretical and analytical models of discrimination.

REFERENCES

- Aigner, Dennis J. and Glen C. Cain. 1977. "Statistical Theories of Discrimination in Labor Markets." *Industrial and Labor Relations Review* 30, 2 (January): 175–187.
- Akerlof, George A. and Janet L. Yellen. 1988. "Fairness and Unemployment." *American Economic Review* 78, 2 (May): 44–49.
- Ashmore, Richard D. and Frances K. Del Boca. 1981. "Conceptual Approaches to Stereotypes and Stereotyping." Pp. 1–35 in D. L. Hamilton (ed.), *Cognitive Procession Stereotyping and Inter-group Behavior*. Hillsdale, N.J.: Erlbaum.
- Austin, William. 1977. "Equity Theory and Social Comparison Processes." Pp. 279–305 in J. Suls and R. Miller (eds.), *Social Comparison Theory: Theoretical and Empirical Perspectives*. Washington, D.C.: Hemisphere Publishing.
- Baron, James N. and William T. Bielby. 1985. "Organizational Barriers to Gender Equality: Sex Segregation of Jobs and Opportunities." Pp. 233–251 in A. S. Rossi (ed.), *Gender and the Life Course*. New York: Aldine.
- Berger, Joseph, Susan J. Rosenholtz, and Morris Zelditch. 1980. "Status Organizing Processes." *Annual Review of Sociology* 6: 479–508.
- Berger, Joseph, Morris Zelditch, Jr., Bo Anderson, and Bernard P. Cohen. 1972. "Structural Aspects of Distributive Justice: A Status Value Formulation." Pp. 119–146 in J. Berger, M. Zelditch, and B. Anderson (eds.), *Sociological Theories in Progress*, vol. 2. Boston: Houghton Mifflin.
- Bielby, William T. and James N. Baron. 1984. "A Woman's Place Is with Other Women: Sex Segregation within Organizations." Pp. 27–55 in B. F. Reskin (ed.), *Sex Segregation in the Workplace: Trends, Explanations, Remedies*. Washington, D.C.: National Academy Press.
- . 1986. "Men and Women at Work: Sex Segregation and Statistical Discrimination." *American Journal of Sociology* 91, 4 (January): 759–799.
- . 1987. "Undoing Discrimination: Comparable Worth and Job Integration." In Christine Bose and Glenna Spitze (eds.), *Ingredients for Women's Employment Policy*. Albany: State University of New York Press.
- Bielby, William T. and Denise D. Bielby. 1988. "She Works Hard for the Money: Household Responsibilities and the Allocation of Work Effort." *American Journal of Sociology* 93, 5 (March): 1031–1059.
- . 1991. "I Will Follow Him: Family Ties, Gender-Role Beliefs, and Reluctance to Relocate for a Better Job." Forthcoming, *American Journal of Sociology* 94.
- Becker, Gary S. 1985. "Human Capital, Effort, and the Sexual Division of Labor." *Journal of Labor Economics* 3, 1, pt. 2 (January): 33–58.
- Callahan-Levy, Charlene M. and Lawrence A. Messe. 1979. "Sex Differences in the Allocation of Pay." *Journal of Personality and Social Psychology* 37, 3 (March): 433–446.
- Coleman, James S. 1990. *Foundations of Social Theory*. Cambridge, Mass.: Harvard University Press.
- Crosby, F. 1982. *Relative Deprivation and Working Women*. New York: Oxford University Press.
- Darity, William A., Jr. 1991. "Efficiency Wage Theory: Critical Reflections on the Neokeynesian Theory of Unemployment and Discrimination." Pp. 39–53 in R.R. Cornwall and Phanindra V. Wunnava (eds.), *New Approaches to Economic and Social Analyses of Discrimination*. New York: Praeger Publishers.
- Darity, William A. and Rhonda A. Williams. 1985. "Peddlars Forever? Culture, Competition, and Discrimination." *American Economic Review: Papers and Proceedings* 75, 2 (May): 256–261.
- Darley, J. M. and R. H. Fazio. 1980. "Expectancy Confirmation Sequences." *American Psychologist* 35, 10 (October): 867–881.
- Hechter, Michael. 1983. *The Microfoundations of Macrosociology*. Philadelphia: Temple University Press.
- Krueger, Alan B. and Lawrence H. Summers. 1987. "Reflections on the Inter-Industry Wage Structure." Ch. 2, pp. 17–47, in K. Lang and J. S. Leonard (eds.), *Unemployment and the Structure of Labor Markets*. New York: Basil Blackwell.
- Lenney, Ellen. 1977. "Women's Self-Confidence in Achievement Settings." *Psychological Bulletin* 84, 1 (January): 1–13.

- Major, Brenda, Dean B. McFarlin, and Diana Gagnon. 1984. "Overworked and Underpaid: On the Nature of Gender Differences in Personal Entitlement." *Journal of Personality and Social Psychology* 47, 6 (December): 1399–1412.
- Marks, Stephen R. 1977. "Multiple Roles and Role Strain: Some Notes on Human Energy, Time, and Commitment." *American Sociological Review* 42, 6.
- Milkman, Ruth. 1987. *Gender at Work: The Dynamics of Job Segregation by Sex During World War II*. Urbana: University of Illinois Press.
- Mincer, Jacob. 1978. "Family Migration Decisions." *Journal of Political Economy* 86, no. 5 (October): 749–773.
- Phelps, E. S. 1972. "The Statistical Theory of Racism and Sexism." *American Economic Review* 62, 4 (September): 659–661.
- Polachek, Solomon W. 1979. "Occupational Segregation Among Women: Theory, Evidence, and a Prognosis." Pp. 137–157 in C. B. Lloyd, E. S. Andrews, and C. L. Gilroy (eds.), *Women in the Labor Market*. New York: Columbia University Press.
- Polachek, Solomon W. and Charng Kao. 1991. "Lifetime Work Expectations and Estimates of Gender Discrimination." Pp. 199–238 in R.R. Cornwall and P.V. Wunnava (eds.), *New Approaches to Economic and Social Analyses of Discrimination*. New York: Praeger Publishers.
- Reskin, Barbara F. and Heidi I. Hartmann (eds.) 1986. *Women's Work, Men's Work: Sex Segregation on the Job*. Washington, D.C.: National Academy Press.
- Shulman, Steven. 1991. "Why Is the Black Unemployment Rate Always Twice as High as the White Unemployment Rate?" Pp. 5–37 in R.R. Cornwall and P.V. Wunnava (eds.), *New Approaches to Economic and Social Analyses of Discrimination*. New York: Praeger Publishers.
- Williamson, Oliver E. 1975. *Markets and Hierarchies*. New York: Free Press.
- . 1985. *The Economic Institutions of Capitalism*. New York: Free Press.

85 ■ Jerry A. Jacobs

Revolving Doors

Sex Segregation and Women's Careers

This chapter is an effort to understand why women do "women's work." In 1985, over two-thirds of the women in the U.S. civilian labor force worked in occupations that were 70 percent or more female.¹ The study of women's work has attracted scholars from many disciplines, from psychology (Gutek, 1985) to history (Kessler-Harris, 1982) to anthropology (Sanday, 1981) to economics (Bergmann, 1986) to sociology (Reskin, 1984). A flurry of recent papers has addressed such topics as trends in occupational segregation (Bianchi and Rytina, 1986; Ja-

cobs, 1986), sex segregation among teenagers in the workplace (Greenberger and Steinberg, 1983), and sex segregation within voluntary organizations (McPherson and Smith-Lovin, 1986), between industries (Tienda, Smith, and Ortiz, 1987), across cities (Abrahamson and Sigelman, 1987), and within firms (Bielby and Baron, 1986). (See Reskin and Hartmann, 1985, for discussion of sex segregation in the workplace.)

This interest has undoubtedly been spurred by the growth in women's labor force participation (Goldin, 1983; Smith and

Jerry A. Jacobs. *Revolving Doors: Sex Segregation and Women's Careers*, pp. 1–4, 8–10, 50–53, 190. Copyright © 1989 by the Board of Trustees of the Leland Stanford Junior University. Reprinted with the permission of the publishers, Stanford University Press.

Ward, 1984). Over 70 percent of women between the ages of 20 and 44 worked for pay in 1985, including nearly half of those with children under one year old. The overall rate of women's labor force participation has grown from 33.9 percent in 1950 to 54.5 percent in 1985 (U.S. Bureau of the Census, 1986). The women's movement has also heightened interest in working women. Legislative policies such as affirmative action, as well as the entry of feminist scholars into academia, have generated research on the disadvantaged position of working women. Most recently, calls for equal pay for work of comparable value have brought public and research attention to the concentration of women in low-paid, female-dominated occupations (Remick, 1984).

This chapter investigates the way women's careers intersect with the sexual division of labor in the workplace. We follow women's (and men's) life histories in order to understand the gender tracking system. We ask whether there is a decisive moment during women's lives when they are channeled into female-dominated occupations, or whether the process is more gradual. Is sex segregation principally due to a single life-course event, such as meeting with a guidance counselor in high school or becoming a parent, or is it the result of a cumulative process extending throughout much of the women's (and men's) lives?

As the title of this chapter suggests, we find surprising rates of mobility for women among male-dominated, sex-neutral, and female-dominated occupations (see Jacobs, 1989). This pattern of change is also evident when the dynamics of sex segregation of career aspirations and college majors are analyzed. Individual mobility is common, yet change in the structure of sex segregation is quite slow. This paradox is the central puzzle of this chapter.

The mobility described here and elsewhere (Jacobs, 1989) does not imply the existence of equal opportunity for women. Quite the

contrary. The rates of mobility we document are roughly proportionate to the distribution of *women* across occupations; they do not imply that women have the same chances as men. For example, the probability of a woman starting her career in a male-dominated occupation is approximately one in three for recent cohorts of women entering the labor force. Our data indicate that, for a woman who changes her occupation, the probability of moving into a male-dominated occupation is close to one in three, no matter whether her previous occupation was in a female-dominated, sex-neutral, or male-dominated field (Jacobs, 1989, chapter 7). This pattern suggests that women in female-dominated occupations who are able to change occupations stand a good chance of moving into a male-dominated occupation. Yet by the same token, women in male-dominated fields who change occupations also stand a good chance of moving to a sex-neutral or female-dominated occupation. Indeed, one of the more disturbing findings documented here (also, see Jacobs, 1989, chapter 7) is a startling rate of attrition of women in male-dominated occupations, even as these occupations appear to be opening their doors to them.

The revolving-door metaphor does not imply that women are getting nowhere. Rather, it suggests that gross mobility far exceeds the overall net change in opportunities for women. In recent years, for every 100 women in male-dominated occupations who were employed in two consecutive years, 90 remained in a male-dominated occupation, while 10 left for either a sex-neutral or female-dominated occupation. At the same time, 11 entered a male-dominated occupation from one of these other occupation groups. Thus, the revolving door sends 10 out for every 11 it lets in.² There has been a slow net accumulation of women in male-dominated occupations, but the net change is small compared to the size of the flows in both directions. If there were less attrition of women from male-dom-

inated occupations, there would be more progress toward integrating men and women in the workplace. . . .

Sex Segregation and Social Control

To maintain any sexual division of labor, a society has to develop mechanisms for slotting men into male-dominated positions and women into female-dominated positions. Once positions or tasks are defined as masculine or feminine, societies must enforce these definitions. Sociologists refer to mechanisms by which societies bring about these results as mechanisms of social control.

The earliest stage of social control is socialization, during which society imbues the young with its prevailing values and beliefs. Central to our interest is vocational socialization, the development of gender identity that bears on the pursuit of careers. While the reproduction of social values may seem a simple matter in preindustrial societies, socialization is problematic in our own society. We argue that social control continues well beyond early socialization, involving the educational system, the decisions and behavior of employees and employers in the labor market, and the influences of family and friends throughout life. We raise questions about the efficacy of various gender-tracking mechanisms that are often assumed to be responsible for the maintenance of sex segregation.

The central thesis of this chapter is that the maintenance of sex segregation depends on a lifelong system of social control. While sex-role socialization is important, since it instills values and goals, it is inadequate by itself to maintain the system of sex segregation. Vocational aspirations have short life-expectancies, and young women frequently aspire to female-dominated occupations at one point in time and male-dominated jobs at another (see Jacobs, 1989, chapter 5). Further, the relationship between the sex type of young women's vocational goals and the sex type of the jobs they obtain is weak (see Jacobs,

1989, chapter 5). These patterns suggest that the influence of sex-role socialization is limited. Thus, sex-role socialization can be viewed as a system of social control for the early years, necessary but not sufficient to account for the persistence of occupational sex segregation.

For those men and women who attend college, social control continues during higher education. Many consider the educational system to be the realm of universalism, the paramount social institution in our society for promoting opportunities for all. Indeed, the declines in sex segregation in higher education have surpassed those in the labor force. Yet, as shown in Jacobs (1989, chapter 6), important gender differences persist in America's colleges. Men and women pursue different majors, and informal social control plays a prominent role on campus.

The field of study in college is an important but not decisive component of gender tracking. Segregation remains fluid during the college years as many students move among male-dominated, sex-neutral, and female-dominated majors. Further, the link between a person's major in college and subsequent occupational pursuits is far from direct for many fields. Thus, the sex segregation of college majors resembles in both structure and process the segregation of aspirations and the segregation of occupations. Like sex-role socialization, sex segregation in the educational system is a necessary but not sufficient cause of gender inequality in the labor force.

Sex segregation in the labor force is the most durable of any context we consider, yet sex-role assignments are not fixed when women and men enter the labor force. Many women move among male-dominated, sex-neutral, and female-dominated occupations. The entry of women into male-dominated fields often does not immediately follow schooling, but instead may follow a period of employment in a female-dominated field or a spell as a homemaker. The patterns of stability

and change in sex segregation are strikingly similar in aspirations, college majors, and careers, at both the aggregate and the individual level.

The problem, then, is accounting for the persistence of sex segregation at the societal or macro level despite its instability or permeability at the individual or micro level. Our resolution of this paradox is that sex roles are subject to continual, systematic, but imperfect control throughout life. Individuals move, but the system remains segregated, owing to the cumulative force of social pressures. . . .

Cumulative Disadvantage Versus Revolving Doors

Two competing versions of the lifelong social control perspective must be distinguished. The more familiar is what we label the “cumulative disadvantage” view, which we contrast with the “revolving door” perspective. The cumulative disadvantage model holds that occupational segregation by sex is the result of the accumulation of obstacles women face in their pursuit of careers. Thus, socialization is compounded by educational inequality and labor market discrimination, the cumulative force of which leaves women in a segregated set of occupations.

This view is clearly implicit in Berryman’s discussion of women in science and engineering (1983). Women face obstacles beginning with ambivalent socialization about the importance of careers and the stereotype that science is an unfeminine pursuit. Women are underrepresented in science and engineering professions because of inadequate preparation in high school, attrition from science and engineering fields in college, and minority status and hostility on the job for those women who pursue such careers.

If the obstacles women confront in pursuing science and engineering careers are characteristic of male-dominated occupations, this model implies a number of specific predictions concerning sex segregation. The first

is that segregation would increase over women’s lifetimes as women continually leave male-dominated pursuits. One would expect to find more segregation as a cohort ages, as the disadvantages women face accumulate. Second, mobility observed between male-dominated and female-dominated occupations would take the form of attrition of women from male-dominated pursuits. Delayed entry into male-dominated occupations would be difficult; at the very least, delayed entries of women into male-dominated occupations should be less frequent than exits. Finally, if declines in sex segregation were to occur, this change would primarily take the form of changes between cohorts. The best one could expect for those already in the pipeline is to reduce the rate of attrition; real change must await a new cohort that is exposed to different socialization, broader opportunities in school, and less hostility on the job. One would not expect increased integration among those already in the labor force, since the cumulative effects of these social control processes are essentially irreversible.

While the cumulative disadvantage model is attractive because of its recognition of the pervasive nature of social control, three striking findings suggest that while social control is continuous, its effects do not increase over the life course. First, occupational segregation by sex does not increase as men and women grow older (see Jacobs 1989, chapter 2). Second, sex-type mobility involves substantial rates of midcareer entry into, as well as exit from, male-dominated occupations (see Jacobs 1989, chapter 7). Third, overall change in segregation is observed across different age groups, and is not simply the result of cohort replacement (see Jacobs 1989, chapter 2). In other words, where the cumulative disadvantage view holds that change must come about principally because a new cohort has different experiences, the evidence suggests that a general weakening of the system of controls broadens opportunities for those already in the labor force. There is as

much change in young women's aspirations over time as there is across new cohorts of women in the process of forming their aspirations. There is as much change for people already in college as there is among those entering college. In recent years, there has been a notable decline in segregation among those already in the labor market (see Jacobs 1989, chapter 2). The insight of the cumulative disadvantage view is that it recognizes the range of constraints imposed on women throughout their lives. Its limitation is that it assumes that the effects of barriers to women's opportunities are permanent and irreversible.

Thus we arrive at a lifelong social control perspective that is not characterized by cumulative disadvantage. In recognition of the dramatic levels of mobility we find, we refer to this view as the "revolving door" model of sex segregation. The revolving door perspective can account for extensive sex-type mobility because it recognizes a variety of stages in the career development process, and it recognizes a host of pressures women face. Most women will face one or more barriers to the pursuit of a career at some point; they are likely to overcome some of these and not others. Thus there are reasons to expect substantial flows of women into and out of male-dominated occupations, while overall the system changes only gradually.

This perspective also accounts for the patterns of change in sex segregation we are beginning to see. If social control is exercised throughout life, social change is likely to be experienced by women of different ages. Change should affect those already in the labor force as well as those entering it and those in the educational system. Theories that emphasize labor market entry as the decisive time for segregation cannot account for decreasing segregation among those already in the labor force. As shown elsewhere (Jacobs 1989, chapter 2), the declines in labor market segregation by sex in the 1970s occurred throughout the age structure and were not simply restricted to labor market entrants.

There is a striking similarity in the dynamics of sex segregation in the development of aspirations, in the segregation of college majors, and in the sex segregation of the occupational structure. All three are characterized by change occurring among different age groups, and all three have striking levels of mobility between male-dominated, sex-neutral, and female-dominated fields. The similarities of these patterns in these diverse contexts and across different stages of life convince us of the utility of the social control/revolving door perspective.

NOTES

1. Calculated from data in *Employment and Earnings* (U.S. Bureau of Labor Statistics, 1986).
2. These figures reflect the analysis of Current Population Survey data discussed in Jacobs (1989, chapter 7).

REFERENCES

- Abrahamson, Mark, and Lee Sigelman. 1987. "Occupational Sex Segregation in Metropolitan Areas," *American Sociological Review*, 52(5): 588–97.
- Bergmann, Barbara. 1986. *The Economic Emergence of Women*. New York: Basic Books.
- Berryman, Sue R. 1983. *Who Will Do Science?* New York: Rockefeller Foundation.
- Bianchi, Suzanne M., and Nancy Rytina. 1986. "The Decline in Occupational Sex Segregation During the 1970s: Census and CPS Comparisons," *Demography*, 23(1): 79–86.
- Bielby, William T., and James N. Baron. 1986. "Men and Women at Work: Sex Segregation and Statistical Discrimination," *American Journal of Sociology*, 91(4): 759–99.
- Goldin, Claudia. 1983. "The Changing Economic Role of Women: A Quantitative Approach," *Journal of Interdisciplinary History*, 13(4): 707–33.
- Greenberger, Ellen, and Laurence D. Steinberg. 1983. "Sex Differences in Early Labor Force Experience: Harbinger of Things to Come," *Social Forces*, 62(2): 467–86.
- Gutek, Barbara. 1985. *Sex and the Workplace*. San Francisco: Jossey-Bass.
- Jacobs, Jerry A. 1986. "Trends in Contact Between Men and Women at Work, 1971–1981," *Sociology and Social Research*, 70(3): 202–6.

- _____. 1989. *Revolving Doors: Sex Segregation and Women's Careers*. Stanford, CA: Stanford University Press.
- Kessler-Harris, Alice. 1982. *Out to Work: A History of Wage-Earning Women in the United States*. New York: Oxford University Press.
- McPherson, J. Miller, and Lynn Smith-Lovin. 1986. "Sex Segregation in Voluntary Associations," *American Sociological Review*, 51(1): 61–79.
- Remick, Helen, ed. 1984. *Comparable Worth and Wage Discrimination: Technical Possibilities and Political Realities*. Philadelphia: Temple University Press.
- Reskin, Barbara, ed. 1984. *Sex Segregation in the Workplace: Trends, Explanations, Remedies*. Washington, D.C.: National Academy of Sciences.
- Reskin, Barbara F., and Heidi I. Hartmann. 1985. *Women's Work, Men's Work: Sex Segregation on the Job*. Washington, D.C.: National Academy of Sciences.
- Sanday, Peggy R. 1981. *Female Power and Male Dominance: On the Origins of Sexual Inequality*. Cambridge: Cambridge University Press.
- Smith, James P., and Michael P. Ward. 1984. *Women's Wages and Work in the Twentieth Century*. Santa Monica, Calif.: Rand Corporation.
- Tienda, Marta, Shelley A. Smith, and Vilma Ortiz. 1987. "Industrial Restructuring, Gender Segregation, and Sex Differences in Earnings," *American Sociological Review*, 52(2): 195–210.
- U.S. Bureau of the Census. 1986. "Women in the American Economy," *Current Population Reports*, P-23, no. 146. Washington, D.C.: U.S. Government Printing Office.
- U.S. Bureau of Labor Statistics. 1986. *Employment and Earnings*. Washington, D.C.: U.S. Government Printing Office.

86 ■ Barbara F. Reskin

Labor Markets as Queues

A Structural Approach to Changing Occupational Sex Composition

The historic persistence of sex segregation (Gross 1968; Beller 1984) made women's dramatic gains during the 1970s in such diverse male occupations as pharmacy, bank management, bartending, and typesetting noteworthy. Why did women disproportionately enter these and a few other male occupations when they made only modest progress in most and lost ground in a few? To answer this question, Patricia Roos and I conducted a two-part study: multivariate analyses of the changing sex composition in all detailed census occupations and in-depth case studies of 14 occupations in which women's representa-

tion increased at least twice as much as it had in the labor force as a whole (see Table 1). This chapter draws on those case studies: bank managers (Bird 1990), bartenders (Detman 1990), systems analysts (Donato 1990a), public relations specialists (Donato 1990b), pharmacists (Phipps 1990a), insurance adjusters/examiners (Phipps 1990b), typesetters/compositors (Roos 1990), insurance salespersons (Thomas 1990), real estate salespersons (Thomas and Reskin 1990), bakers (Steiger and Reskin 1990), book editors (Reskin and Roos 1990), print and broadcast reporters, and accountants.¹ Although the case studies

Barbara F. Reskin. "Labor Markets as Queues: A Structural Approach to Changing Occupational Sex Composition," in *Micro-macro Linkages in Sociology*, edited by Joan Huber, pp. 170–186, 188–192. Copyright © 1991 by Sage Publications, Inc.

Table 1. Percent Females in Feminizing Occupations, 1970, 1980, 1988

| <i>Occupation</i> | <i>1970</i> % | <i>1980</i> % | <i>1988</i> % | <i>1970–1988 Increase</i> % |
|-------------------------------|------------------|------------------|------------------|--------------------------------|
| Financial managers | 19.4 | 31.4 | 42.4 | 23.0 |
| Accountants/auditors | 24.6 | 38.1 | 49.6 | 25.0 |
| Operation/systems analysts | 11.1 | 27.7 | 39.6 | 28.5 |
| Pharmacists | 12.1 | 24.0 | 31.9 | 19.8 |
| Public relations specialists | 26.6 | 48.8 | 59.1 | 32.5 |
| Editors and reporters | 41.6 | 49.3 | n.a. | n.a. |
| Insurance sales | 12.9 | 25.4 | 29.7 | 16.8 |
| Real estate sales | 31.2 | 45.2 | 48.5 | 17.3 |
| Insurance adjusters/examiners | 29.6 | 60.2 | 72.2 | 42.6 |
| Bartenders | 21.2 | 44.3 | 49.6 | 28.4 |
| Bakers | 25.4 | 40.7 | 47.8 | 22.4 |
| Bus Drivers | 28.3 | 45.8 | 48.5 | 20.2 |
| Typesetters/compositors | 16.8 | 55.7 | 73.9 | 57.1 |
| Labor force | 38.0 | 42.6 | 45.0 | 7.0 |

revealed that widely different factors precipitated these occupations' feminization, the process of change can be encompassed in a single perspective—queuing. This chapter begins by outlining the queuing approach. It then shows how the determinants of occupational feminization conformed to queuing processes. Finally, it argues that the queue model as I have developed it offers a structural approach to understanding change in job composition.

Queuing: An Overview

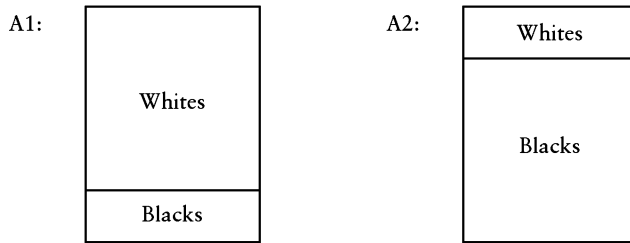
Simply stated, a queuing perspective views labor markets as composed of labor queues and job queues that reflect, respectively, employers' ranking of possible workers and workers' ranking of jobs.² The idea of the *labor* queue originated in Lester Thurow's (1969) work on race differences in unemployment and was elaborated by Hodge (1973) and Lieberman (1980). Rotella's (1977) analysis of the feminization of clerical work, Lieberman's (1980) study of racial and ethnic inequality, and Strober's (1984) theory of sex segregation implicitly invoked *job* as well as *labor* queues. The queuing perspective presented here views occupational composition as resulting from the simultaneous

operation of job and labor queues (also see Reskin and Roos 1990). Employers hire workers from as high in the labor queue as possible, and workers accept the best jobs available to them, so the most desirable jobs go to the most preferred workers, less attractive jobs go to workers lower in the labor queue, and the most lowly workers end up jobless or in jobs others have rejected.

Three structural properties characterize both job and labor queues: (a) the ordering of their elements, (b) whether or not their elements overlap, and (c) their shape. By *elements*, I mean the constituent units—groups of potential workers or jobs. Thus the *ordering of elements* simply indicates in what order workers rank possible jobs and employers rank prospective workers. I pass over how workers order jobs in the labor queue as something both sociologists and workers understand. In ordering the labor queue, employers consider potential productivity and cost.³ However, the indirect effect of workers' sex on employers' estimations of the former and its direct influence on how employers rank workers result in labor queues that encompass and often act as *gender queues*.

The absolute and relative numbers of elements in a queue determine its *shape*. Thus the number of prospective workers in each

Panel A. Hypothetical labor queues ordered by race for predominantly white and predominantly black labor markets, respectively



Panel B. Hypothetical job queues ordered by nonmanual-manual work for predominantly nonmanual and predominantly manual occupational structures, respectively

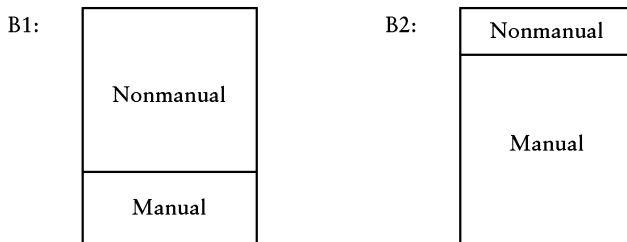


Figure 1. Illustration of Variation in the Shape of Labor and Job Queues

subgroup in a labor market sets the shape of the labor queue, and the number of jobs at each level in the job queue fixes its shape. Panels A and B of Figure 1 show how the shape of labor and job queues can vary while their order remains constant. This variation influences each group’s probable access to occupations of varying desirability. (Put differently, it shows how a queue’s shape influences each occupation’s chance of recruiting workers from particular groups.) For example, in a society with relatively few workers in the preferred group (A2) and few desirable jobs (B2), preferred groups will monopolize better or worse jobs than persons from their group normally command. In consequence, when good jobs sharply outnumber highly ranked workers as in panels A2 and B1, employers must fill the better jobs with workers from lower in the labor queue than they usually hire. In contrast, when the *job queue* is bottom heavy, only the highest-ranked workers get desirable jobs, and workers moderately high in the labor queue are forced to

settle for less. As Lieberman (1980, p. 297) has shown, both the *absolute* and the *relative* sizes of each group in the labor queue affect lower-ranked workers’ chances of getting desirable jobs. The larger a subordinate group relative to the size of the preferred group, the more costly it is for employers to deny its members good jobs (Hodge 1973).

Whether or not *elements overlap* denotes the strength of rankers’ preferences for one element over another. Among some employers, group membership is the overriding consideration in ordering the labor queue, and rankers favor persons from the group they prefer, regardless of qualifications. For other labor queues, employers are indifferent to group membership except to break ties between otherwise equally qualified prospects. Figure 2 illustrates variation in the intensity of raters’ preferences with respect to workers’ race in three hypothetical labor queues. The space between the races in panel A depicts the ranking for employers who invariably hire the lowest-ranked white worker over the

Panel A. Racial group membership is an overriding consideration to rankers. Employers hire applicants as qualified as possible, but hire unqualified whites before hiring highly qualified blacks.

| | | | | | |
|---------------|--------------------|----------------|---------------|--------------------|----------------|
| Blacks Low | Blacks Moderate | Blacks High | Whites Low | Whites Moderate | Whites High |
|---------------|--------------------|----------------|---------------|--------------------|----------------|

Level of Qualification

Panel B. Racial group membership is a minor consideration to rankers. Employers hire the most qualified applicants, but within levels of qualification, give white applicants an edge over equally qualified blacks.

| | | | | | |
|---------------|---------------|--------------------|--------------------|----------------|----------------|
| Blacks Low | Whites Low | Blacks Moderate | Whites Moderate | Blacks High | Whites High |
|---------------|---------------|--------------------|--------------------|----------------|----------------|

Level of Qualification

Panel C. Racial group membership is an intermediate consideration to rankers. Employers will prefer a more qualified white over a less qualified black, but will hire very qualified blacks over unqualified whites.

| | | | | | |
|---------------|--------------------|---------------|----------------|--------------------|----------------|
| Blacks Low | Blacks Moderate | Whites Low | Blacks High | Whites Moderate | Whites High |
|---------------|--------------------|---------------|----------------|--------------------|----------------|

Level of Qualification

Figure 2. Illustration of Variation in the Intensity of Raters' Preferences with Respect to Race

best black worker; the overlapping groups in panel B depict the situation in which raters' aversion to blacks is slight. Panel C illustrates an intermediate situation in which employers prefer white workers over equally qualified blacks but will set aside racial biases to hire very talented blacks over mediocre whites. Where preferences are weak, substantial job-level integration exists. In contrast, when preferences are overriding, all members of the favored group precede all members of the disfavored group in the labor queue, and the labor market is totally segmented.

Workers' job preferences also vary in intensity. Some workers may categorically prefer any job in a high-ranked occupation to

working in a lower occupation (expressed, for example, in rejecting manual jobs in favor of nonmanual work). Alternatively, they may be more attuned to specific job characteristics than their overall rank, occasionally eschewing jobs in occupations usually reserved for the preferred group and appropriating more desirable jobs within lower-ranked occupations that usually go to less preferred groups. Strong preferences mean that workers invariably prefer one type of job (e.g., professional or nonmanual work) over another.

Changes in any of these structural features of queues—employers' rankings of workers or workers' ranking of jobs, the intensity of either groups' preferences, and the relative

distributions of workers or jobs—can transform occupations' composition.

Queuing Models as Structural Approaches to Occupational Composition

Queuing emphasizes how employers rank groups of potential workers as well as how workers rank jobs. Because rankings simply boil down to preferences, is *queuing* not simply a fancy way to say that workers' and employers' preferences affect their labor market decisions? What does the queuing perspective add to the neoclassical economic approach to occupational segregation?

The neoclassical economic approach assumes (a) that workers' labor market decisions represent attempts to maximize their lifetime earnings, (b) that differences in the sexes' distributions across jobs stem primarily from women's family responsibilities, (c) that employers act to maximize revenues, and (d) that, in the short run, discriminatory "tastes" sometimes divert employers from their economic best interests, but, in the long run, competitive disadvantages that accrue from indulging economically irrational tastes lead discrimination to decline over time (Becker 1957). Thus traditional economic analysis treats sex segregation as the *aggregate* outcome of the mostly rational choices of *individual* workers and employers.

A queuing approach to changing occupational sex composition differs in several ways. First, queuing emphasizes the *collective* nature of sex segregation that results from *socially structured rankings by groups in conflict*. In contrast to the neoclassical perspective, queuing highlights the roles played by power and conflict between groups with contradictory interests in shaping occupations' composition. As I show below, the queuing perspective elucidates the role of group power in determining occupational access.

Second, it takes seriously the effects of noneconomic factors on workers' rankings of

occupations and employers' rankings of prospective workers. It reminds us that working conditions, autonomy, social standing, career opportunities, and sex composition influence how workers appraise jobs and that employers' prejudices, stereotypes, and desire to preserve their own and other men's advantages influence how they rank workers.

Third, the queuing perspective assumes that the sexes rank most occupations similarly, so it predicts women's influx into accessible male occupations in view of their superior rewards. In explaining occupational feminization, it directs our attention to changes in employers' need for women and to declining barriers to women's access to traditionally male jobs. In other words, the queuing perspective redirects us from characteristics of female workers to structural properties of labor markets that in turn are shaped by the preferences of employers and male workers.

Power and Coalitions in Ordering the Labor Queue

In examining how conflict affects occupations' composition, we must bear in mind three points: First, three groups with competing interests have a stake in which sex predominates in an occupation—employers (who are overwhelmingly male), male workers, and female workers. Second, the greater social and economic power of employers and of men give them a substantial advantage in the struggle over what sex heads the queue for desirable occupations. Third, the outcome of the struggle over what sex dominates an occupation—occupations' sex compositions—represents a *de facto* coalition between competing groups.

Employers. Given their orientation toward minimizing costs and maximizing labor docility, employers should prefer female workers for all jobs. That so few have done so reminds us that other considerations affect

male employers' labor decisions. Foremost among these is the personal stake of both male employers and male workers in excluding women from male-labeled jobs to preserve sex differentiation and hence male privilege in all spheres. Because sex differentiation both legitimates and facilitates unequal treatment, physically segregating the sexes at work and assigning them different tasks preserves male dominance by legitimating unequal wages, insulating men from female competition, supporting men's dominance in other public realms and the home, and shoring up myths of essential sex differences and male superiority (Reskin 1988). As Cockburn (1988, p. 41) put it, "Behind occupational segregation is gender differentiation, and behind that is male power."

Other considerations inclining employers to preserve good jobs for men include (a) the ability of male workers—and sometimes customers—to penalize employers for hiring women for "men's" jobs (Bergmann and Darity 1981; Strober 1988); (b) organizational inertia that makes following standard practices the course of least resistance; (c) custom—in other words, jobs' sex labels—combined with the uncertainties attendant on radically altering employment practices (Cohn 1985; Strom 1987; Figart and Bergmann 1989, p. 36); (d) gender solidarity with male workers;⁴ and (e) forestalling later threats to their own jobs. As long as employers can operate profitably, preserving male jobs can be a fairly cheap amenity, especially if their competitors also practice male preference (Stolzenberg 1982; Strober 1988). Thus, in reserving male-dominated jobs for men, employers at once appease male workers, who could make hiring women costly; simplify organizational decisions; and shore up the gender hierarchy from which they benefit.

Male Workers. Male workers' disadvantaged economic position vis-à-vis employers and their advantaged sex status divide their interests. They stand to gain in the long run from

acting collectively with *all* workers, but, in the short run, men are vulnerable to job and wage competition from women who must settle for lower wages. Equal job opportunities for women can also cost men domestic benefits they derive from sex inequality in the workplace and can threaten the gender hierarchy from which working men benefit (Hartmann 1976; Reskin 1988). Thus, for the most part, male workers do better in the short run by monopolizing male jobs than by welcoming female coworkers with whom they could collectively fight for better pay. Only when large numbers of women already have their foot in the door, or when labor shortages or outside agencies have imposed integration, are male workers freed to pursue their class interests collectively with female workers (Hirsch 1986; Milkman 1987; Cobble forthcoming). Under such conditions, male unions have tried to organize women and advocated equal pay.

Coalitions Between Employers and Male Workers. Male workers' monopoly of most occupations stems from the de facto coalition between male employers and male workers. Although I use the term *coalition* figuratively, employers' power to hire workers and assign them jobs means that ordinarily male workers can resist women's entry only with their male bosses' active contrivance (Baker 1964; Hartmann 1976; Cockburn 1988). The result of this coalition between male workers and employers is also seen in feminizing occupations in women's confinement to the least desirable jobs, thereby providing male workers a refuge from integration in still-male jobs (Cockburn 1988). For example, when outside pressure forced a large publishing firm to promote women to jobs as editorial assistants, it created a new, higher rank into which it promoted male former editorial assistants (Osterman 1979).

Despite the forces that preserve sex segregation, it broke down in the occupations we studied when male employers or male workers

withdrew from their coalition. Employers never violated a coalition by directly substituting women for men, but they set aside customary practices, their own sex biases, and concern with the reactions of male workers when their firms' survival depended on cutting labor costs or averting unionization. In the occupations we studied, increased competition and the concomitant need to cut costs as well as vulnerability to work disruptions by unions precipitated this action. For example, industry-wide demand for increased profits during the high-inflation 1970s prompted insurance companies to cut costs by standardizing the work of insurance adjusters and examiners. Heightened competition for customers following deregulation forced banks to provide more customer services while trying to increase productivity and keep the lid on wages. In typesetting and composing, the International Typographical Union's frequent recourse to strikes spurred newspaper publishers to develop a printing technology that the union could not control, and then they hired women for the newly clericalized typesetting and composing jobs. It is important, however, that, in occupations that became less attractive to men for reasons not tied to employers' actions (book editing, real estate sales, retail pharmacy, and bartending), employers did not counter by improving male jobs. Employers' economic interests prevailed, and they replaced men with cheaper women.

In sum, the case studies suggested that employers put women ahead of men in labor queues *only* when labor costs became a pressing concern or when action by women or government regulators raised the cost of male preference.⁵ Faced with economic exigencies, employers' willingness to donate profits to support white male privilege went by the wayside, and they struck a bargain with women workers.

Coalitions Between Male Employers and Female Workers. Why women should enter

into de facto coalitions with male employers is no mystery. Disadvantaged relative to both male workers and male employers, women stand to gain from cooperating with either group. Employers are more likely to accept a coalition with women when their numbers are large. Cobble's (forthcoming) comparison of women's efforts early in this century to preserve the right to serve drinks with their efforts to tend bar illustrates the importance of numbers. Women waitresses' moderate representation within union locals gave them the clout to block proposed restrictions against their serving alcoholic drinks. In contrast, women's small numbers doomed their efforts to tend bar until they could invoke the Civil Rights Act. In the 1970s, women capitalized on their numbers in the book publishing, newspaper, broadcasting, and banking industries to challenge their exclusion from male jobs or their confinement in low-status posts. Large numbers mean that women can better spot their concentration in the least desirable jobs and more effectively protest it. Thus women's increased share in labor queues for male occupations contributed to their gains both by increasing the odds that employers would get down to women in the hiring queue and by making them more powerful adversaries.

Resistance by Male Workers. Male workers' stake in monopolizing the best jobs encourages them to oppose employers' efforts to integrate women. To this end, they may try to reduce women's productivity by denying them training and information or withhold their own labor (Doeringer and Piore 1971; Bergmann and Darity 1981).⁶ Whether men win the battle depends on their power base. For example, early in the Industrial Revolution, before strong unions existed, employers could easily feminize jobs (Kessler-Harris 1982, p. 266; Cohn 1985; Milkman 1987; Cockburn 1988). However, once unions were well established, they could preserve their domains. They did so in bartending,

baking, and typesetting and composing as well as many other occupations throughout most of the century.

Obviously, men did not ultimately prevent women's entry in the occupations we studied. In most of those occupations, by abandoning jobs, male workers gave employers the go-ahead to hire women. Sometimes men do not see women as a threat. For instance, although men fought women's integration into clerical jobs in the British railroads (Cohn 1985), they did not resist women's movement in similar jobs in the United States because integration occurred slowly, and Pullman reserved the best jobs for men (Hirsch 1986, p. 33). In rapidly growing occupations, men did not resist women's entry because growth ensured that good jobs were plentiful, and within-occupation segregation assured men's access to the best jobs.

Thus the case studies of accounting and systems analysis revealed no systematic efforts to exclude women. Male incumbents in high-turnover occupations like book editing, bartending, and residential real estate sales had no stake in what happened in the long run. High turnover also means workers were too poorly organized to resist integration. Finally, in reorganizing work, employers had an effective tool in phasing out men. By changing the work process to incorporate female-labeled tasks, employers could disguise the fact that they were filling what had been men's jobs with women (Davies 1975, p. 282; Rotella 1977, pp. 162, 165) and recast them as women's work. For instance, after transforming the work of adjusters, industry trade journals construed the ideal insurance adjuster as a woman.⁷

However, men did effectively resist women's employment in the most desirable specialties within desegregating occupations, such as production baking, commercial real estate, and industrial pharmacy—partly by informal pressures that discouraged women's entry. I suspect, however, that women's con-

tinued exclusion depended more on employer discrimination in job assignment and promotion practices than on male workers' actions.

Conclusions

In summary, the queuing model proposed here holds that (a) employers rank prospective workers in terms of their potential productivity as well as their personal characteristics, but that they are also influenced by current employees and others who can impose costs for hiring or failing to hire women; (b) shortages—whether from job growth or from a job's inability to attract enough customary workers—prompt employers to hire workers from lower in the labor queue; (c) male workers affect women's access to jobs through their ability to preempt jobs, their power to enforce their monopoly over desirable jobs, and their ability, in abandoning jobs, to bestow them on workers lower in the labor queue; and (d) women's search for better jobs leads them to individually and collectively challenge their exclusion from men's jobs and to move into male lines of work that become open. The case study data support these predictions. I caution readers that our selection of case studies of occupations in which women's gains were exceptionally large limits the generalizability of these findings. However, most occupations included specialties or work settings that had feminized at varying rates, so the case studies provide variation on the dependent variable. Nonetheless, the bases for women's movement into occupations not marked by rapid feminization may have differed. Patricia Roos and I addressed this question in statistical analyses of all 503 detailed occupations (Reskin and Roos 1985).

The case studies support two empirical generalizations both of which are consistent with the operation of queuing processes. First, in most of the occupations we studied, after they had become less attractive to men,

employers hired disproportionate numbers of women. Second, within these nominally desegregating occupations, women tended to be relegated to female enclaves, while men retained the most desirable jobs.⁸ These findings confirm Bielby and Baron's (1986) contention that greater balance in the sex composition of *occupations* does not necessarily imply decreased segregation of *jobs* and also establish the fallacy of inferring declines in sex segregation from occupational-level data. Yet social analysts cannot ignore such data, because policy makers and the media use them to assess change, and, as W. I. Thomas said, situations men define as real are real in their consequences. Exaggerated conclusions about women's progress in male occupations support claims that governmental agencies no longer need to intervene in the workplace to ensure equal treatment. Thus it is incumbent on researchers to determine whether or not the trends that superficial comparisons imply are genuine.

NOTES

I am grateful to Ross Boylan, Lowell L. Hargens, Heidi Hartmann, and Patricia A. Roos for their comments on earlier drafts of this chapter and to James N. Baron and Ronnie Steinberg for helpful discussions of the issues it addresses. The larger study that gave rise to the ideas in this chapter is a collaborative project with Patricia A. Roos. The case studies on which this chapter is based appear in our book, *Job Queues, Gender Queues: Explaining Women's Inroads into Male Occupations* (1990).

1. I conducted the last two unpublished studies. Unless otherwise noted, assertions about these occupations are based on the case studies whose full citations are in the references. For a description of the case study method, see Reskin and Roos (1990).

2. This analysis takes for granted the existence of multiple labor markets, each composed of a labor queue of potential workers and a job queue of available jobs.

3. Incumbent workers can affect the ordering of the labor queue, for example, through seniority rules that restrict employers' hiring prerogatives.

4. This should be especially common among bosses who share a similar background with male

workers, previously held the jobs in question, and currently work with or remain friendly with their former coworkers. Two examples are blue-collar supervisors or sales managers.

5. Unless it is backed by threats from regulatory agencies, pressure from women rarely suffices to persuade employers to assign women to male jobs, partly because it elicits opposition from male workers.

6. Employers can rarely forestall resistance by feminizing entire work teams. Seniority rules may prevent their replacing men, they may doubt whether enough qualified women are available, and they may be unable to train women because male unions control training or because training is too long and expensive.

7. Kessler-Harris (1986) and Strober and Arnold (1987) also recounted how banks exploited sex stereotypes to justify hiring women for banking jobs as well as to exclude them from the same jobs.

8. For further evidence on this point, see Chapter 3 in Reskin and Roos (1990).

REFERENCES

- Baker, Ross. 1964. "Entry of Women into Federal Job World—At a Price." *Smithsonian* 8:82–91.
- Becker, Gary. 1957. *The Economics of Discrimination*. Chicago: University of Chicago Press.
- Beller, Andrea. 1984. "Trends in Occupational Segregation by Sex, 1960–1981." Pp. 11–26 in *Sex Segregation in the Workplace: Trends, Explanations, Remedies*, edited by B. Reskin. Washington, DC: National Academy Press.
- Bergmann, Barbara R., and William Darity. 1981. "Social Relations, Productivity, and Employer Discrimination." *Monthly Labor Review* 104:47–49.
- Bielby, William T., and James N. Baron. 1986. "Men and Women at Work: Sex Segregation and Statistical Discrimination." *American Journal of Sociology* 91:759–99.
- Bird, Chloe. 1990. "High Finance, Small Change: Women's Increased Representation in Bank Management." Pp. 145–66 in B. F. Reskin and P. A. Roos, *Job Queues, Gender Queues: Explaining Women's Inroads into Male Occupations*, Philadelphia: Temple University Press.
- Cobble, Dorothy Sue. Forthcoming. "'Drawing the Line': The Construction of a Gendered Workforce in the Food Service Industry." In *Work Engendered*, edited by Ava Baron. Ithaca: Cornell University Press.
- Cockburn, Cynthia. 1988. "The Gendering of Jobs: Workplace Relations and the Reproduction of Sex Segregation." Pp. 29–42 in *Gender Segregation at*

- Work*, edited by S. Walby. Milton Keynes, England; New York: Open University Press.
- Cohn, Samuel. 1985. *The Process of Occupational Sex-Typing: The Feminization of Clerical Labor in Great Britain*. Philadelphia: Temple University Press.
- Davies, Margery W. 1975. "Women's Place Is at the Typewriter: The Feminization of the Clerical Labor Force." Pp. 279–96 in *Labor Market Segmentation*, edited by R. Edwards, M. Reich, and D. Gordon. Lexington, MA: D. C. Heath.
- Detman, Linda. 1990. "Women Behind Bars: The Feminization of Bartending." Pp. 241–56 in B. F. Reskin and P. A. Roos, *Job Queues, Gender Queues: Explaining Women's Inroads into Male Occupations*. Philadelphia: Temple University Press.
- Doeringer, Peter B., and Michael J. Piore. 1971. *Internal Labor Markets and Manpower Analysis*. Lexington, MA: D. C. Heath.
- Donato, Katharine M. 1990a. "Programming for Change? The Growing Demand for Women Systems Analysts." Pp. 167–82 in B. F. Reskin and P. A. Roos, *Job Queues, Gender Queues: Explaining Women's Inroads into Male Occupations*. Philadelphia: Temple University Press.
- _____. 1990b. "Keepers of the Corporate Image: Women in Public Relations." Pp. 129–44 in B. F. Reskin and P. A. Roos, *Job Queues, Gender Queues: Explaining Women's Inroads into Male Occupations*. Philadelphia: Temple University Press.
- Figart, Deborah M., and Barbara Bergmann. 1989. "Facilitating Women's Occupational Integration." Paper prepared for the U.S. Department of Labor, Commission on Workforce Quality and Labor Market Efficiency, American University.
- Gross, Edward. 1968. "Plus ça Change. . . The Sexual Segregation of Occupations over Time." *Social Problems* 16:198–208.
- Hartmann, Heidi I. 1976. "Capitalism, Patriarchy, and Segregation by Sex." *Signs* 1: 137–70.
- Hirsch, Susan E. 1986. "Rethinking the Sexual Division of Labor: Pullman Repair Shops, 1900–1969." *Radical History Review* 3:26–48.
- Hodge, Robert W. 1973. "Toward a Theory of Racial Differences in Employment." *Social Forces* 52:16–31.
- Kessler-Harris, Alice. 1982. *Out to Work*. New York: Oxford University Press.
- _____. 1986. "Women's History Goes to Trial: EEOC vs. Sears, Roebuck and Co." *Signs* 1:767–79.
- Lieberson, Stanley. 1980. *A Piece of the Pie*. Berkeley: University of California Press.
- Milkman, Ruth. 1987. *Gender at Work*. Urbana: University of Illinois Press.
- Osterman, Paul. 1979. "Sex Discrimination in Professional Employment: A Case Study." *Industrial and Labor Relations Review* 32: 451–64.
- Phipps, Polly A. 1990a. "Occupational and Industrial Change: Prescription for Feminization." Pp. 111–28 in B. F. Reskin and P. A. Roos, *Job Queues, Gender Queues: Explaining Women's Inroads into Male Occupations*. Philadelphia: Temple University Press.
- _____. 1990b. "Occupational Resegregation: A Case Study of Insurance Adjusters and Examiners." Pp. 224–40 in B. F. Reskin and P. A. Roos, *Job Queues, Gender Queues: Explaining Women's Inroads into Male Occupations*. Philadelphia: Temple University Press.
- Reskin, Barbara F. 1988. "Bringing the Men Back In: Sex Differentiation and the Devaluation of Women's Work." *Gender & Society* 2:58–81.
- Reskin, Barbara F., and Patricia A. Roos. 1985. "Collaborative Research on the Determinants of Changes in Occupations' Sex Composition Between 1970 and 1980." National Science Foundation Grant no. 85-NSF-SES-85-12452. Manuscript.
- _____. 1990. *Job Queues, Gender Queues: Explaining Women's Inroads into Male Occupations*. Philadelphia: Temple University Press.
- Roos, Patricia A. 1990. "Hot Metal to Electronic Composition: Gender, Technology, and Social Change." Pp. 275–98 in B. F. Reskin and P. A. Roos, *Job Queues, Gender Queues: Explaining Women's Inroads into Male Occupations*. Philadelphia: Temple University Press.
- Rotella, Elyce. 1977. *From Home to Office: U.S. Women at Work, 1870–1930*. Ann Arbor: UMI Research Press.
- Steiger, Thomas, and Barbara F. Reskin. 1990. "Baking and Baking off: Deskilling and the Changing Sex Make Up of Bakers." Pp. 257–74 in B. F. Reskin and P. A. Roos, *Job Queues, Gender Queues: Explaining Women's Inroads into Male Occupations*. Philadelphia: Temple University Press.
- Stolzenberg, Ross. 1982. "Industrial Profits and the Propensity to Employ Women Workers." Paper presented at the Workshop on Job Segregation by Sex, Committee on Women's Employment and Related Social Issues, National Research Council, Washington, DC.
- Strober, Myra. 1984. "Toward a General Theory of Occupational Sex Segregation." Pp. 144–56 in *Sex Segregation in the Workplace: Trends, Explanations, Remedies*, edited by B. Reskin. Washington, DC: National Academy Press.
- _____. 1988. "The Processes of Occupational Segregation: Relative Attractiveness and Patriarchy."

- Paper presented at the meetings of the American Educational Research Association, New Orleans.
- Strober, Myra, and Carolyn Arnold. 1987. "The Dynamics of Occupational Segregation Among Bank Tellers." Pp. 107–48 in *Gender in the Workplace*, edited by C. Brown and J. Pechman. Washington, DC: Brookings Institution.
- Strom, Sharon Hartman. 1987. "Machines Instead of Clerks': Technology and the Feminization of Book-keeping, 1910–1950." Pp. 63–97 in *Computer Chips and Paper Clips*. Vol. 2, edited by H. Hartmann. Washington, DC: National Academy Press.
- Thomas, Barbara J. 1990. "Women's Gains in Insurance Sales: Increased Supply, Uncertain Demand." Pp. 183–204 in B. F. Reskin and P. A. Roos, *Job Queues, Gender Queues: Explaining Women's Inroads into Male Occupations*. Philadelphia: Temple University Press.
- Thomas, Barbara J., and Barbara F. Reskin. 1990. "A Woman's Place Is Selling Homes: Women's Movement into Real Estate Sales." Pp. 205–24 in B. F. Reskin and P. A. Roos, *Job Queues, Gender Queues: Explaining Women's Inroads into Male Occupations*. Philadelphia: Temple University Press.
- Thurow, Lester. 1969. *Poverty and Discrimination*. Washington, DC: Brookings Institution.
- _____. 1972. "Education and Economic Equality." *The Public Interest* 28:66–81.

87 ■ David B. Grusky and Asaf Levanon

Four Gloomy Futures for Sex Segregation

For those of us trained in the structural-functionalism of the 1960s and 1970s (e.g., Parsons 1970), it is perhaps especially difficult to shake the simple view that ascription is destined to wither away, if only gradually and fitfully. The well-known presumption of the structural-functional tradition is that racial, gender, and class discrimination will be gradually undermined by the inefficiency of discrimination (e.g., Becker [1971] 1957), the diffusion of bureaucratic personnel policies (e.g., Dobbin et al. 1993), or the spread of egalitarian sensibilities and the consequent delegitimation of discriminatory tastes (e.g., Jackson 1998). The purpose of this chapter is

to suggest that such arguments, however normatively attractive they may be, overlook many complexities in the mechanisms by which gender inequality is generated and maintained. Although obviously one cannot rule out the possibility that the future of gender inequality will unfold over the very long run in accord with a structural-functional formulation, there are several other futures that also seem plausible, at least in the more meaningful short run. The future of gender-based ascription is in this regard a very open one, perhaps more so than is the case with other types of ascription, such as those based on race or social origins.

This is a commissioned chapter that draws heavily on material in previous publications (Maria Charles and David B. Grusky, 2004, *Occupational Ghettos: The Worldwide Segregation of Women and Men*, Stanford, CA: Stanford University Press; Maria Charles and David B. Grusky, 2007, "Egalitarianism and Gender Inequality," pp. 327–342 in *The Inequality Reader: Contemporary and Foundational Readings in Race, Class, and Gender*, edited by David B. Grusky and Szonja Szelenyi, Boulder, CO: Westview Press; David B. Grusky, Forthcoming, "A Benign Future for Occupational Sex Segregation?" in *Inequality, Religion, and Differentiation*, edited by Bernice McNair Barnett, Urbana: University of Illinois Press; Francine D. Blau, Mary C. Brinton, and David B. Grusky, 2006, "The Declining Significance of Gender?," pp. 3–34 in *The Declining Significance of Gender?*, edited by Francine D. Blau, Mary C. Brinton, and David B. Grusky, New York: Russell Sage; Asaf Levanon and David B. Grusky, 2007, "The Making of an Occupational Ghetto," Working paper, Stanford Center for the Study of Poverty and Inequality, Stanford University).

Why even attempt such a speculative discussion? After a half-century of dramatic reductions in the gender pay gap and other forms of gender inequality, recent evidence reveals that change is stalling on some fronts and slowing on others (see Blau, Brinton, and Grusky 2006), and important decisions must therefore be made as to whether new and more aggressive policy initiatives should be undertaken. It may be useful in making those decisions to chart the various pathways and endpoints that appear to lie before us and to ask whether certain initiatives might make some endpoints more likely than others. We will proceed, then, by attempting to identify the most plausible endpoints, obviously leaving unaddressed the question of how one might come to prefer some endpoints over others.

In taking on this task, special attention is properly paid to the future of occupational sex segregation, given that the gender pay gap and other gender differentials in life chances are largely a function of such segregation (e.g., Petersen and Morgan, Ch. 88). The occupational structure is in this sense the backbone of the gender stratification system and hence a natural starting point for any discussion of its future.

The Two Dimensions of Segregation

The task of making sense of the future of sex segregation is best addressed by laying out the competing logics that underlie late-industrial segregation regimes. As Charles and Grusky (2004) have argued, two distinct cultural principles are interwoven to generate the contemporary pattern of sex segregation, the first being the essentialist presumption that women and men have fundamentally different tastes and proclivities and are accordingly best suited for different types of occupations, and the second being the vertical presumption that men are more status-worthy than women and accordingly well suited for prestigious positions that require the most substantial human capital investments. These two cul-

tural tenets are undergirded by a wide array of proximate processes on both the supply and demand sides. For example, the essentialist presumption that men are especially well suited for jobs entailing strength is not just internalized by employees and revealed in their preferences for particular jobs, but is also internalized by employers and revealed in their tastes for discrimination at the point of hiring, firing, and promotion. Similarly, the vertical presumption that men are more status-worthy is not just expressed in their supply-side decisions to invest heavily in human capital, but is also expressed in the demand-side decisions of employers to privilege men in the competition for high-status occupations (see Charles and Grusky 2004, pp. 15–23, for further discussion of the mechanisms underlying these cultural principles).

In contemporary segregation regimes, a core essentialist claim is that women excel in service, nurturance, and interpersonal interaction while men excel in interaction with things (rather than people) and in strenuous, physical, or demanding labor. This cultural tenet is expressed in the worldwide tendency for women to disproportionately fill nonmanual occupations and men to disproportionately fill manual ones. That is, because nonmanual occupations are associated with skills in which women are presumed to excel (e.g., interaction with people), and because manual occupations are associated with skills in which men are presumed to excel (e.g., strength, interaction with things), the forces of demand (e.g., discrimination) and supply (e.g., aspirations) will generate much segregation across the manual-nonmanual divide.

The distinctive feature of this particular form of essentialism is that it allocates women into occupations that are relatively high in pay and status (i.e., nonmanual occupations) and are therefore preferred relative to the manual alternative. We may accordingly refer to this form of essentialism as female-advantaging. Although Charles and Grusky (2004) focused almost exclusively on the female-advantaging

variant of essentialism, it is well to bear in mind that many forms of essentialism are in fact male-advantaging. Most obviously, the presumption that men are more suited to analytical or mathematical tasks has precisely this male-advantaging effect, as the occupations that rely on analytical or mathematical skills (e.g., technical and professional occupations) are typically higher in status and pay than those that do not rely on such skills.

The contemporary segregation regime is therefore generated by three processes: (1) the “vertical” cultural presumption that men are well suited for the substantial human capital investments that underwrite entry into high-wage, high-prestige, and otherwise desirable occupations; (2) the “female-advantaging” essentialist presumption that men are well suited for (relatively less desirable) manual occupations requiring strength, robustness, and interaction with things; and (3) the “male-advantaging” essentialist presumption that men are well suited for (relatively desirable) occupations requiring analytical and mathematical skills. If one then analyzes a $J \times 2$ cross-classification of occupation by gender, this interpretation may be formalized as follows:

$$m_j = \alpha + \beta X_j + \gamma Y_j + \delta Z_j \quad (1)$$

where m_j refers to the logged ratio of men to women in the j^{th} occupation, X_j refers to the vertical location of the j^{th} occupation (as indexed by wages, prestige, or some other measure of desirability), Y_j refers to the quantitative score of the j^{th} occupation on a female-advantaging trait (e.g., requirements for strength), Z_j refers to the quantitative score of the j^{th} occupation on a male-advantaging trait (e.g., requirements for analytical or mathematical skills), α pertains to the baseline level of female labor force participation (when X_j , Y_j , and Z_j all equal zero), and β , γ , and δ pertain to the effects of the three types of occupational traits on levels of segregation.

In any credible model, the full complement of segregative forces could not likely be repre-

sented without several vertical dimensions (e.g., wages, prestige), several female-advantaging essentialist traits (e.g., strength, robustness), and several male-advantaging essentialist traits (e.g., mathematical skills, analytic skills). The simplified model of equation (1) is in this sense purely didactic. We resort here to a stylized model because we wish to represent late-industrial segregation in its generic form rather than that prevailing in any particular country or time period (see Levanon and Grusky 2007 for a related model applied to the United States).

The operation of this model can be illustrated with hypothetical data for each of the three occupational dimensions (i.e., X_j , Y_j , Z_j). The data graphed in Figure 1 may be understood as stylized representations of occupational variability in (a) socioeconomic status, (b) requirements for mathematical or analytical skills (i.e., a male-advantaging essentialist trait), and (c) requirements for strength or robustness (i.e., a female-advantaging essentialist trait).¹ The stacked bars of Figure 1 thus represent how the socioeconomic and essentialist traits characterizing each occupation combine to produce an observed level of segregation. Because prior research reveals that female-advantaging essentialism is an especially important determinant of segregation (e.g., Levanon and Grusky 2007), Figure 1 rests on the assumption that γ , the coefficient for female-advantaging essentialism, is twice as large as β and δ (and hence the stacked bars represent the quantities X_j , $2Y_j$, and Z_j). This assumption expresses the well-known importance of the manual-nonmanual divide in contemporary segregation regimes (e.g., Charles and Grusky 2004).

The most striking feature of Figure 1 is that many occupations contain within them starkly opposing forces for male and female representation. The laboring occupations, for example, are “feminized” by virtue of their low socioeconomic status (i.e., the vertical dynamic) and analytic requirements (i.e., the male-advantaging essentialist dynamic), but are simultaneously “masculinized” by virtue of their substantial strength requirements (i.e.,

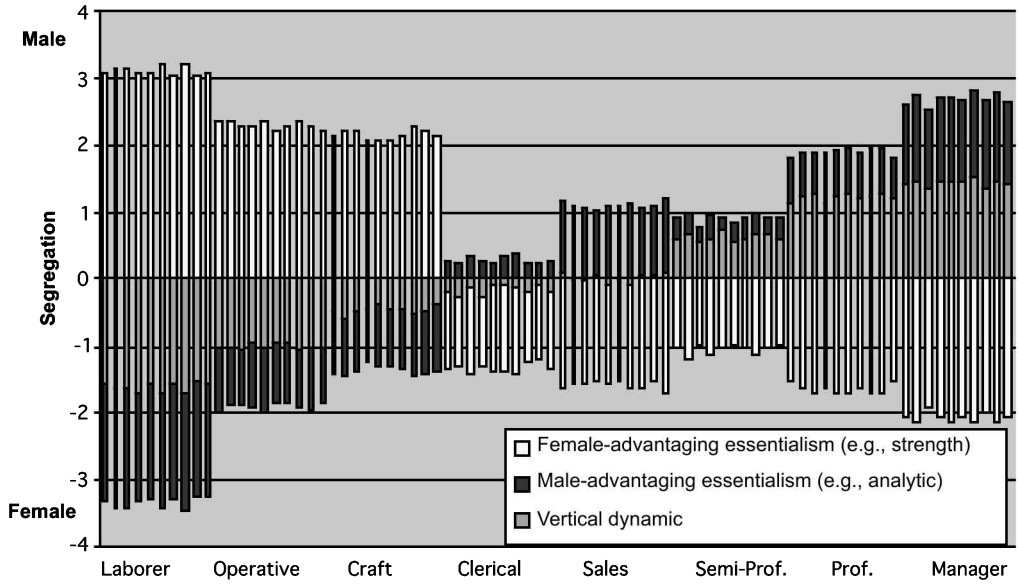


Figure 1. Components of Segregation

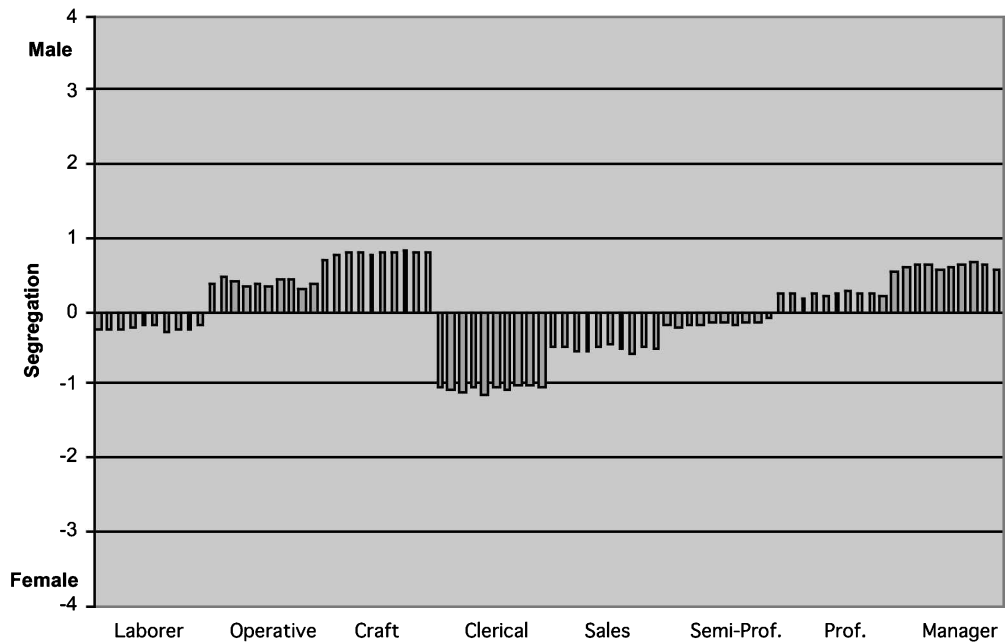


Figure 2. Overall Segregation Pattern

the female-advantaging essentialist dynamic). The combined effect of these opposing forces, as represented by adding X_j , $2Y_j$, and Z_j , together, is a slight net female overrepresentation (see Figure 2). At the other end of the

class distribution, we find that semiprofessional occupations are feminized by virtue of low strength requirements but masculinized by virtue of relatively high status and analytic requirements. Again, these countervailing

Table 1. A Typology of Segregation Regimes

| Vertical Principle | Essentialist Principle | |
|--------------------|-------------------------|------------------------|
| | <i>High</i> | <i>Low</i> |
| <i>High</i> | Type I: Dual inequality | Type II: Queuing |
| <i>Low</i> | Type III: Essentialist | Type IV: Dual equality |

forces combine to generate a net female overrepresentation (see Figure 2), although such an outcome is in this case achieved with a very different mixture of occupational traits. The overall pattern yielded in Figure 2, although entirely stylized, is nonetheless consistent with the representation of late-industrial segregation that Charles and Grusky (2004) have elsewhere described.

This simple analytic setup, which is hopefully an uncontroversial account of the hybrid structure of late-industrial segregation, can serve as a tool for uncovering the implications of possible changes in the mechanisms underlying segregation. It is useful in particular to consider how changes in the relative prominence of essentialist and vertical principles may play out (see Table 1). Although the contemporary “dual inequality” regime rests on both principles, the question at hand is whether there are forces at work that will change the prevailing mixture of essentialist and vertical principles. The following sections take on this question by considering whether any of the regimes distinguished in Table 1 might emerge over the next half-century or so.

Essentialism and the Rise of a Type III Regime

In predicting the future, a conventional approach is simply to extrapolate recent developments, an approach that in this case suggests a continued weakening of the vertical principle. The main cultural force behind such weakening, the diffusion of egalitarian commitments, can of course be dated to the Enlightenment but has arguably intensified over the last half-century. As egalitarianism spreads, women

and men are increasingly viewed as having the same rights and responsibilities, and rules, laws, and regulations aimed at reducing gender inequality in families and the workplace are increasingly adopted. This diffusion of liberal egalitarianism acts to erode vertical segregation through changes in the supply and demand for labor. On the supply side, the rise of liberal egalitarianism means that women will come to be more fully committed to the formal labor force and, by virtue of this commitment, decide to make human capital investments that qualify them for occupations that are similar in pay and prestige to “men’s occupations.” At the same time, demand-side discrimination against women should also weaken, not merely because employers will gradually shed their tastes for vertical discrimination against women but also because, in a society where egalitarianism has spread and become institutionalized, there are substantial legal, financial, and public-relations penalties imposed on discriminatory firms. These developments should in turn precipitate “feedback effects” (Blau, Ferber, and Winkler 2002) whereby new cohorts of women come to appreciate that vertical discrimination has receded and that substantial human capital investments are now more likely to yield a payoff. It is precisely such multiplier effects that can convert small shocks to the system into large changes in gender inequality.

Will this ongoing diffusion of liberal egalitarianism also precipitate a decline in essentialism? There is good reason to doubt that it will. As Charles and Grusky (2004) argue, essentialism can readily adapt to liberal egalitarianism, transmuting itself into a conservative “different but equal” understanding of the ap-

titudes and abilities of women and men (also, Charles and Grusky 2007). Within the essentialist worldview, women and men are understood as having very different tastes, aptitudes, and aspirations, an understanding that can be reconciled with liberal egalitarianism insofar as the (relatively few) men and women with gender-atypical aspirations can compete fairly in a gender-neutral contest. For a liberal egalitarian, one must defend the right of women to freely and fairly compete for any occupation to which they aspire, but there is no corollary obligation to examine how those aspirations were formed or why they may putatively differ from those of men. This mix of liberal egalitarianism and essentialism therefore leads to a Type III essentialist regime rather than any simple withering away of “ascription” (Parsons 1970).

The persistence of essentialism can be understood as playing out on both the supply and demand sides. On the supply side, liberal egalitarians will view aspirations as well beyond their purview, and the persistence of gender differences in such aspirations won't, as a result, be scrutinized or challenged to the extent that they would under more radical egalitarian commitments. On the demand side, the liberal egalitarian commitment delegitimizes all forms of pure discrimination, but it does not as directly challenge statistical discrimination that rests on essentialist presumptions about gender differences in aptitudes. In a world in which women have disproportionately invested in nurturance and service, essentialist stories about intrinsic gender differences in aptitudes have ample room to flourish, and employers may therefore reason that gender provides a good signal of capabilities in nurturing and service. It follows that liberal variants of egalitarianism serve principally to undermine the assumption of male primacy rather than gender essentialism; and, consequently, essentialist forms of segregation may prove to be quite resistant to the spread of liberal egalitarianism. With the rise of liberal egalitarianism,

women will develop an ever-strengthening commitment to the paid labor market, but they will do so in ways that are consistent with their essentialist preferences, with the essentialist sanctions imposed by others, and with the essentialist prejudices of employers.

If liberal egalitarianism leaves essentialism largely untouched, are there any other forces that might call essentialism into question? Although one obviously cannot rule out the possibility of an antiessentialist revolution, there is good reason to believe that, at least in the short term, essentialism will live on and remain a main interpretive lens through which gender inequality is understood. The essentialist worldview remains, after all, deeply embedded in families (with their gender-specific socialization practices), work organizations (with their discriminatory hiring practices), and all manner of other institutional contexts. By way of (trivial) example, consider the continuing practice among American fast-food restaurants of providing gender-specific toys to children, a practice of interest only because it is widely diffused and evidently unobjectionable to all but a small minority of gender progressives. If these same restaurants distributed toys on the basis of racial or class standing, the practice would be deemed absurd at best and racist or classist at worst. This example suggests that, at least in the United States, it is less legitimate to interpret racial or class-based inequalities in essentialist terms than to interpret gender segregation and inequality in these terms. In this sense, there is a deep structure to essentialist segregation that seemingly makes it an organic feature of modern economies, even as pressures for equalization may mount in other domains of the stratification system (see Charles and Grusky 2004 for details).

If indeed a Type III regime is in ascendancy, one might reasonably ask how sex segregation will be patterned under such a regime. This question can be explored by recalculating the predicted segregation values after reducing the parameter for vertical

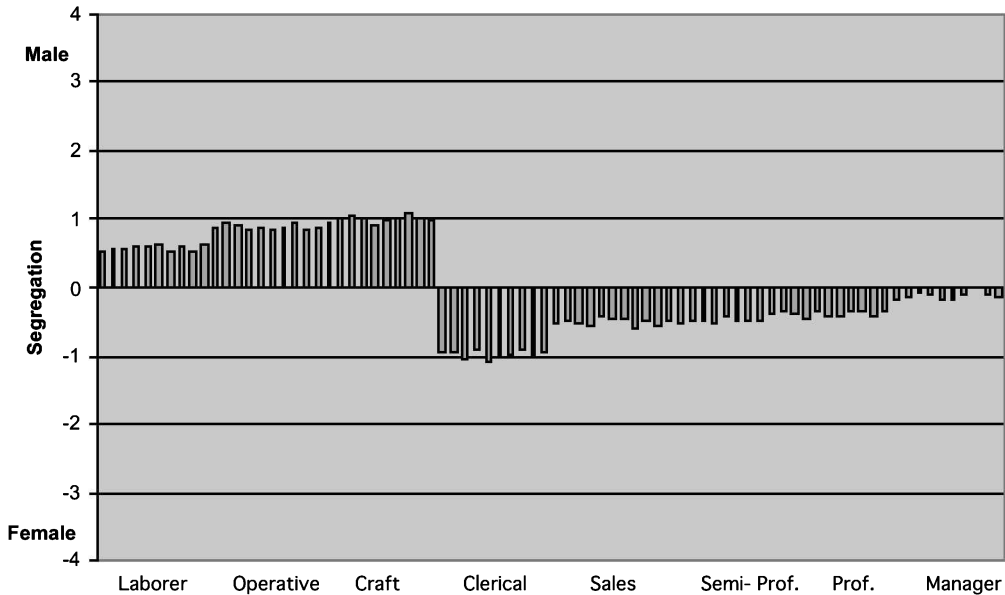


Figure 3. Type III Regime (weakened vertical segregation)

segregation (i.e., β). In Figure 3, this coefficient has been (arbitrarily) halved, while the two essentialist parameters remain unchanged. Although the vertical principle is clearly visible in both the manual and non-manual sectors of Figure 3, it is so weakened that the most desirable nonmanual occupations have become disproportionately female and the least desirable manual occupations have become disproportionately male. Under this formulation, the parameter for female-advantaging essentialism remains unchanged, but even so the manual-nonmanual divide is cast into sharper relief because it is no longer camouflaged by other, competing principles of segregation. The segregation under the resulting Type III regime therefore takes on an increasingly sectoral character.

The Opt-Out Revolution and the Persistence of Type I Segregation

The foregoing story stresses that liberal egalitarianism, for all its success in weakening vertical segregation, is not exactly tailor-made for the task of taking on essentialist

segregation. Although the persistence of such segregation is inconsistent with structural-functional stories that have all forms of gender inequality declining of a piece, the story is nonetheless one of gradual desegregation that continues to play out until vertical segregation is completely eliminated. Is there an alternative story on offer that questions whether even vertical segregation will wither away? Is it plausible that our current Type I regime will persist over the long run rather than gradually transform into a Type III regime?

The main foundation for such a prediction is to be found in currently fashionable stories about an ongoing “opt-out revolution” (e.g., Belkin, Ch. 78; Story 2005). Under these formulations, women with high investments in human capital and seemingly attractive career options are presumed to be increasingly opting out of high-stress careers in favor of recommitting to their children, spouses, and domestic responsibilities more generally. The presumed precipitants of such opting-out are either (1) a disenchantment with the shallow and superficial rewards that careers can offer

and a rediscovery of the more fundamental rewards of childrearing, or (2) a recognition that, because gender discrimination and institutionalized barriers to career advancement persist in the workplace (e.g., a glass ceiling), career investments won't necessarily have the payoff that had been initially anticipated. The result in either case is that women will come to make less substantial investments in human capital and that the historic erosion in vertical segregation will slow down or even reverse itself.

In judging the plausibility of the opt-out story, one of course has to first appreciate that the data on which it is based are exceedingly thin, surely too thin to provide any foundation for convincing extrapolation. The original opt-out essays were based on convenience samples of women in elite universities (Belkin, Ch. 78; Story 2005), and subsequent and more rigorous research reveals only a slight decline in female labor force participation in the post-2000 period, a decline that pertains equally to mothers and nonmothers alike (e.g., Boushey, Ch. 79; also see Weeden 2004). Furthermore, even if the dynamics described by opt-out scholars are indeed at work, they may not imply anything more than a minor and temporary slowdown in the rate of decline in vertical segregation. It is relevant in this regard that the opt-out narrative doesn't typically rest on any claim to the effect that demand-side discrimination is in fact intensifying. Rather, the narrative simply presumes that women overestimated the likely payoff to their human capital investments, an overestimate predicated on the assumption that discrimination and other barriers to workplace advancement would erode more quickly than in fact they did. Ironically, the forces for retrenchment under this story are set in motion not because of any real deterioration in the opportunities for women, but only because opportunities prove not to be as substantial as had been anticipated. The more important point, however, is that these forces

for retrenchment merely imply a temporary readjustment in human capital investments, a readjustment that brings them into alignment with real payoffs rather than the earlier highly optimistic estimates of the rate of decline in discrimination. This one-time readjustment implies a corresponding one-time deterioration in human capital investments and occupational outcomes but not a more fundamental change in the trend line thereafter.

The proponents of the opt-out story seem not to have appreciated that their story, even if true, thus implies a seemingly minor one-time readjustment rather than a long-term threat to the gender revolution. Is there any way to embellish the story in ways that render it more threatening? Indeed there is: It need merely be argued that the opt-out phenomenon, no matter how limited it may be, becomes prominent and celebrated enough to precipitate an ongoing and more fundamental change in how the aptitudes of women are understood. That is, insofar as employers and others come to believe that women are increasingly bowing out of high-status occupations, such observations could conceivably reinforce the older cultural principle that men are especially well suited to high-status pursuits. This recommitment to the old principle of male primacy will in turn fuel (1) an increase in statistical discrimination by employers, and (2) a consequent reluctance on the part of women to make human capital investments that don't in the end pay off. It follows that feedback effects are the consummate double-edged sword. Indeed, just as they can magnify and reinforce initial declines in segregation and convert small developments into a revolution, so too they can magnify and reinforce initially small counter-developments and transform them into a full-blown counterrevolution. The most perverse version of this embellished opt-out story has employers reassessing the aptitudes of women not because of any objective increase in opting-out behavior but

because of celebrated (yet incorrect) claims of an objective increase. However unfounded those claims may be, the consequent take-off in statistical discrimination would alter the payout to human capital, and women would be obliged to take this changing payout into account when making their investment decisions. This especially perverse version of the opt-out story falls foursquare into the category of self-fulfilling prophecy.

The Decline of Essentialism and the Move Toward Type IV Equality

To this point, our discussion has focused on whether vertical segregation will continue to wither away, and our argument that essentialist segregation is, by contrast, relatively resistant to change has not been exposed to much critical scrutiny. Is there any reason to believe that essentialist segregation might possibly wither away and move us toward Type IV segregation? Will a deeper form of egalitarianism ultimately emerge and delegitimize (1) the tendency of women and men to view their aptitudes in essentialist terms and develop aspirations in accord with these self-understandings, and (2) the tendency of employers to make hiring, promotion, and firing decisions in ways that are likewise affected by their essentialist presuppositions?

In addressing these questions, it has to be appreciated that liberal egalitarianism, for all its popularity, is hardly the only cultural commitment in play. Although liberal egalitarianism is easily reconciled with essentialism, it is of course possible that a new form of more radical antiessentialist egalitarianism will come to supplant liberal egalitarianism. This type of revolution is precisely the mission of a small vanguard population of gender liberals and feminists. Within this subpopulation, much emphasis is placed on socializing children in gender-neutral ways, with daughters in particular being encouraged to become mathematicians, engineers, scientists, and other similarly gender-atypical occupations.

For such radical egalitarians, it is not enough to simply encourage girls to invest heavily in human capital, but additionally there is some sentiment for encouraging them to overcome essentialist presuppositions by making human capital investments that are gender-atypical in content. This sentiment may be understood, then, as delegitimizing not just vertical segregation but essentialist segregation as well.

The key question at hand is whether the hopes and aspirations of these vanguard parents can be realized in the near term. It is no easy task to convert such abstract political commitments into truly gender-free socialization given that children are exposed daily to more conventional gender messages from peers, teachers, films, video games, popular music, and television. Moreover, even though such radical egalitarians are no doubt intellectually committed to antiessentialism, most of them were themselves raised under the influence of deeply essentialist precepts; and their commitment to antiessentialism is therefore best viewed as an ideological veneer that is layered on top of subtle, subconscious, and perhaps more profound essentialist presuppositions (e.g., Reskin, Ch. 82). This essentialist foundation will surely color their interactions with everyone, including their own children, and make it difficult to succeed in delivering truly gender-free socialization. The more fundamental point is of course that relatively few parents even have an antiessentialist veneer of this sort. For children of these less “enlightened” parents, essentialist socialization within the home will simply mirror essentialist socialization outside it, and little in the way of meaningful change can be anticipated.

This is not to deny that incipient antiessentialist forces are clearly at work and that gradual and possibly fitful movement toward a more radical form of egalitarianism can be anticipated. However, because our commitment to antiessentialism is far less profound at this point than our commitment to liberal egalitarianism, it seems likely that more headway will be made in reducing segregation of the

vertical variety than segregation of the essentialist variety. The resulting “different but equal” solution bears, perhaps ironically, some resemblance to the vision of gender equality that comparable worth activists have long pursued. The Type III regime is nonetheless achieved through different means than “comparable worth” adjustments; that is, the movement to Type III segregation occurs as women enter high-prestige occupations that have an essentialist “female profile” (e.g., pediatrician, sociologist, personnel manager), whereas the comparable worth vision was to be secured without any occupational mobility at all, just an upward adjustment of rewards to characteristically female occupations. The end result in either case is a new form of “different but equal” segregation.

The Decline of Female-Advantaging Essentialism

If a wholesale decline in essentialism is not in our immediate future, is there any reason to anticipate a decline in at least some forms of essentialism? There have clearly been non-trivial shifts over the last half-century in the types of traits that have been regarded as essentially male or essentially female. The question that then emerges is whether we should anticipate further changes over the next half-century in the content of our essentialist beliefs and in the structure of essentialist segregation.

In posing this question, it bears recalling that men were linked to manual labor in an especially prominent way during a period of history, that of early industrialism, in which manual labor was a central factor of production (see Kessler-Harris 1982; Milkman 1987). By contrast, manual labor is now becoming a less fundamental part of production, particularly in late-industrial economies that are experiencing deindustrialization through labor-reducing technological change and the globalization of production. It is not necessarily plausible that men will stand idly

by as “men’s work” of the conventional manual sort becomes peripheralized. Over the next half-century, the transition to a post-manual system may begin to reduce male commitment to those pieces of the essentialist package, especially an emphasis on male physicality, that assign men to increasingly devalued pursuits. It must in this context be appreciated that essentialism rests in part on self-assessments of competence and accordingly requires some complicity among those making those judgments. Will men remain complicit even when doing so profoundly disadvantages them?

If men are able to gradually disown such female-advantaging essentialism, they would succeed in harnessing essentialism to work almost entirely on their behalf. Although progressives and other “gender egalitarians” might assume that *any* breakdown in essentialist beliefs should be applauded, a decreasing emphasis on male physicality has ambiguous implications for women’s status because it involves undermining one of those relatively rare manifestations of essentialism that benefits women rather than men. If the male physicality assumption breaks down, this female-advantaging effect is lost; and the remaining forms of essentialist segregation disadvantage women (in terms of pay and prestige) quite systematically. For example, the essentialist presumption that men have superior mathematical or analytical skills generates a disproportionately large number of male accountants, scientists, programmers, and engineers, all occupations that are relatively well paid and high in prestige.

The implications of such a breakdown in female-advantaging essentialism are depicted in Figure 4. When the parameter for female-advantaging segregation is trimmed in half (i.e., reduced to .5), the resulting segregation profile (see Figure 4) takes on a quite dramatic vertical form in which men dominate the good jobs and women are relegated to the bad ones. Although the profile appears to take on a purely vertical form, in fact both

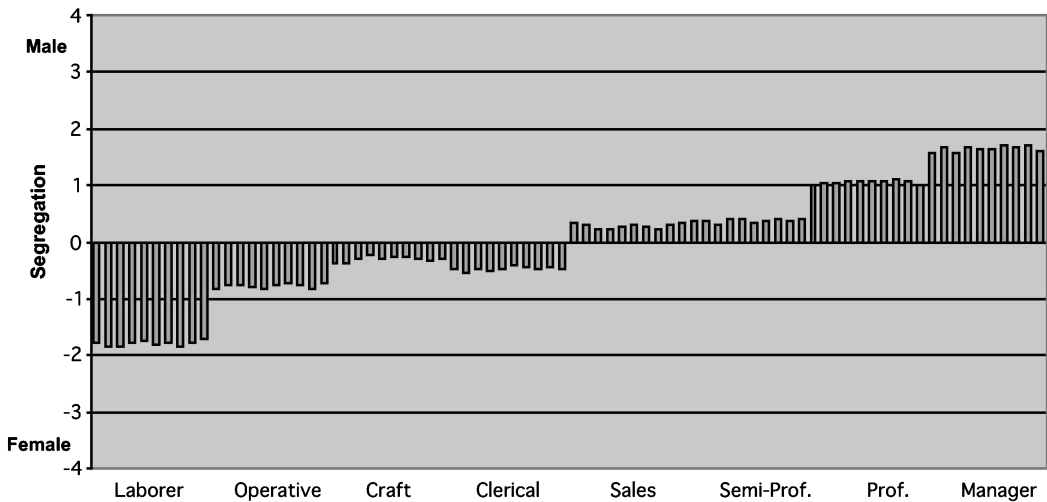


Figure 4. Weakened Female-Advantaging Segregation

essentialist and vertical processes underlie it, a complication that is largely concealed because the female-advantaging form of essentialism has been suppressed. The remaining male-advantaging variant of essentialism combines with the vertical dynamic to produce a segregation regime that would have to be understood as a real setback for women. The simple moral here is that, much as we think of essentialism as regressive, in fact its female-advantaging variant plays a role in making segregation as palatable as it now is. If the female-advantaging variant begins to wither away, and there is good reason to believe that it might, the segregation regime may begin to assume a more hostile form for women. This threat to the gender revolution may well be more serious than the celebrated opt-out revolution: It is hardly implausible, after all, to suppose that men will successfully shake off old disadvantaging definitions of masculinity and adopt new definitions that bring more power, prestige, and rewards.

The Rise of Male-Advantaging Micro-Essentialism

The final possibility worth considering is that male-advantaging essentialism will not

just persist over the next half-century but in fact strengthen. This forecast is most plausibly developed in the context of the argument, already outlined above, that vertical segregation will gradually weaken as liberal egalitarianism spreads. As may be recalled, our core claim is that vertical segregation is partly undermined by liberal egalitarianism, whereas essentialist segregation can coexist quite organically with liberal egalitarianism. Although this line of argument implies that essentialism will persist, it does not provide any reason to believe that it will increase.

Is there room for a more aggressive argument about how essentialist segregation will come to increase? The main rationale for such an argument is that the spectacular flow of women into professional, managerial, and other relatively desirable sectors of the occupational structure might itself precipitate a reactive resegregation within those sectors. As the two genders come together in the workplace, deeply entrenched essentialist precepts may inform decisions about how tasks should then be divided among them, with the result being a strengthened “micro-essentialism” informed by presumed gender differences in skills, aptitudes, and tastes.

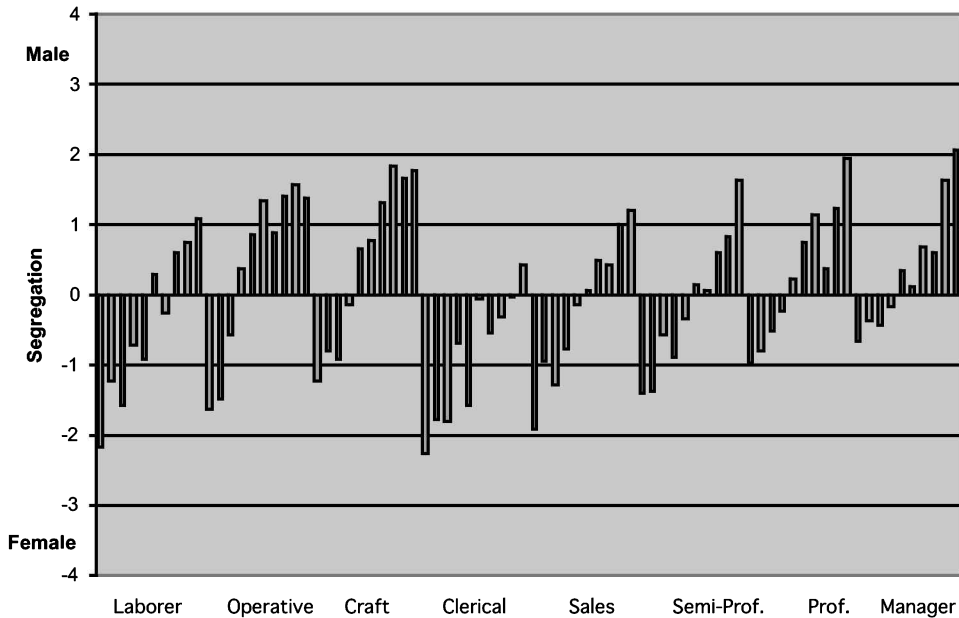


Figure 5. Strengthened Male-Advantaging Micro-Essentialism

There is much anecdotal and qualitative evidence supporting just such an account: (1) the rising number of female physicians appears to have generated new female-dominated ghettos in nurturant work (e.g., pediatrics); (2) the recent (minor) influx of U.S. women into road-construction occupations has been coupled with personnel practices that shunt new entrants into positions that are physically undemanding or people-oriented (e.g., “flagman”); and (3) the rising number of female lawyers has been accommodated by allocating them into family practice and other specialties presumed to require female-typed skills. In each of these cases, initial headway has been made in reducing vertical segregation, only to find that the newly integrated occupations seemingly resegment along essentialist lines at a lower suboccupational level.

This resegmentative process, if allowed to play out, might lead to a regime of the sort represented in Figure 5. In generating this figure, the values in Figure 3 were simply augmented with a “noise parameter” that creates ghettos of extreme segregation within each

major class (e.g., craft, professional), a noise parameter that rests on the further assumption that male ghettos tend to be relatively high in status and female ghettos tend to be relatively low in status. This constraint represents the very plausible claim that within-class essentialism will be male-advantaging. The second main assumption underlying Figure 5 is that classes comprise detailed occupations that are very heterogeneous in their segregative practices; that is, whereas Figure 3 allowed for only trivial amounts of within-class segregation, Figure 5 ratchets up the amount of such segregation and gives it a distinctly vertical cast.²

The end result is again a highly segregated regime that is nonetheless nominally consistent with liberal egalitarianism. Although radical egalitarians may complain, for example, about the disproportionate number of female pediatricians and their relatively low pay, their concerns will never gain traction because of the compelling essentialist story that legitimates such segregation as well as the inequality in remuneration that so often goes with it. Namely, liberal egalitarians will

understand essentialist segregation as arising, in part, from the exogenous tastes of women, implying that the wage penalty they suffer is compensated by the extra utility they derive from realizing their tastes for nurturant work. These liberal egalitarian defenders will also continue to be attracted to the logic of matching women to occupations that appear to exploit their talents. By liberal egalitarian logic, the right of women to exploit their competencies and realize their aspirations should be vigorously defended, whereas the social processes that shape and form these aspirations and competencies are regarded as sacrosanct and well beyond the proper purview of policy.

Conclusions

The standard evolutionary accounts of our time treat gender inequality as a vestige of an earlier economic system (e.g., Parsons 1970), as an economic irrationality that the discipline of the market will gradually eliminate (e.g., Becker 1971 [1957]), or as an inevitable casualty of our growing commitment to liberal egalitarianism (e.g., Jackson 1998). These accounts continue to be rehearsed by some (e.g., Jackson 1998) and challenged by others (e.g., Blau, Brinton, and Grusky 2006). The main question posed here is whether a new set of narratives might be developed by treating essentialism and liberal egalitarianism as two coexisting logics in deep tension.

In building any narrative, it is conventional to privilege certain tendencies and downplay others, thus making the preferred pathway seem especially plausible. The tack instead taken here is to identify a constellation of trajectories and thereby expose straightforwardly the inevitable ambiguities and complexities in attempting to make claims about the future. Although several “long-shot trajectories” are considered (e.g., an opt-out revolution), the most plausible narratives feature the role of essentialism in

generating sex segregation and gender inequality.

The narratives featured here also take into account the powerful effects of liberal egalitarian ideology in shaping tastes and beliefs (e.g., the decline of discriminatory tastes), motivating and legitimating political reform (e.g., antidiscrimination legislation), and underwriting various types of organizational reform (e.g., the rise of bureaucracy). There is no denying the dynamic role of liberal egalitarianism in legitimating far-reaching change. It is nonetheless important to recognize that liberal egalitarianism, for all its seductive appeal, is not the only dynamic at work. Rather, because it coexists with an essentialist logic that is no less deeply institutionalized, it is unlikely that sex segregation regimes will develop as simply as standard evolutionary accounts have it.

The resulting narratives laid out here will hopefully play the role of cautionary tales. If standard evolutionary accounts lull us into a sense of security about the fundamentally benign forces at work, essentialist-laced narratives make it clear that any number of wrong-turns (by the logic of radical egalitarianism) may instead play out. These “wrong-turns” are at least plausible. It is hardly alarmist, for example, to suggest that men will gradually shed those parts of the essentialist package, like the male physicality assumption, that serve to allocate them into ever more peripheralized pursuits. If essentialism indeed evolves in such terms, women will become less protected from male competition for nonmanual occupations; and one of the few essentialist advantages that women now have will be lost. It is equally plausible that the historic feminization of the professional and managerial sectors will trigger new resegregation of the “micro-essentialist” variety. As women stream into these sectors, compensatory forms of resegregation may surface at the detailed occupational level, forms that will likely bear the imprint of male-advantaging essentialism.

The plausibility of these and other essentialist-laced scenarios make it altogether clear that a liberal egalitarian logic is not the only logic in play. When liberal egalitarianism is fused with essentialism, a “different but equal” conception of gender and social justice is generated, and even extreme forms of segregation can be taken as consistent with this conception.

NOTES

1. These three variables have each been standardized to a mean of zero and a standard deviation of one.
2. We suspect that micro-essentialism of this sort is most likely to emerge within occupations (e.g., lawyer, doctor) rather than within the big social classes of Figure 5 (e.g., laborer, operative, craft). This complication has been suppressed in Figure 5 for the purpose of simplicity.

REFERENCES

- Becker, Gary S. 1971 [1957]. *The Economics of Discrimination*, 2nd ed. Chicago: University of Chicago Press.
- Blau, Francine D., Mary C. Brinton, and David B. Grusky. 2006. “The Declining Significance of Gender?” Pp. 3–34 in *The Declining Significance of Gender?*, edited by Francine D. Blau, Mary C. Brinton, and David B. Grusky. New York: Russell Sage.
- Blau, Francine D., Marianne A. Ferber, and Anne E. Winkler. 2002. *The Economics of Women, Men, and Work*, 4th ed. Upper Saddle River, NJ: Prentice-Hall.
- Charles, Maria, and David B. Grusky. 2004. *Occupational Ghettos: The Worldwide Segregation of Women and Men*. Stanford, CA: Stanford University Press.
- _____. 2007. “Egalitarianism and Gender Inequality.” Pp. 327–42 in *The Inequality Reader: Contemporary and Foundational Readings in Race, Class, and Gender*, edited by David B. Grusky and Szonja Széleányi. Boulder: Westview Press.
- Dobbin, Frank, John R. Sutton, John W. Meyer, and W. Richard Scott. 1993. “Equal Opportunity Law and the Construction of Internal Labor Markets.” *American Journal of Sociology* 88, pp. 396–427.
- Jackson, Robert Max. 1998. *Destined for Equality: The Inevitable Rise of Women’s Status* (Cambridge, MA: Harvard University Press)
- Kessler-Harris, Alice. 1982. *Out to Work: A History of Wage-Earning Women in the United States*. New York: Oxford University Press.
- Levanon, Asaf, and David B. Grusky. 2007. “The Making of an Occupational Ghetto.” Working paper, Stanford Center for the Study of Poverty and Inequality, Stanford University.
- Milkman, Ruth. 1987. *Gender at Work: The Dynamics of Job Segregation by Sex During World War II*. Urbana: University of Illinois Press.
- Parsons, Talcott. 1970. “Equality and Inequality in Modern Society, or Social Stratification Revisited.” Pp. 14–72 in *Social Stratification: Research and Theory for the 1970s*, edited by Edward O. Laumann. Indianapolis: Bobbs-Merrill.
- Story, Louise. 2005. “Many Women at Elite Colleges Set Career Path to Motherhood.” *New York Times*, Sept. 20.
- Weeden, Kim. 2004. “Profiles of Change: Sex Segregation in the United States, 1910–2000.” Pp. 131–78 in *Occupational Ghettos: The Worldwide Segregation of Women and Men*, by Maria Charles and David B. Grusky. Stanford, CA: Stanford University Press.

Gender Gap in Wages

88 ■ Trond Petersen and Laurie A. Morgan

The Within-Job Gender Wage Gap

There are three types of discrimination that can produce wage differences between men and women. (1) Women are differentially allocated to occupations and establishments that pay lower wages. This may involve discrimination partly through differential access to occupations and establishments—that is, the matching process at the point of hire—and partly through subsequent promotions. We call this process *allocative discrimination*. (2) Occupations held primarily by women are paid lower wages than those held primarily by men, although skill requirements and other wage relevant factors are the same. This process, which we call *valuative discrimination*, is addressed by comparable worth initiatives. (3) Women receive lower wages than men within a given occupation and establishment. We call this process *within-job wage discrimination*. Allocative and valuative discrimination involve the segregation of men and women into different occupations, establishments, or both, and may occur without within-job wage discrimination. Thus, although it may be the case that where men and women share the same jobs they receive the same pay, they simply do not share the same jobs all that often. One conjecture

currently accepted by many researchers is that wage differences are less an issue of within-job wage discrimination and more a matter of allocative and valuative processes. That is, the segregation of women into lower paying occupations, establishments, or both, and lower pay in occupations held primarily by women are more important than pay differences within the same job in explaining the gender wage gap. Treiman and Hartmann (1981, 92–93) write, “Although the committee recognizes that instances of unequal pay for the same work have not been entirely eliminated, we believe that they are probably not now the major source of differences in earnings.”

This conjecture is drawn primarily from a large literature that focuses on pay differences across and within occupations. One pattern of findings is that the wage gap between men and women becomes smaller as occupational controls become finer, suggesting that a large proportion of the wage gap is explained by occupational distribution (Treiman and Hartmann 1981; Reskin and Roos 1990). For example, Treiman and Hartmann (1981, 33–39) explained 10 to 20 percent of the raw gap using 222 occupa-

Trond Petersen and Laurie A. Morgan. This chapter draws heavily on material in a previous publication (Trond Petersen and Laurie A. Morgan. “Separate and Unequal: Occupation-Establishment Sex Segregation and the Gender Wage Gap,” in *American Journal of Sociology* 101 [September 1995], pp. 329–361. Copyright © 1995 by the University of Chicago Press). Used with permission from the University of Chicago Press.

tional categories and 35 to 40 percent using 479 categories. These studies usually draw on data from the Census or national probability samples that allow no analysis of practices in specific establishments. Additional evidence suggests that, within occupations, the distribution of women across firms or establishments also accounts for some portion of the wage gap. For example, Blau (1977) found that in 11 clerical occupations, differences in men's and women's wages were larger between than within establishments.

Yet the prevailing conjecture remains a conjecture. It has not been shown that men and women receive equal pay within given occupations in given establishments. What has been shown is that sex segregation is extensive and pervasive (Bielby and Baron 1984; Petersen and Morgan 1995), but not the extent to which sex segregation accounts for the wage gap or that, when sex segregation is absent, the sexes receive more or less equal treatment. To confirm such a claim, one needs data on wages of men and women in the same detailed occupational group or position within the same establishment. Such data are not widely available except for isolated establishments.

This article reports a large-scale empirical investigation of wage differences between men and women within the same detailed occupational position within the same establishment. We use establishment-level data from a wide variety of industries. In each establishment, individual-level wage data for a large array of detailed occupational groups were collected, providing more accurate wage as well as occupational data than probably any other surveys available, except in some case studies of single establishments (e.g., Hartmann 1987). We focus first on production and clerical employees in 16 U.S. industries in the 1974 to 1983 period, primarily 1974 to 1978, analyzing data on 787,577 employees, 705 industry-specific occupations, 6,057 establishments,

and 71,214 occupation-establishment pairs. Second, we focus on seven professional and three administrative occupations across a broad range of industries in 1981, analyzing data on about 740,000 employees distributed across 2,162 establishments and 16,433 occupation-establishment pairs.

We make no attempt to settle the important conceptual issues that go along with the empirical patterns we address, namely the sources of observed patterns, neither from the demand side—that is, discriminatory behavior by employers—or the supply side—that is, behaviors by employees and prospective employees (see England 1992, chap. 2). Nevertheless, our results have implications for the kinds of theoretical issues that are most in need of being addressed and for the type of data that need to be collected and analyzed.

Data

We use two large-scale data sets. The first data set comes from 16 Industry Wage Surveys (IWS) conducted by the U.S. Bureau of Labor Statistics (BLS) in the period 1974 to 1983 (see, e.g., U.S. Department of Labor 1976), corresponding to industry codes at three and more digits as defined in the *Standard Industrial Classification Manual* (see U.S. Executive Office of the President 1987). Eleven industries were surveyed in 1974 to 1978, whereas 5 were surveyed in 1980 to 1983. The populations for the surveys and the sampling from the populations are described in several U.S. Department of Labor publications (see Petersen and Morgan 1995, table 1). Of the 16, 11 are manufacturing industries, and 5 are service industries.

In each industry, the BLS drew a sample of several hundred establishments, often covering a large proportion of the establishments in the industry. For each establishment, information was obtained from establishment records both on establishment characteristics

and on a large number of the production and/or clerical workers in the establishment. Within each industry, only a selection of occupations were surveyed—on average 42 occupations per industry. The occupations were selected by the BLS in order to provide a wide representation of production or clerical occupations in an industry. The individual-level data provide information on each individual in the relevant occupation and establishment. Excluded from the data collection were professional and managerial employees. Because these occupations may have exhibited wider variations in wages, even at the occupation-establishment level, there may have been less occupation-establishment level variation in wages here than in samples including professionals and managers.

For each employee surveyed, information was obtained on sex, occupation (an industry-specific code), method of wage payment (incentive- or time-rated), and hourly earnings. No information was collected on race, age, experience, or education. The occupational classification is unusually detailed, corresponding in many cases to nine digits in the *Dictionary of Occupational Titles* (see U.S. Department of Labor 1977). In other cases, the titles are specific to the BLS data, based on industry-specific codes, but are usually as detailed as the nine-digit titles in the *Dictionary of Occupational Titles*.

Wage data are straight-time hourly wages in 13 industries and full-time weekly earnings in the other three, excluding premium pay for overtime and work on weekends, holidays, and late shifts. Thus, we do not conflate pay earned on regular hours with pay earned on overtime and irregular hours, making the wage data less prone to bias than virtually any other study used for assessing wage discrimination. Men work more overtime hours than women, due either to preference for or better access to overtime hours, and overtime hours are usually paid at a higher rate. Non-production bonuses, such

as year-end bonuses, are also excluded, whereas incentive pay is included.

The second data set we use is the National Survey of Professional, Administrative, Technical, and Clerical (PATC) employees in 1981, also conducted by the BLS (U.S. Department of Labor 1981). The sampling and data collection design for this survey is similar to that in the IWS data. We use the data on weekly pay for full-time employees in seven professional and three administrative occupations.

Methods

We report all statistics separately for each of the 16 industries in the IWS data.¹ The raw (average) wages, either hourly or full-time weekly earnings, for women and men in an industry are given by \bar{w}_f and \bar{w}_m , and the raw proportional wage between women and men is $w_{r,r} = \bar{w}_f / \bar{w}_m$. The average wages for women and men in occupation o are $\bar{w}_{o,f}$ and $\bar{w}_{o,m}$, and the proportional wage is $w_{o,r} = \bar{w}_{o,f} / \bar{w}_{o,m}$. The average wages for women and men in establishment e are $\bar{w}_{e,f}$ and $\bar{w}_{e,m}$, and the proportional wage is $w_{e,r} = \bar{w}_{e,f} / \bar{w}_{e,m}$. The average wages for women and men in occupation-establishment pair oe are $\bar{w}_{oe,f}$ and $\bar{w}_{oe,m}$, and the proportional wage is $w_{oe,r} = \bar{w}_{oe,f} / \bar{w}_{oe,m}$. These computations can only be done for units that are integrated by sex. We multiply each proportional wage by 100 in order to get the average female wage as percent of average male wage at the relevant level. This percentage is referred to as the relative wage in what follows.

We make three computations with relative wages. Each computation answers the following question: suppose sex segregation—by occupation, establishment, or occupation-establishment—were abolished, what then would the remaining relative wage be? These computations are defined in the appendix.

At the occupation level, we report the average of the relative wage across all sex-

integrated occupations. We do the same at the establishment level. At the occupation-establishment level, we likewise report the average of the relative wage ($w_{oe,r} \times 100$) across all sex-integrated occupation-establishment units (*oe*). The female wage gap at each of these levels is then given as 100 minus the average of the relative wage at each level. This equals the percent by which women are paid less than men.

The remaining wage gap at the occupation-establishment level can reasonably be interpreted as an estimate of the upper bound on the amount of within-job wage discrimination, the gap one would observe in the absence of occupation-establishment sex segregation. It is an upper bound because the remaining gap at that level could have been caused by other factors, such as differences in experience and human capital between men and women in a given category. The difference between the raw wage gap and the wage gap at the occupation-establishment level can thus be interpreted as being attributable to occupation-establishment sex segregation. This part of course need not be caused by discrimination alone.

Results

Table 1 gives first in column 2 the raw relative wage and then in columns 3 to 5 the relative wages controlling for occupation, for establishment, and for occupation-establishment, computed from equations (1) to (4) in the appendix. Column 1 shows the percentage female in each industry. Column 2 shows that the average of the industry-specific raw relative wages between men and women was 81 percent. Thus the average of the industry-specific female wage gaps was 19 percent, where the wage gap is the percentage by which women earned less than men, computed as 100 minus the relative wage.

Column 3 shows that the female wage gap is reduced to an average, across the 16 indus-

tries, of 8 percent when we control for occupation, using equation (2). Column 4 shows that the female wage gap is reduced, across the 16 industries, to an average of 15 percent when we control for establishment, using equation (3). Column 5 shows that the female wage gap is reduced to about 0 to 4 percent when we control for the occupation-establishment pair, using equation (4). The average of the industry-specific female wage gaps is 1.7 percent, after controlling for the occupation-establishment pair.

These results are striking. Controlling for occupation or establishment alone reduces the wage gap somewhat, but not drastically. However, controlling for the occupation-establishment pair reduces the wage gap to a point of virtually no difference between men and women. The remaining female wage gap is on average 1.7 percent across the 16 industries, even without controlling for any individual-level characteristics such as education, age, seniority, and race. Occupation-establishment segregation accounts better for wage differences between men and women than any other set of variables studied in the literature on wage differences.²

Columns 6 to 8 give a decomposition of the raw wage gap, calculated as 100 minus the relative wage from column 2, into the percentage that is due to occupation, to establishment, and to occupation-establishment segregation, when each dimension is considered alone. These results are based on equations (5) to (7) in the appendix. Column 6 shows that occupational sex segregation alone accounts for about 64 percent of the wage gap, somewhat more than other studies have shown (e.g., Treiman and Hartmann 1981; Reskin and Roos 1990). Column 7 shows that establishment sex segregation alone accounts for about 24 percent of the wage gap. This is a new result. To date, occupational segregation has been the primary focus of gender wage gap studies. Column 8 shows that occupation-establishment

Table 1. Women's Wages Relative to Men's With and Without Occupation, Establishment, and Occupation-Establishment Controls

| <i>Industry</i> | %F | <i>Women's Wages as a Percentage of Men's</i> | | | <i>Percent of Raw Wage Gap Explained by:</i> | | | |
|------------------------------------|------|---|------|-------|--|-------|------|---------|
| | | Raw | Occ | Est | Occ-Est | Occ | Est | Occ-Est |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Men's and Boys' Shirts | 92.5 | 84.0 | 92.7 | 83.1 | 98.0 | 54.5 | -5.4 | 87.5 |
| Hospitals | 84.8 | 105.3 | 95.8 | 103.7 | 101.0 | 178.7 | 30.7 | 81.3 |
| Banking* | 82.8 | 64.7 | 97.3 | 65.5 | 98.1 | 92.4 | 2.3 | 94.6 |
| Life Insurance* | 75.9 | 61.5 | 98.0 | 61.4 | 98.9 | 94.8 | -0.3 | 97.1 |
| Wool Textiles | 55.4 | 87.3 | 93.8 | 91.7 | 98.6 | 51.3 | 34.8 | 89.0 |
| Cotton Man-Made Fiber Textiles | 54.6 | 89.2 | 96.5 | 90.5 | 99.7 | 67.5 | 11.8 | 97.2 |
| Miscellaneous Plastics Products | 51.9 | 77.5 | 87.4 | 80.6 | 97.8 | 44.1 | 14.0 | 90.2 |
| Hotels and Motels | 51.5 | 87.0 | 91.9 | 89.1 | 99.7 | 37.9 | 16.4 | 97.7 |
| Computer and Data Processing* | 44.5 | 67.5 | 98.4 | 81.4 | 98.9 | 95.1 | 42.8 | 96.6 |
| Wood Household Furniture | 35.9 | 85.4 | 91.3 | 90.4 | 95.7 | 40.3 | 34.2 | 70.5 |
| Textile Dyeing and Finishing | 19.8 | 82.7 | 92.2 | 90.2 | 99.0 | 55.0 | 43.4 | 94.2 |
| Machinery | 15.5 | 74.3 | 88.8 | 89.0 | 99.6 | 56.4 | 57.2 | 98.4 |
| Nonferrous Foundries | 9.4 | 81.3 | 84.1 | 84.4 | 97.5 | 15.1 | 16.7 | 86.7 |
| Paints and Varnishes | 5.6 | 85.1 | 92.6 | 91.8 | 96.6 | 50.2 | 44.8 | 77.1 |
| Industrial Chemicals | 2.5 | 83.2 | 93.2 | 90.1 | 97.1 | 59.6 | 41.2 | 82.8 |
| Fabricated Structural Steel | 0.6 | 80.7 | 86.0 | 81.5 | 96.9 | 27.4 | 4.1 | 83.9 |
| Average Across Industries** | 42.7 | 81.0 | 92.5 | 85.3 | 98.3 | 63.8 | 24.3 | 89.1 |

Note: Column 1, %F, gives the percent female in the industry. Column 2 gives the raw relative wage, from equation (1) in the appendix. Columns 3–5 give the average of the within occupation–within establishment and within occupation-establishment relative wages, computed from equations (2)–(4). Columns 6–8 give the percent of the raw wage gap, defined as 100 minus the number in column 2, that can be attributed to occupation, establishment, and occupation-establishment segregation, computed from equations (5)–(7).

*This is based on weekly wages for full-time employees.

**Each column gives the unweighted average across industries of the percentages in the respective column.

sex segregation alone accounts for a large percentage of the wage gap, ranging from a low of 70.5 percent to a high of 98.4 percent. In 14 of 16 industries, more than 80 percent of the wage gap is explained by occupation-establishment segregation, and among seven of those more than 90 percent is explained, whereas in two industries only 70 percent and 77 percent of the gap is explained.

Again, we have the striking result that the interaction of occupation and establishment segregation not only explains dramatically more of the wage gap than either establishment or occupational segregation alone, it also explains most of the wage differences between men and women. Had men and women been equally distributed on occupation-establishment pairs, on the average 89 percent

of the wage gap would have disappeared, assuming that other forms of discrimination would not have emerged.

One point requires discussion here. The female wage gap for the workers in our sample is about 20 percent. This is substantially less than the about 40 percent gap for all full-time workers during the period of the surveys (e.g., Goldin 1990, 61). There are several reasons for this discrepancy. One is that prior studies look at median or average annual earnings for full-time employees. It is probable that men on the average work more hours per year than women and that they also work more overtime hours. Therefore, differences between men and women are likely overstated. Another reason for this difference is that the occupations covered here comprise a relatively narrow range of the overall spectrum. In particular, professional and managerial occupations are excluded from the sample. The excluded occupations are expected to be on average higher paid than those included and to be occupied primarily by men. As a result, we have, when computing the raw wage gap of 19 percent, some degree of *de facto* control for occupation.

The latter results were based on data mostly for blue-collar, clerical, and some technical employees. It might be suspected that the wage gap among professional, administrative, and managerial employees is greater. However, when we turn our attention to analysis of wage differences among these employees using PATC data, the results are similar to those in the IWS data. The wage gap in the PATC data is just 3.1 percent among men and women in the same position (occupation and rank) and the same establishment.

Conclusion and Discussion

The findings of this article are simple to summarize. First, wage differences within

given occupation-establishment pairs were relatively small: on average 1.7 percent among blue-collar and clerical as well as some technical employees, and on average 3.1 percent in seven professional and three administrative occupations, ranging from 1 percent in the lower to 5 percent in the higher ranks in an occupation. Thus, occupation-establishment segregation, not within-job wage discrimination, was the driving force for observed wage differences.³ Second, establishment segregation was important for wage differences between men and women, although not as important as occupational segregation, a result we could compute only for the blue-collar, clerical, and some technical employees in the IWS data.

The first finding is important. It shows that occupation-establishment segregation accounted for more of the gender wage gap than any other variable or set of variables currently used in studying the gender wage gap. Occupational segregation alone tends to account for about 40 percent of the wage gap, and human capital and other variables also account for about 40 percent (e.g., Treiman and Hartmann 1981, chap. 2). But no set of variables, either individual or structural, accounts for as much as 89 percent, as occupation-establishment segregation did here in the case of blue-collar, clerical, and some technical employees in the IWS data.

This first finding establishes the conjecture already made in the literature, but not yet documented: wage differences are to a larger extent generated by occupation-establishment segregation than by within-job wage discrimination. Along with the first, the second finding shows the need to study establishment as well as occupational segregation.

The implications of the findings are straightforward. In terms of policy, allocative and valiative processes should be given the most attention, and within-job wage discrimination, which is covered by the Equal

Pay Act of 1963 and which has been the implicit or explicit focus of much discussion and research, should receive less. Future research on differential wage attainment between men and women should be refocused as follows. The emphasis should be less on within-job wage discrimination and more on three prior processes: (1) the entry of employees into occupations and establishments, that is, the differential access of men and women to positions during the initial hiring or matching process, an allocative mechanism that is not easy to research; (2) career advancement within establishments, that is, the differential rates of promotion for men and women, also an allocative mechanism, but one which is more easily studied (Spilerman 1986); and (3) how jobs occupied primarily by women tend to be paid less than those occupied by men—the comparable worth issue, or what we refer to as valuative discrimination, also a line of research well under way (see England 1992).

Two issues arise in the identification of allocative processes as responsible for the gender wage gap. The first is whether segregation patterns are due primarily to discrimination or to differences in productive capacities. The meaning of our results for theory and policy depends on which mechanism operated here. This cannot be settled with our data and is obviously a task for future research. It requires information about productive capacities and would require an analysis of the matching between these and particular jobs.

The second issue concerns the role of supply-side sources of differential attainment between men and women. Of particular interest are the constraints put on women's career attainment from family obligations and household choices. Although these constraints may result most proximally from a traditional household division of labor, not directly from discriminatory behaviors by

employers, household decisions are made in light of labor market opportunities, so the two are interdependent. The role of supply-side behaviors and their interrelationship with employer behaviors in generating the observed occupation-establishment segregation is in need of research.

Appendix

The *raw* relative wage between women and men is given as

$$w_{(r,r)} = \frac{\bar{w}_f}{\bar{w}_m} \times 100. \quad (1)$$

The relative wage controlling for occupation obtains as

$$w_{(o,r)} = \frac{1}{N_{o(I)}} [\sum_{o=1}^{N_{o(I)}} (w_{o,r})] \times 100. \quad (2)$$

Here, $N_{o(I)}$ is the number of occupations in which both men and women are present.

The establishment-level relative wage obtains as

$$w_{(e,r)} = \frac{1}{N_{e(I)}} [\sum_{e=1}^{N_{e(I)}} (w_{e,r})] \times 100. \quad (3)$$

Here, $N_{e(I)}$ is the number of establishments where both men and women are present.

The occupation-establishment-level relative wage obtains as

$$w_{(oe,r)} = \frac{1}{N_{oe(I)}} [\sum_{oe=1}^{N_{oe(I)}} (w_{oe,r})] \times 100. \quad (4)$$

Here, $N_{oe(I)}$ is the number of occupation-establishment pairs in which both men and women are present.

The raw wage gap obtains as 100 minus the number in (1). The percentage of the raw wage gap due to occupational segregation alone is given by

$$\%w_{(o,r)} = \frac{w_{(o,r)} - w_{(r,r)}}{100 - w_{(r,r)}} \times 100. \quad (5)$$

The percentage of the raw wage gap due to establishment segregation alone is given by

$$\%w_{(e,r)} = \frac{w_{(e,r)} - w_{(r,r)}}{100 - w_{(r,r)}} \times 100. \quad (6)$$

The percentage of the raw wage gap due to occupation-establishment segregation alone is given by

$$\%w_{(oe,r)} = \frac{w_{(oe,r)} - w_{(r,r)}}{100 - w_{(r,r)}} \times 100. \quad (7)$$

NOTES

The chapter is a version of Petersen and Morgan (1995) that was excerpted and partially rewritten for this volume by Petersen and Morgan with the help of David Grusky.

1. Unfortunately, we cannot compare wages across industries, because the industry data come from different years and hence reflect inflation as well as general wage increases.

2. Although Groshen (1991) and Tomaskovic-Devey (1993) report similar results on wage differences, neither author reported the wage gap at the occupation-establishment level.

3. Similar results are shown in Petersen et al. (1997) for Norway in 1984 and 1990.

REFERENCES

- Bielby, William T., and James N. Baron. 1984. "A Woman's Place Is with Other Women: Sex Segregation Within Organizations." Pages 27-55 in *Sex Segregation in the Workplace: Trends, Explanations, Remedies*, edited by Barbara F. Reskin. Washington, D.C.: National Academy Press.
- Blau, Francine D. 1977. *Equal Pay in the Office*. Lexington, MA: Lexington Books.
- England, Paula. 1992. *Comparable Worth: Theories and Evidence*. Hawthorne, NY: Aldine de Gruyter.
- Goldin, Claudia. 1990. *Understanding the Gender Gap: An Economic History of American Women*. New York: Oxford University Press.
- Groshen, Erica L. 1991. "The Structure of the Female/Male Wage Differential: Is It Who You Are, What You Do, or Where You Work?" *Journal of Human Resources* 26(3):457-72.
- Hartmann, Heidi I. 1987. "Internal Labor Markets and Gender: A Case Study of Promotion." Pages 59-92 in *Gender in the Workplace*, edited by Clair Brown and Joseph Pechman. Washington, D.C.: Brookings Institution.
- Petersen, Trond, and Laurie A. Morgan. 1995. "Separate and Unequal: Occupation-Establishment Sex Segregation and the Gender Wage Gap." *American Journal of Sociology* 101(2): 329-61.
- Petersen, Trond, Vemund Snartland, Lars-Erik Becken, and Karen Modesta Olsen. 1997. "Within-Job Wage Discrimination and the Gender Wage Gap, The Case of Norway." *European Sociological Review* 13(2):199-215.
- Reskin, Barbara F., and Patricia Roos (Eds.). 1990. *Job Queues, Gender Queues: Explaining Women's Inroads into Male Occupations*. Philadelphia: Temple University Press.
- Spilerman, Seymour. 1986. "Organizational Rules and the Features of Work Careers." *Research in Social Stratification and Mobility* 5:41-102.
- Tomaskovic-Devey, Donald. 1993. *Gender and Race Inequality at Work: The Sources and Consequences of Job Segregation*. Ithaca, NY: ILR Press.
- Treiman, Donald J., and Heidi I. Hartmann (Eds.). 1981. *Women, Work, and Wages: Equal Pay for Jobs of Equal Value*. Washington, D.C.: National Academy Press.
- U.S. Department of Labor. 1976. *Industry Wage Survey: Miscellaneous Plastics, September 1974*. Bureau of Labor Statistics, Bulletin 1914. Washington, D.C.: U.S. Government Printing Office.
- _____. 1977. *Dictionary of Occupational Titles*. 4th ed. Bureau of Employment Security. Washington, D.C.: U.S. Government Printing Office.
- _____. 1981. *National Survey of Professional, Administrative, Technical, and Clerical Pay, March 1981*. Bureau of Labor Statistics, Bulletin 2108. Washington, D.C.: U.S. Government Printing Office.
- U.S. Executive Office of the President. 1987. *Standard Industrial Classification Manual*. 1987. Office of Management and Budget. Washington, D.C.: U.S. Government Printing Office.

89 ■ Paula England

Devaluation and the Pay of Comparable Male and Female Occupations

The terms *comparable worth* and *pay equity* refer to a form of sex discrimination that went virtually unrecognized until the 1980s. Evidence abounds that jobs filled mostly by women have pay levels lower than they would be if the jobs were filled mostly by men. This is seen as sex discrimination by advocates of the principle of comparable worth, which posits that jobs that are equally demanding and onerous and of equal value to society should be paid equally. The theory positing that employers assign lower wages to some jobs because they are filled largely with women has come to be called the devaluation perspective. Petersen and Saporta (2004) have referred to this type of discrimination as “valuative discrimination.”

At first glance, the issue of comparable worth sounds like the more familiar issue of “equal pay for equal work,” which refers to men and women in the same job with the same seniority, performing the same work equally well, but being paid differently. Petersen and Saporta (2004) have called this “within job discrimination,” and it has been illegal in the United States since the 1963 Equal Pay Act and Title VII of the 1964 Civil Rights Act. However, comparable worth is distinct because it refers to comparisons between the pay in different jobs that entail at least some distinct tasks. Comparable worth

advocates compare jobs that are largely male with those that are largely female (which I will sometimes, for brevity, refer to as “male” and “female” jobs) and, in some cases, claim that the difference between the pay of the two jobs results from gender bias in wage setting rather than from other job characteristics. Needless to say, this approach creates a thorny issue in distinguishing when two distinct jobs are comparable in the sense that we would expect them to pay the same in the absence of sex discrimination.

The wage discrimination at issue in a comparable worth scenario is also distinct from discrimination in hiring, initial job placement, and promotion, all of which Petersen and Saporta (2004) refer to as “allocative discrimination.” Such discrimination is a violation of Title VII of the Civil Rights Act of 1964. Hiring discrimination against women seeking to enter traditionally male jobs is one (although not the only) reason for occupational sex segregation. There are supply-side processes at work as well, such as those that lead women and men to major in different fields in college. Without segregation of jobs, female jobs could not be given a discriminatory pay level. But engaging in discrimination on the basis of sex in setting the pay levels assigned to male and female jobs is analytically distinct from engaging in hiring

discrimination on the basis of sex. Moreover, valuative discrimination may exist whether or not the segregation came from allocative discrimination or from supply-side forces.

Some examples may help the reader to visualize the sorts of comparisons at issue in comparable worth. In 1975, nurses in Denver, Colorado, sued the city claiming that their jobs paid less than male jobs such as tree trimmer and sign painter. It would be hard to argue that the latter two jobs require as much skill or are as demanding as nursing. Women workers for the city of San Jose, California, discovered that secretaries were generally earning less than workers in male jobs that required no more than an eighth grade education, including, for example, men who washed cars for the city. Eventually, women in San Jose succeeded in getting the city to do a job evaluation study. The results showed, to choose some examples, that nurses earned \$9,120 less annually than fire truck mechanics and that legal secretaries made \$7,288 less annually than equipment mechanics. In 1985, the California School Employees Association complained that school librarians and teaching assistants (female jobs) were paid less than custodians and groundskeepers (male jobs). In recent years, the city of Philadelphia, Pennsylvania, was paying practical nurses (mostly women) less than gardeners (mostly men). These are not atypical examples. In addition, one is hard-pressed to come up with a single example of a male job paying less than a female job that reasonable people would find comparable in skill, effort, and difficulty of working conditions. Nor are these differences in pay a result of men averaging more years of seniority than women, since in the above comparisons (from England 1992: 2), a constant level of experience was assumed.

U.S. courts have ruled that, under most conditions, this type of valuative discrimination and lack of comparable worth is not a violation of U.S. law (Nelson and Bridges 1999; England 1992: Chapter 5). Studies on data from a single employer like those reviewed here are

not seen as evidence of illegal discrimination in wage setting. At present there is no legal remedy for this kind of discrimination unless there is a “smoking-gun” statement by employers that the reason they paid one job less than another was simply because it was filled with women. Statistical evidence of consistently biased pay levels is not enough. However, occasionally, some of the job evaluation studies have led private and city or state government entities to revise their pay scales to equalize salaries in jobs deemed comparable.

Economists propose two explanations for the lower pay of occupations with a high percent of female workers. The first is the theory of compensating differentials, which posits that the full pay of a job consists of both pecuniary (wage) and nonpecuniary compensation (the utility experienced from doing the work itself). Jobs with more comfortable, less hazardous working conditions can be filled with workers easily at lower wages, *ceteris paribus*. If women care more about nonpecuniary rewards (such as avoiding physical danger or having mother-friendly work conditions such as flexible hours or on-site child care centers) than men, and men focus more on maximizing earnings, then women will trade off earnings for amenities by choosing safer, more mother-friendly jobs. Most tests have failed to find that greater nonpecuniary amenities explain much of the lower pay of women’s jobs (England 1992; Glass 1990; Glass and Camarigg 1992; Jacobs and Steinberg 1990; Kilbourne et al. 1994). The idea that women might choose lower-paying jobs for other compensations seems on first glance consistent with the finding that mothers earn less than nonmothers, even after controlling for part-time work status, experience, and seniority (Budig and England 2001; Lundberg and Rose 2000; Waldfogel 1997). But neither Glass (1990) nor Glass and Camarigg (1992) found women’s jobs to have more mother-friendly characteristics than men’s jobs. Similarly, Budig and England (2001) could not find any job characteristics (except part-time

status) that reduced the motherhood wage penalty. Nonmothers were about as concentrated in female jobs as mothers. Thus, oddly enough, in the United States, segregation by sex does not seem to stem from women choosing or being consigned to more mother-friendly jobs. The “motherhood penalty” explains only some of the sex gap in pay, and this portion is largely unrelated to the effect of occupational sex segregation on the sex gap in pay. Thus, although the general theory of compensating differentials may have some merit, it cannot explain most of the lower pay of female jobs.

A second explanation offered by economists for the lower pay in female jobs is crowding. Bergmann (1986) argues that women’s jobs pay less because they are crowded. According to this view, women seeking to enter male occupations face sex discrimination in hiring (the allocative discrimination discussed above), leading to a supply of applicants for traditionally female jobs larger than it would be in the absence of hiring discrimination as women denied entrance to male jobs crowd the female jobs. The excess supply lowers wages in female jobs. Although this is plausible, it is very difficult to test directly.

Sociologists have proposed the devaluation thesis, which posits that gender bias leads to valuing (and thus paying) female jobs less than comparably skilled male jobs (England et al. 2000; Steinberg 2001). According to these sociologists, it doesn’t matter if segregation arises from allocative discrimination (in hiring and job placement) or not; employers engage in valuative discrimination once the segregation has occurred. Of course, the same evidence offered for this view would also be consistent with crowding. However, sociologists believe it is unlikely that hiring is the only place where bias occurs; it occurs in wage-setting between jobs as well.

Evidence that the sex composition of an occupation or job affects its wage level supports the devaluation theory. Such effects of sex composition, after statistically controlling for

the other factors already discussed, have led some researchers to conclude that employers set lower wages (relative to job demands) when jobs are filled largely by women. One type of study takes the U.S. Census Bureau’s detailed occupational categories as units of analysis. The measures include the percent female of those in the occupation, average or median male wage, median female wage, average education of incumbents, and a myriad of measures of skill demands from the *Dictionary of Occupational Titles*. These studies examine the effect of occupational percent female on average male and female wages (separately for men and women) after adding all the educational, skill, and working condition demands as control variables whose effects are adjusted out statistically. These studies find that both men and women earn less when in a more female occupation (England 1992; England et al. 2001). That is, in a regression predicting median wage, percent female of the occupation is significant and negative. (See Filer 1989 for a contrary finding.)

Other studies use individuals as units of analysis, and find that individuals in more female occupations earn less, after similar statistical controls to those described above (England et al. 2000; see Tam 1997, 2000 for a contrary view). Other studies take person/years as units of analysis, and percent female in occupations as contextual variables. The studies with person/years as units employ person-fixed effects to remove omitted variable bias; that is, they use persons as their own control and generally show that the same person earns more when moving to a more male-dominated occupation or earns less when moving to a more female-dominated occupation (Kilbourne et al. 1994; MacPherson and Hirsch 1995).

Other studies use jobs from a single employer as cases and show that female jobs pay less than male jobs when a job evaluation (of skill demands and working conditions) finds these jobs comparable (reviewed in England 1992).

All of these studies suggest that jobs pay either men or women less if they are filled mostly by women. The low pay of the largely female jobs cannot be entirely explained by how much skill the job requires or how onerous the working conditions are, because measures of these factors are controlled in the regression analyses that find a significant negative effect of the percent female in the occupation or job. These studies convince many sociologists that employers' evaluation of jobs is gendered—and that devaluation infects the wage levels of female jobs. Sociologists posit that this devaluation is perpetuated through cultural, institutional, and market mechanisms. Although the evidence of devaluation is quite strong, there is little direct evidence of the mechanisms through which it occurs.

I close with a somewhat speculative description of the process by which the devaluation of female jobs has probably occurred. At certain points in the life of organizations, particularly when they are first formed, or undergo significant restructuring, the relative wages of various jobs are somewhat up for grabs, although not unaffected by market forces. At these points, gender bias infects the decisions managers or consultants make about the wage levels for various jobs. In part, this results from a cognitive distortion in underestimating the relative contribution to profits or other organization goals of the work done in female jobs. In part, this bias on the part of organizational decision makers results from generalizing to women's jobs the relatively low status accorded women as persons by our culture. Later, bureaucratic inertia takes over, setting the wage consequences of these cultural biases in stone; the relative wage levels of jobs within an organization often remain quite constant for decades. Relative wage scales get institutionalized. If any attempt is made to change relative wage levels in a way favorable to female occupations, individual or collective action by male workers may increase the wages in male-domi-

nated jobs over those of female-dominated jobs again. Market processes then reproduce all these effects. That is, if most employers have at some point inscribed these biases into their wage structures, this affects the "going wage" in predominantly female and male occupations that other new organizations must pay to hire workers of typical quality into the occupation. These market forces dictate that even an organization not beset by these cognitive and cultural biases will be forced to pay a higher wage to recruit in male occupations than comparable female occupations. Although they could choose to pay more than they need to in female occupations—more than the market or average wage—few profit-minded employers will do so. In these ways, gender bias is perpetuated over time. Absent significant collective action or new legislation that requires comparable worth in pay scales, the devaluing social and inertial forces described above may keep the pay penalty for working in female jobs in force indefinitely.

REFERENCES

- Bergmann, Barbara. 1986. *The Economic Emergence of Women*. New York: Basic Books.
- Budig, Michelle J., and Paula England. 2001. "The Wage Penalty for Motherhood." *American Sociological Review* 66: 204–225.
- England, Paula. 1992. *Comparable Worth: Theories and Evidence*. New York: Aldine.
- England, Paula, Jennifer Thompson, and Carolyn Aman. 2001. "The Sex Gap in Pay and Comparable Worth: An Update." Pp. 551–556 in Ivar Berg and Arne Kalleberg, eds., *Sourcebook on Labor Markets: Evolving Structures and Processes*. New York: Plenum.
- England, Paula, Joan M. Hermsen, and David A. Cotter. 2000. "The Devaluation of Women's Work: A Comment on Tam." *American Journal of Sociology* 105: 1741–1751.
- Filer, Randall K. 1989. "Occupational Segregation, Compensating Differentials and Comparable Worth." Pp. 153–170 in Michael, Robert T., Heidi I. Hartmann, and Brigid O. Farrell, eds., *Pay Equity: Empirical Inquiries*. Washington, DC: National Academy Press.

- Glass, Jennifer. 1990. "The Impact of Occupational Segregation on Working Conditions." *Social Forces* 68: 779–796.
- Glass, Jennifer, and Valerie Camarigg. 1992. "Gender, Parenthood, and Job-Family Compatibility." *American Journal of Sociology* 98: 131–151.
- Jacobs, Jerry, and Ronnie Steinberg. 1990. "Compensating Differentials and the Male-Female Wage Gap: Evidence from the New York State Comparable Worth Study." *Social Forces* 69: 439–468.
- Kilbourne, Barbara S., Paula England, George Farkas, Kurt Beron, and Dorothea Weir. 1994. "Returns to Skill, Compensating Differentials, and Gender Bias: Effects of Occupational Characteristics on the Wages of White Women and Men." *American Journal of Sociology* 100: 689–719.
- Lundberg, Shelley, and Elaina Rose. 2000. "Parenthood and the Earnings of Married Men and Women." *Labour Economics* 7:689–710.
- MacPherson, D. A., and B. T. Hirsch. 1995. "Wages and Gender Composition: Why Do Women's Jobs Pay Less?" *Journal of Labor Economics* 13: 426–471.
- Nelson, Robert, and William Bridges. 1999. *Legalizing Gender Inequality*. New York: Cambridge University Press.
- Petersen, Trond, and Ishak Saporta. 2004. "The Opportunity Structure for Discrimination." *American Journal of Sociology* 109: 852–901.
- Steinberg, Ronnie. J. 2001. "Comparable Worth in Gender Studies." Pp. 2293–2397 in N. J. Smelser and P. B. Baltes, eds., *International Encyclopedia of the Social and Behavioral Sciences*, vol. 4. London: Elsevier.
- Tam, Tony. 1997. "Sex Segregation and Occupational Gender Inequality in the United States: Devaluation or Specialized Training?" *American Journal of Sociology* 102: 1652–1692.
- Tam, Tony. 2000. "Occupational Wage Inequality and Devaluation: A Cautionary Tale of Measurement Error." *American Journal of Sociology* 105: 1752–1760.
- Waldfogel, Jane. 1997. "The Effect of Children on Women's Wages." *American Sociological Review* 62: 209–217.

90 ■ Tony Tam

Why Do Female Occupations Pay Less?

Background

For the past two decades, scholars of gender inequality in the labor market have sought to understand why occupations with more female workers tend to pay less. This question addresses a robust finding about occupational sex composition and wages: studies based on different samples, units of analysis, and statistical techniques almost always find significant wage effects of occupational sex composition even after accounting for a wide range of personal attributes and occupational

characteristics. The evidence of lower wages in "female occupations" (i.e., occupations that are disproportionately female) may reflect subtle gender discrimination. Instead of wage discrimination against *female workers*, gender discrimination may take the form of wage discrimination against *female work*.

The finding of persistent wage effects of occupational sex composition is socially important because it offers crucial support for comparable worth policies. The comparable worth movement has long identified occupa-

tional wage inequality as a major source of gender inequity in the labor market. The movement claims that there is widespread evaluative discrimination against female work, as suggested by research in the laboratory and by surveys of prestige ratings of occupations. Thus the typical comparable worth policy in the United States seeks to equalize wages for occupations (within a firm) that differ mainly in sex composition.

Human Capital and Specialization

The main competing interpretations in the literature on occupational sex composition effects are (a) the devaluation hypothesis that female work is deemed less important and thereby *culturally devalued* (England et al. 1988), and (b) the specialized human capital hypothesis that female occupations require, on average, less substantial investments in *specialized skills* (Tam 1997). The devaluation hypothesis simply refers to a general cultural devaluation of women's labor—a special case of discrimination against women. In a female occupation, *all* workers in the occupation are subject to the devaluation effect. By contrast, the specialized human capital hypothesis claims that the apparent wage penalties against female occupations arise mainly from occupational differences in specialized training. In essence, the same worker is expected to receive different wages for occupations associated with different training requirements for specialized human capital, which may be occupation specific, industry specific, or firm specific.

Human capital may be general (e.g., reading skills and health) or specialized (e.g., knowledge of computer science). Although any investment in human capital is costly, an investor in specialized human capital faces an additional opportunity cost because specialized skills by definition have restricted applicability. For instance, a medical doctor cannot expect to recoup the training cost of attending medical school by becoming a high

school teacher, even if s/he may one day find teaching in high school the most personally gratifying occupation. But a specialized skill also tends to earn higher returns than does a general skill that requires the same amount of training time. The lower pay attached, on average, to female occupations may thus arise because such occupations require less substantial investments in specialized human capital. The sources of sex segregation by specialized training requirement are outside the purview of this paper, but one might hypothesize that (a) employers for occupations requiring specialized human capital may avoid hiring women, (b) specialized occupations discourage potential female recruits because the work environment is often hostile (by virtue of employer or coworker behavior), and (c) women have relatively stronger incentive than men to make training investments that can pay off in many lines of work.

The devaluation hypothesis of course does not deny the existence of market forces in wage determination. The key point of contention is whether gendered cultural bias has sufficiently distorted the market's wage-setting process to generate substantial, unexplained sex composition effects. The specialized human capital hypothesis claims that it does not and that the apparent sex composition effects can be explained by taking account of occupational differences in specialized human capital. This hypothesis is a plausible application of human capital theory but, unfortunately, has never been properly tested in the literature.

Weak Old Tests and Strong New Evidence

The empirical tests necessary for adjudicating between the two hypotheses are surprisingly difficult to achieve, as is the measurement of the extent and degree of discrimination in general (Heckman 1998). There are many traps that would invalidate a test or misguide the interpretation of a finding.¹ Recently I

TABLE 1
Description of Variables

| <i>Variable</i> | <i>Description</i> | <i>Mean</i> | <i>SD</i> |
|-----------------|--|-------------|-----------|
| Wage | Hourly wage is the ratio of weekly earnings and usual weekly hours. | 9.679 | 5.497 |
| SMSA | 1 = residing in a standard metropolitan statistical area, 0 = otherwise. | .725 | .446 |
| South | 1 = residing in the South, 0 = otherwise. | .303 | .459 |
| West | 1 = residing in the West, 0 = otherwise. | .181 | .385 |
| Midwest | 1 = residing in the Midwest, 0 = otherwise. | .272 | .445 |
| Veteran | 1 = ever in the armed forces, 0 = otherwise. | .160 | .366 |
| Union | 1 = union member or under union contract, 0 = otherwise. | .187 | .390 |
| School | Years of schooling completed. | 14.028 | 2.632 |
| Experience | Potential work experiences, estimated by age – school – 5. ¹ | 17.702 | 12.177 |
| Tenure | Years with current employer. | 6.779 | 7.721 |
| Full-time | 1 = usually work at least 35 hours a week, 0 = otherwise. | .829 | .376 |
| PWOM | Women as a proportion of workers in respondent's occupation. | .465 | .324 |
| SVP | Years of specific vocational training for an occupation. ² | 1.977 | 1.611 |

Note: The data are for whites in the expanded May 1988 CPS. Most of the valid *N*s are 23,019 or slightly less. Valid *N*s for experience and tenure are 19,666.

1. For consistency, negative values of experience are recoded "0."

2. From England and Kilbourne's (1988) match of SVP (*DOT*, 4th ed.) with three-digit 1980 census occupations.

have reexamined the evidence for the devaluation hypothesis and showed that there are strong reasons to question its validity (Tam 1997).

The first reason is a crucial bias in the evidential strategy. The establishment of general empirical support for the devaluation hypothesis essentially follows a "residual" approach because of the difficulty in developing any direct measure of devaluation in survey data. That is, the empirical support for the hypothesis crucially hinges on the existence of substantial and unexplained (i.e., residual) wage penalties against female occupations. The more comprehensive are the theoretically relevant control variables and the larger are the residual effects among workers with comparable characteristics, the stronger is the evidence in favor of the devaluation hypothesis. Virtually all quantitative tests of the devaluation hypothesis adopt this strategy. Nevertheless, this evidential strategy

is biased in favor of the devaluation hypothesis because of the potential for omitted-variable bias and, most important, measurement errors in correlated control variables (Tam 2000). Past tests and evidence have been produced and interpreted without taking any account of measurement error and thus fail to provide critical support for devaluation.

The second reason for skepticism is new counter-evidence that simultaneously contradicts the devaluation hypothesis and confirms the specialized human capital hypothesis (Tam 1997, 2000). This evidence uses the May 1988 Current Population Survey (CPS) as the main source of data. Appended to the CPS are two constructed variables: the sex composition (PWOM) of three-digit census occupations, and occupation-specific training time (SVP, measured in years) derived from the fourth edition of the Dictionary of Occupational Titles (DOT). Table 1 describes the key variables used in the analysis.

TABLE 2
 Regressions of Log Hourly Wages on Demographic, Human Capital, and Industry Variables

| | <i>White Women</i> | | | | <i>White Men</i> | | | |
|-------------------------|--------------------|-----------------|-----------------|------------------|------------------|-----------------|-----------------|-----------------|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| PWOM | -.158 (.016) | -.187 (.016) | .011# (.017) | -.007# (.018) | -.224 (.018) | -.094 (.019) | -.141 (.018) | .002# (.019) |
| SVP | | | .088 (.004) | .078 (.004) | | | .065 (.003) | .069 (.003) |
| INDUSTRY | | Yes (56.792) | | Yes (50.145) | | Yes (51.883) | | Yes (54.429) |
| Adjusted R ² | .403 | .470 | .440 | .496 | .452 | .503 | .482 | .532 |

Note: Variables included in all models are South, West, Midwest, SMSA, veteran, union, school, experience and its square, tenure and its square, and full-time. INDUSTRY denotes the control for 22 industries: construction, education, financial/insurance/real estate, wholesale, retail trade, medical (except hospital), hospital, agriculture, mining, manufacturing of durable goods, manufacturing of nondurable goods, postal, other transportation, other utilities, private household service, business/repair, personal services (except private household), entertainment/recreation, welfare/religious, other professional services, forestry/fisheries, and public administration/armed forces. Nos. in parentheses are standard errors, except that F-statistics are reported for industry effects. $N = 9,269$ for (1) and (2), 9,265 for (3) and (4), 10,383 for (5) and (6), 10,372 for (7) and (8). Data are for whites in the May 1988 CPS.

Not significant at the .05 level; other presented coefficients are significant at the .05 level.

Table 2 presents the core results that show the extent to which occupation-specific (SVP) and/or industry-specific (INDUSTRY) training can explain PWOM effects. The first major finding pertains to white female workers. The gross PWOM effect (-.255, not shown in the table) is substantially reduced (-.158) after controlling for a set of basic variables. This PWOM effect is still negative and substantial, apparently confirming a wage disadvantage against female occupations. However, after adding occupational differences in two measures of specialized human capital, column 4 of table 2 shows that there is virtually no occupational sex composition effect left to be explained. Comparing the coefficients of PWOM across columns 1-4 of table 2 reveals that SVP is the key explanatory variable. After controlling for SVP, the PWOM effects are dramatically reduced to .011 and -.007 without and with control for industry effects, respectively. This finding therefore contradicts the devaluation hypothesis at any reasonable level of statistical significance.

The devaluation hypothesis implies that women's work is devalued for all incumbents (i.e., men and women alike) and via the same cultural process. Columns 5-8 of table 2 examine this implication. After controlling for specialized human capital (SVP and INDUSTRY), the PWOM effect among white men becomes practically zero (.002), as was the case for white women (-.007). Equally significant, the model with INDUSTRY alone (column 6) or SVP alone (column 7) fails to fully account for the PWOM effect (-.094 and -.141, respectively) among white men. These results therefore suggest that the sex composition effects among men and women are attributable to gender-specific occupational differences in specialized human capital.

These two results strongly suggest that the PWOM effects are mediated by specialized human capital differences and that the sources of the wage effects of sex composition are gender specific. Both findings contradict the devaluation hypothesis. Moreover, both findings are remarkably robust across racial groups, as

shown in Tam (1997, p. 1674). There is absolutely no wage penalty against female occupations: not among women or men, not among whites or blacks.²

These striking results are hard to reconcile with a devaluation story but are consistent with the specialized human capital hypothesis. Occupational sex composition is just a handy proxy for unobserved specialized human capital investment. In a related article (Tam 2000), I have further demonstrated that even seemingly supportive findings for the devaluation hypothesis can be accurately predicted (i.e., numerically reproduced) under the specialized human capital hypothesis by assuming a modest level of measurement errors. In sum, the accumulated evidence to date is remarkably strong for the specialized human capital hypothesis.

Positional Capital and Allocative Inequality

Jobs are highly differentiated, interconnected, and form a social structure. Some jobs are more conducive to further opportunities for specialized training, and only certain jobs will generate returns to the specialized skills of a worker. Not surprisingly, jobs are often the focus of career investment, just as human capital is. The capital value of a job is high if it provides the necessary gateway to the accumulation of specialized human capital or the unique promotion path to other jobs that can make the most use of the skills of a worker. The “positional capital” of a worker is made up of the worker’s portfolio of ever-occupied jobs that can provide long-term returns in addition to income (Tam 1998). To the extent that male and female occupations offer different positional capital, any discriminatory job allocation by sex could have profound long-term impacts on women’s labor market outcomes even in the absence of devaluation of female occupations.

NOTES

This chapter draws on analyses and arguments that are presented in more detail elsewhere (Tam 1997, 2000). The research presented here was supported by Academia Sinica through the Organization-Centered Society Project and the Institute of European and American Studies.

1. For instance, it is poorly understood that market prices are not determined by the preferences of the average or majority of buyers and sellers, but set at the margin by the final set of exchanges that align supply and demand. The result is that discriminatory intentions, attitudes, and behavior can be pervasive while market wages remain competitive (nondiscriminatory) under a wide variety of scenarios.

2. As Tam (1997) has further demonstrated, these results are reinforced by the fact that there is a large wage advantage, instead of disadvantage, for female occupations among women working part-time. Also, see Tam (1997) for a discussion of various fallback defenses of the devaluation hypothesis, all of which can be shown to be untenable.

REFERENCES

- England, Paula, George Farkas, Barbara S. Kilbourne, and Thomas Dou. 1988. “Explaining Occupational Sex Segregation and Wages: Findings from a Model with Fixed Effects.” *American Sociological Review* 53:544–558.
- England, Paula, and Barbara S. Kilbourne. 1988. *Occupational Measures from the Dictionaries of Occupational Titles for 1980 Census Detailed Occupations* (MRDF). ICPSR no. 8942. Ann Arbor, Mich.: Inter-University Consortium for Political and Social Research.
- Heckman, James J. 1998. “Detecting Discrimination.” *Journal of Economic Perspectives* 12 (2):101–116.
- Tam, Tony. 1997. “Sex Segregation and Occupational Gender Inequality in the United States: Devaluation or Specialized Training?” *American Journal of Sociology* 102:1652–1692.
- . 1998. “Getting Ahead in the Labor Market: The Positional Capital Approach.” Paper presented at the 14th World Congress of Sociology, Montreal, Canada, July 26–August 1.
- . 2000. “Occupational Wage Inequality and Devaluation: A Cautionary Tale of Measurement Error.” *American Journal of Sociology* 105:1741–1760.

91 ■ Francine D. Blau and Lawrence M. Kahn

The Gender Pay Gap

Have Women Gone as Far as They Can?

After half a century of stability in the earnings of women relative to men, there has been a substantial increase in women's relative earnings since the late 1970s. One of the things that make this development especially dramatic and significant is that the recent changes contrast markedly with the relative stability of earlier years.

These post-1980 earnings changes are also interesting because, when you compare women to their male counterparts, gains have been prevalent across a wide spectrum. So, for example, at first much of the female gains were centered on younger women, but now, while the gains may be a bit larger for younger women, women of all ages have narrowed the pay gap with men. The same broad progress is visible when we look at the trends in the gender pay gap by education. Less-educated women have narrowed the pay gap with less-educated men and highly educated women have narrowed the pay gap with highly educated men.

The earnings gains of women are particularly remarkable because they have occurred during a period when overall wage inequality was rising. That is, the difference in pay between workers with high wages and workers with low wages has widened considerably over the past 25 years or so. And yet, women, a low paid group, have nonetheless been able to narrow the pay gap with a relatively higher paid group, men.

The foregoing supports our initial observation that there has been important, significant progress for women. On the other hand, however, there is still a gender pay gap. Women continue to earn considerably less than men on average. It is also true that convergence slowed noticeably in the 1990s after women had especially gained relative to men in the 1980s. Although there were some larger gains for women in the early 2000s, the long-run significance of this recent experience is unclear. With the evidence suggesting that convergence has slowed in recent years, the possibility arises that the narrowing of the gender pay gap will not continue into the future. Moreover, there is evidence that although discrimination against women in the labor market has declined, some discrimination does still continue to exist.

Trends in the Gender Pay Gap

In this section we look in more detail at the trends in the relative wages of women. Figure 1 presents data drawn from published government statistics on female-to-male earnings ratios of full-time workers. We focus on full-time workers to adjust for gender differences in hours worked. This is important because women are more likely than men to work part-time. Ideally we would like a measure of wages or an hourly rate of pay. Unfortunately, we do not have a similar long

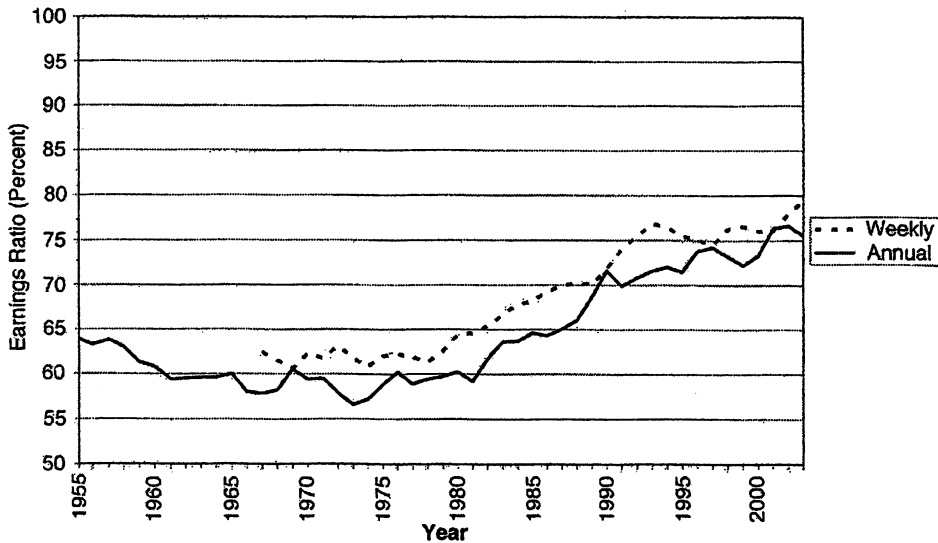


Figure 1. Female-to-Male Earnings Ratios of Full-Time Workers 1955–2003

data series for hourly wages. Thus, we focus here on the earnings of full-time workers.

The figure gives the gender earnings ratio for two data series available from published government statistics. Again, both pertain to the relative earnings of female and male full-time workers. The first, the annual earnings series, is based on annual earnings data on workers who are employed year round as well as full time. The second, the weekly earnings series, is based on the earnings of full-time workers over the survey week, regardless of how many weeks per year the individual works. The annual earnings series has been available for the longest time period, 1955 to 2003; the weekly earnings series has been available for a somewhat shorter period, 1967 to 2003.

While the exact figure for the gender earnings ratio differs a bit for the two series, they both tell the same story in terms of the trends. Until the late 1970s or early 1980s there was a remarkable constancy in the ratio, at around 60%. There were some year-to-year fluctuations, but the ratio hovered around the 60% level. Indeed, if there was any discernible trend, it was a decrease in the ratio between 1955 and 1960. Then, over the 1980s, we see a period of strong, sus-

tained increase in the ratio. This rising trend prevailed through perhaps 1990 or 1993, depending on the series. However, during the 1990s, the pace of convergence in both the annual and the weekly earnings series slowed and both series behaved more erratically. The pace of change picked up again in the early 2000s.¹ However, as noted above, the long-run significance of this recent experience is unclear. It may signal a resumption of a strong, long-run trend towards convergence in male–female earnings or may prove to be of only short duration.²

Abstracting from the differential trends over the various subperiods and focusing on the period since the late 1970s as a whole, the gains have been quite remarkable, especially viewed in terms of the long constancy in the gender ratio that preceded this time. So, for example, based on the weekly earnings series, the gender ratio rose from 61.3% in 1978 to 79.4% in 2003. Again, much of this increase was accomplished in a relatively short period of time, with the ratio reaching 76.8% by 1993. Of course, the 77–79% figure remains below earnings parity. Thus, clearly all sources of the pay differential between men and women have not been eradicated.

How do we explain these earnings gains for women? To address this question as well as to understand why women continue to earn less than men, we need to first consider the basic factors that explain the gender pay gap to begin with.

Economists' Explanations for the Gender Pay Gap

The Role of Qualifications and Discrimination

Economists point to a number of factors that could be important in explaining the lower earnings of women compared to men, but traditionally have focused on two primary factors. Following Juhn, Murphy, and Pierce (1991), we call these “gender-specific” factors in that they relate specifically to differences between women and men, either in their qualifications or how they are treated. With regard to qualifications, the human capital model has been especially important in pointing out the potential role played by education and experience.

The gender gap in educational attainment was never particularly large in the United States. The biggest difference historically was that, although women were more likely to graduate from high school than men, they were less likely to go on to college and graduate education. Moreover, men tended to concentrate in career-oriented fields of study such as engineering, law, medicine and business that led to relatively high earnings. These educational differences have decreased quite a bit in recent years, especially at the college level where women are actually now over half of college students; women have also greatly increased their representation in traditionally-male professional fields. Thus gender differences in education levels have never explained a large portion of the overall gender pay gap; most recently, in some samples gender differences in years of schooling favor women.

The qualification that has proven to be quite important is work experience because tradi-

tionally women moved in and out of the labor market based on family considerations. Before World War II, most women left the labor market permanently when they got married and had children. In the immediate post-war period, a pattern arose whereby older married women returned to the labor market after their children were in school or grown. An even bigger change has occurred in the past 20 to 30 years as increasing numbers of women, including married women, started staying in the labor force fairly continuously even when they had small children at home. Today, even the majority of women with children a year or less in age are participating in the labor force. Nonetheless, on average, women have less work experience than men and that difference in qualifications is quantitatively important in explaining the gender pay gap.

Jacob Mincer and Solomon Polachek (1974) have done especially important work in highlighting the role of labor market experience in explaining the gender pay gap. Given the traditional division of labor by gender in the family, women tend to accumulate less labor market experience than men. Further, because women anticipate shorter and more discontinuous work lives, they have lower incentives to invest in market-oriented formal education and on-the-job training. Their resulting smaller human capital investments lower their earnings relative to those of men. An additional way in which the traditional division of labor may disadvantage women is that the longer hours women spend on housework may also decrease the effort they put into their market jobs compared to men, controlling for hours worked, and hence also reduce their productivity and wages (Becker, 1985).

To the extent that women choose occupations for which on-the-job training is less important, gender differences in occupations are also expected. Women may especially avoid jobs requiring large investments in skills that are unique to a particular enterprise, because the returns to such investments are reaped only as long as one remains with that employer. At

the same time, employers may be reluctant to hire women for such jobs because the firm bears some of the costs of such firm-specific training, and fears not getting a full return on that investment.

However, even controlling for experience and whatever other qualifications can readily be measured, there tends to be a pay difference between men and women that is not explained and is potentially due to discrimination. Gary Becker (1971; 1st ed., 1957) has been especially instrumental in developing analyses of labor market discrimination. Although he was looking at differences between blacks and whites, the idea of prejudice and its negative consequences are readily transferable to women versus men. Becker conceptualized discriminatory preferences as the desire to maintain social distance from the discriminated group. It may at first seem odd to hypothesize that men would not like to associate with women on the job when they generally live together with women in families. However, the issue here may be more one of socially appropriate roles than of the desire to maintain social distance, as Becker postulated was the case with race.

Standard models in economics suggest discrimination can arise in a variety of ways. In Becker's model, discrimination is due to the discriminatory tastes of employers, co-workers, or customers. Alternatively, in models of "statistical discrimination," differences in the treatment of men and women arise from average differences between the two groups in the expected value of productivity (or in the reliability with which productivity may be predicted), which may lead employers to discriminate on the basis of that average (see for example, Aigner & Cain, 1977). Finally, discriminatory exclusion of women from "male" jobs can result in an excess supply of labor in "female" occupations, depressing wages there for otherwise equally productive workers, as in Bergmann's (1974) "overcrowding" model.

The typical approach to analyzing the sources of the gender pay gap is to estimate

wage regressions specifying the relationship between wages and productivity-related characteristics for men and women. The gender pay gap may then be statistically decomposed into two components: one due to gender differences in measured characteristics, and the other "unexplained" and potentially due to discrimination. Such empirical studies provide evidence consistent with both human capital differences and labor market discrimination in explaining the gender pay gap.

However, any approach that relies on a statistical residual will be open to question as to whether all the necessary explanatory variables were included in the regression. For example, even if measured human capital characteristics can explain only a portion of the wage gap between men and women, it is possible that unmeasured group differences in qualifications may explain part of the residual. If men are more highly endowed with respect to these omitted variables then we would overestimate discrimination. Alternatively, if some of the factors controlled for in such regressions—like occupation and tenure with the employer—themselves reflect the impact of discrimination, then discrimination will be underestimated. Moreover, if women face barriers to entry into certain occupations, they may have higher unmeasured productivity than men in the same jobs. This factor would also suggest an underestimate of discrimination if we controlled for occupation.³

Using the residual from a regression to estimate the effects of discrimination will also run into trouble if feedback effects are important. Even small initial discriminatory differences in wages may cumulate to large ones as men and women make decisions about human capital investments and time allocation in the market and the home on the basis of these wage differentials.

Results of statistical studies of the gender pay gap may nonetheless be instructive. Representative findings from analyses of this type may be illustrated by results from a recent paper of ours (Blau & Kahn, 2006). Using

data from the Panel Study of Income Dynamics (PSID), which contains information on actual labor market experience for a large, nationally representative sample, we found a wage differential between male and female full-time workers in 1998 of 20%. The restriction to full-time workers is designed to focus on male and female workers who are as similar as possible.⁴

The impact of gender differences in characteristics on the male-female wage differential is shown in Table 1. The variables considered include indicators of “human capital,” that is, those relating to education and experience, as well as measures of occupation, industry and union status. (Race is also included as a control variable, but its effect is small since the proportion of each race group in the full-time sample is about the same for men and women.)

As would be expected, women’s lesser amount of labor market experience is found to be a significant determinant of the gender wage differential, explaining 11% of the gender gap in wages. This reflects a 3.5 year difference in full-time experience between men and women, which, though smaller than in previous years, is still a substantial factor explaining the wage gap. Interestingly, women in this sub-sample are found to have higher educational attainment than men, which (as indicated by the negative sign in the table) works to *lower* the gender wage gap by 7%. Putting this somewhat differently, gender differences in educational attainment do not help to explain the gender wage gap, but rather work slightly in the opposite direction. While in the population as a whole, men’s educational attainment is still somewhat higher than women’s, when we focus on a sub-sample of the population which is not only employed, but employed full time, women have a slight edge.

Finally, gender differences in occupation and industry are substantial and help to explain a considerable portion of the gender wage gap. Men are more likely to be in blue-collar jobs and to work in mining, construc-

Table 1. Contribution to the Wage Differential Between Men and Women of Differences in Measured Characteristics, 1998

| <i>Characteristics</i> | <i>Percent Explained</i> |
|------------------------|--------------------------|
| Educational attainment | -6.7 |
| Labor force experience | 10.5 |
| Race | 2.4 |
| Occupational category | 27.4 |
| Industry category | 21.9 |
| Union status | 3.5 |
| Unexplained | 41.1 |
| Total | 100.0 |
| Wage differential (%) | 20.3 |

Source: Calculated from data presented in Blau and Kahn (2006).

tion, or durable manufacturing; they are also more likely to be in unionized employment. Women are more likely to be in clerical or professional jobs and to work in the service industry. Taken together, these variables explain 53% of the gender wage gap—27% for occupation, 22% for industry, and an additional 4% for union status.⁵

Although these findings suggest that gender differences in work-related characteristics are important, they also indicate that qualifications are only part of the story. The proportion of the wage differential that is *not* explained by these types of productivity-related characteristics includes the impact of labor market discrimination, although as mentioned above, the residual may also include the effects of gender differences in unmeasured productivity levels or non-wage aspects of jobs. In this case, 41% of the gender gap cannot be explained even when gender differences in education, experience, industries, occupations, and union status are taken into account. We can consider the results of this study somewhat differently by focusing on the gender wage ratio. The actual (“unadjusted”) gender wage ratio is 80%; that is, women’s wages are, on average, 80% of men’s wages. If women had the same human capital characteristics (that is, education and experience), racial composition, industry and occupational distribution, and union coverage

as men, the “adjusted” ratio would rise to 91% of men’s wages. Thus, while measured characteristics are important, women still earn less than similar men even when all measured characteristics are taken into account. And, as we suggested above, including controls for occupation, industry, and union status may be questionable to the extent that they may be influenced by discrimination.

Nonetheless, the residual gap, however measured, may well reflect factors apart from discrimination. One that has received particular attention recently is the impact of children on women’s wages, since evidence of a negative effect of children on wages has been obtained, even in analyses which control for labor market experience (Waldfogel, 1998). The reason may be that, in the past, having a child often meant that a woman withdrew from the labor force for a substantial period, breaking her tie to her employer and forgoing the returns to any firm-specific training she might have acquired, as well as any rewards for having made an especially good job match. Given the sharp increase in the labor force participation of women with young children that has occurred since the 1960s, this factor may have been of growing importance in influencing the aggregate gender gap. However, the greater availability of parental leave, legally mandated in the United States since 1993, may well mitigate the effect of this factor on more recent cohorts. Indeed, Waldfogel finds that the negative effect of children on wages is substantially reduced for mothers who have maternity leave coverage.

Some studies of discrimination have taken different approaches to the question, thus avoiding some of the problems of traditional analyses. First, two studies have applied traditional econometric techniques to especially homogeneous groups and employed extensive controls for qualifications, thus minimizing the effect of gender differences in unmeasured productivity characteristics. Wood, Corcoran, and Courant (1993) studied grad-

uates of the University of Michigan Law School classes of 1972–1975, 15 years after graduation. The gap in pay between women and men was relatively small at the outset of their careers, but 15 years later, women graduates earned only 60% as much as men. Some of this difference reflected choices that workers had made, including the propensity of women lawyers to work shorter hours. But, even controlling for current hours worked, as well as an extensive list of worker qualifications and other covariates, including family status, race, location, grades while in law school, and detailed work history data, such as years practiced law, months of part-time work, and type and size of employer, a male advantage of 13% remained. In a similar vein, Weinberger (1998) examined wage differences among recent college graduates in 1985. Her controls included narrowly defined college major, college grade point average, and specific educational institution attended. She found an unexplained pay gap of 10 to 15% between men and women.

A second set of studies used an experimental approach. Neumark (1996) analyzed the results of a hiring “audit” in which male and female pseudojob seekers were given similar résumés and sent to apply for jobs waiting on tables at the same set of Philadelphia restaurants. In high-priced restaurants, a female applicant’s probability of getting an interview was 40 percentage points lower than a male’s and her probability of getting an offer was 50 percentage points lower. A second study examined the impact of the adoption of “blind” auditions by symphony orchestras in which a screen is used to conceal the identity of the candidate (Goldin & Rouse, 2000). The screen substantially increased the probability that a woman would advance out of preliminary rounds and be the winner in the final round. The switch to blind auditions was found to explain 25% of the increase in the percentage female in the top five symphony orchestras in the United States, from less than 5% of all musicians in 1970 to 25% in 1996.

Third, several recent studies have examined predictions of Becker's (1971) discrimination model. Becker and others have pointed out that competitive forces should reduce or eliminate discrimination in the long run because the least discriminatory firms, which hire more lower-priced female labor, would have lower costs of production and should drive the more discriminatory firms out of business. For this reason, Becker suggested that discrimination would be more severe in firms or sectors that are shielded to some extent from competitive pressures. Consistent with this reasoning, Hellerstein, Neumark, and Troske (2002) found that, among plants with high levels of product market power, those employing relatively more women were more profitable. In a similar vein, Black and Strahan (2001) report that, with the deregulation of the banking industry beginning in the mid-1970s, the gender pay gap in banking declined as men's wages fell by considerably more than women's (12% vs. 3%). This suggests that during the period of regulation, banks shared the rents fostered by regulation primarily with men. It was thus men who lost the most in the shift to deregulation. And, Black and Brainerd (2004) find that increasing vulnerability to international trade reduced apparent gender wage discrimination in concentrated industries, again as predicted by Becker's (1971) model.

Finally, additional evidence on discrimination comes from court cases. A number of employment practices which explicitly discriminated against women used to be quite prevalent; including marriage bars restricting the employment of married women (Goldin, 1990), and the intentional segregation of men and women into separate job categories with associated separate and lower pay scales for women (e.g., *Bowe v. Colgate-Palmolive Co.*, 416 F.2d 711 [7th Cir. 1969]; *IUE v. Westinghouse Electric Co.*, 631 F.2d 1094 [3rd Cir. 1980]). While many such overt practices have receded, recent court cases

suggest that employment practices still exist which produce discriminatory outcomes for women.

For example, in 1994, Lucky Stores, a major grocery chain, agreed to a settlement of \$107 million after Judge Marilyn Hall Patel found that "sex discrimination was the standard operating procedure at Lucky with respect to placement, promotion, movement to full-time positions, and the allocation of additional hours" (*Stender v. Lucky Stores, Inc.* 803 F. Supp. 259; [N.D. Cal. 1992]; King 1997). And, in 2000, the U.S. Information Agency agreed to pay \$508 million to settle a case in which the Voice of America rejected women who applied for high-paying positions in the communications field. A lawyer representing the plaintiffs said that the women were told things like, "These jobs are only for men," or "We're looking for a male voice" (FEDHR, 2000). A final example is the 1990 case against Price Waterhouse, a major accounting firm, in which the only woman considered for a partnership was denied, even though, of the 88 candidates for partner, she had brought in the most business. Her colleagues criticized her for being "overbearing, 'macho' and abrasive and said she would have a better chance of making partner if she would wear makeup and jewelry, and walk, talk and dress 'more femininely.'" The Court found that Price Waterhouse maintained a partnership evaluation system that "permitted negative sexually stereotyped comments to influence partnership selection" (BNA, 1990; Lewin, 1990).

Oftentimes, economists serve as expert witnesses in court cases alleging discrimination. Their analyses, when publicly available, provide a window into discriminatory practices that still exist to some extent in the labor market, although there is of course likely to be disagreement between experts employed by each side in the type of evidence that is relevant or in the interpretation of the evidence. For example, the Lucky

Stores case cited above generated an interesting exchange summarized in Taylor (2001).⁶

Labor economist John Pencavel testified for the plaintiffs, the women who brought the suit. He found that women at Lucky earned between 76 percent and 82 percent as much as Lucky's male workers earned. Pencavel found that women were regularly placed in jobs that paid less than jobs given male coworkers, although there was no significant difference between the education and experience of the workers. There was little difference in the wages of the male and female workers within each type of job; but some jobs paid more than others and women happened to be assigned to the lower-paying jobs.

Joan Haworth, another labor economist, was an expert witness for the defendant, Lucky Stores. She reported survey evidence showing that Lucky's assignment of women and men to different jobs reflected differences in the work preferences of men and women. Thus, Lucky justified its job assignments by arguing that there was a gender difference in attitudes toward work. Lucky argued that its employment policies were based on observed differences in the career aspirations of male and female employees. For example, one manager at Lucky testified that women were more interested in cash register work and men were more interested in floor work.

As we noted above, Judge Marilyn Hall Patel decided the case in favor of the plaintiffs. With respect to the evidence cited above, she wrote: "The court finds the defendant's explanation that the statistical disparities between men and women at Lucky are caused by differences in the work interests of men and women to be unpersuasive." An interesting aspect of this case is that both sides agreed that male and female employees received equal pay for equal work and that the pay differential was associated with pay differences across occupations. They differed, however, over the source of the occupational

differences: the choices of women vs. discrimination. This disagreement mirrors the alternative explanations economists offer in general for wage and occupational difference between men and women: differences in qualifications based on the choices men and women make versus discrimination which limits the opportunities and pay of women compared to men.

Some additional evidence supporting discrimination as a source of the type of occupational differences cited above is provided by a recent study of eight years of data from an unidentified regional grocery chain on gender differences in job titles and wage rates (Ransom & Oaxaca, 2005). As in the case of the Pencavel analysis summarized above, Ransom and Oaxaca find a pattern of gender differences in initial job assignment and upward mobility within the firm that "generally penalized women, even when the analysis account[ed] for individuals' characteristics" (p. 219). While one might again dispute the reason for these differences, the authors found that job segregation of women and men was dramatically lower in the period after the company lost a discrimination suit (1984) and reached a settlement (1986) in which it initiated affirmative action policies. This implies that it was possible to find women interested in higher-level jobs, leading one to doubt that such segregation was entirely voluntary.

These cases emphasize the role of occupational segregation by sex within firms in producing pay differences between men and women. Pencavel explicitly notes that there was little difference in pay between men and women in the same job. It is worth noting that economists and sociologists who have examined this issue across a wider range of firms have tended to come to a similar conclusion: pay differences between men and women in the same narrowly-defined occupational categories within the same firm tend to be small (Blau, 1977; Groshen, 1991; Petersen & Morgan, 1995; and Bayard, Heller-

stein, Neumark, & Troske, 1999). However, even when men and women are in the same occupation, they tend to be segregated by firm, and such establishment segregation contributes substantially to the gender pay gap.

The Role of Wage Structure

In earlier work, building on a framework suggested by Chinhui Juhn, Kevin Murphy, and Brooks Pierce (1991), we point out that there is another factor that needs to be considered when analyzing gender differences in pay, and that is what we call wage structure (Blau & Kahn, 1996 and 1997). We define wage structure as being the market returns to skills and the rewards for employment in particular sectors of the economy. Market returns to skills denote the premiums the market determines for being a more experienced worker or a more highly educated worker, etc. Rewards for employment in particular sectors of the economy refer to the fact that, for example, unionized workers tend to earn more than comparable nonunionized workers or workers in some industries—durable goods, manufacturing for example—may earn more than similarly-qualified workers in other industries, say services. In addition, considerable research suggests that predominantly female occupations pay less, even controlling for measured personal characteristics of workers and a variety of characteristics of occupations, although the interpretation of such results remains in some dispute.⁷

We distinguish wage structure from gender-specific factors because the idea is that these are the returns to skills or the rewards for working in a particular industry or occupation regardless of whether you are male or female. Why should wage structure affect the gender pay gap? To see how, let's think a bit more about the two factors we discussed earlier—gender differences in qualifications and labor market discrimination. Suppose women do have less experience, on average, than men do. Then, the higher the return to experience the larger the gender pay gap will

be. Or, suppose that jobs staffed primarily by women do pay less than jobs staffed primarily by men. Then, the higher the premium for being in a male occupation the larger the gender pay gap will be.

This is interesting because these market returns have in fact varied over time. In the last 25 years or so, the market returns to skills, like those acquired with work experience, have increased. So this is a factor that, taken alone, would have worked to increase the gender pay gap. The rewards to being in male occupations and industries have increased as well, and that factor, taken alone, would have increased the pay gap as well. So, one question that we have raised in our research is: How have women been able to successfully swim against the tide of rising returns to skills and rising rewards to being in particular industries and occupations? That is, how have they managed to narrow the pay differential with men in the face of the adverse trends in wage structure that have worked against them?

Before looking at the results of our research addressing these questions, let's consider the issue of why the returns to skills have been increasing. There is a fairly broad consensus among economists (though not complete unanimity) that within countries like the United States, one of the main reasons that the returns to skills have been rising is that the demand by employers for skilled workers has been rising relative to the demand for unskilled workers. Why has this occurred? There are at least two reasons. The one that that we would put the most weight on is technological change. The information and telecommunications revolution has worked to put more of a premium on skill, at least thus far. There are other scenarios possible, but thus far it has increased the demand for skilled workers compared to less skilled workers. The other reason—we would put less weight on it although it has also played a role—is international trade. Today, less skilled workers in the United States are to

some extent competing against less skilled workers from around the world; many of them are available at much lower wages. Factors in addition to demand shifts that appear to have also played a role are—a decline in the union movement since unions tend to push for more egalitarian pay structures, the falling real value the minimum wage (adjusted for inflation, the minimum wage is actually lower today than it was in the 1970s), an influx of unskilled immigrants, and a decrease in the rate of growth of college-educated workers.

While rising returns to skills may be hypothesized to widen the gender pay gap, all else equal, it is possible that the demand shifts discussed above may have favored women relative to men in certain ways, and thus contributed to a decrease in the unexplained gender pay gap (Blau & Kahn, 1997; Welch, 2000). Technological change is believed to have caused within-industry demand shifts that favored white collar workers in general (Berman, Bound, & Griliches, 1994). Given the traditional male predominance in blue-collar jobs, this shift might be expected to benefit women relative to men. Similarly, to the extent that the spread of computer technology is an important source of recent technological change, the observation that women are more likely than men to use computers at work suggests that women as a group may have benefited from shifts in demand associated with computerization (Autor, Katz, & Krueger, 1998; Weinberg, 2000). Diffusion of computers likely also benefits women because computers restructure work in ways that de-emphasize physical strength (Weinberg, 2000).

Explaining the Trends: The 1980s

Returning to the trends in the gender pay gap—how do we explain them? To answer this question, we summarize results from Blau and Kahn (1997 and 2006). Using data from the PSID (we reported on some of our results above), we analyzed women's

wage gains over the 1980s (1979–1989), which, as we saw in Figure 1, was a period of exceptionally rapid closing of the gender wage gap. We found that higher rewards to skills did indeed retard wage convergence during this period but this was more than offset by improvements in gender-specific factors.

Of particular importance was the decline in the experience difference between men and women: the gender gap in full-time experience fell from 7.5 to 4.6 years over this period. Shifts in major occupations played a significant role too, as the employment of women as professionals and managers rose relative to men's, while their relative employment in clerical and service jobs fell. Women's wages also increased relative to men's because of deunionization (the decline of unions). Deunionization had a larger negative impact on male than female workers because men, who have traditionally been more likely than women to be unionized, experienced a larger decrease in unionization than women. Another factor that worked to increase the gender pay ratio substantially was a decrease in the "unexplained" portion of the gender differential—that is, a decline in the pay difference between men and women with the same measured characteristics (i.e., experience, education, occupation, industry, and union status).

Taken together, changes in qualifications and in the unexplained gap worked to increase the gender wage ratio substantially. Working in the opposite direction, however, were changes in wage structure (or returns to characteristics) that favored men over women during this period. Of particular importance were a rise in the return to experience (since women have less of it) and increases in returns to employment in industries where men are more highly represented. These shifts in labor market returns by themselves would have reduced the gender ratio substantially. Thus, in order for the wage gap to decline, the factors favorably af-

fecting women's wages had to be large enough to more than offset the impact of unfavorable shifts in returns. This was indeed the case, so that the gender pay gap did decline over the 1980s.

Can we say anything about the reasons for the decline in the unexplained gender wage gap that occurred over the 1980s? Such a shift may reflect a decline in labor market discrimination against women, but also an upgrading of women's *unmeasured* labor market skills, a shift in labor market demand favoring women over men, or changes in the composition of the labor force due to the pattern of labor force entries or exits. Indeed all of these factors may well have played a role, and all appear credible during this period.

First, since women improved their relative level of measured skills, as shown by the narrowing of the gap in full-time job experience and in occupational differences between men and women, it is plausible that they also enhanced their relative level of unmeasured skills. For example, women's increasing labor force attachment may have encouraged them to acquire more on-the-job training or encouraged their employers to offer them more training. Evidence also indicates that gender differences in college major, which have been strongly related to the gender wage gap among college graduates (Brown & Corcoran, 1997), decreased over the 1970s and 1980s (Blau, Ferber, & Winkler, 2002); the marketability of women's education has probably improved. The male-female difference in SAT math scores has also been declining, falling from 46 points in 1977 to 35 points in 1996 (Blau, Ferber, & Winkler, 2002), which could be another sign of improved quality of women's education.

Second, the argument that discrimination against women declined in the 1980s may seem less credible than that their unmeasured human capital characteristics improved, since the federal government scaled back its antidiscrimination enforcement ef-

fort during the 1980s (Leonard, 1989). However, as women increased their commitment to the labor force and improved their job skills, the rationale for statistical discrimination against them diminished; thus it is plausible that this type of discrimination decreased. Further, in the presence of feedback effects, employers' revised views can generate additional increases in women's wages by raising women's returns to investments in job qualifications and skills. To the extent that such qualifications are not fully controlled for in the statistical analysis used to explain the change in the gender wage gap, this may also help to account for the decline in the "unexplained" gap. Another possible reason for a decline in discrimination against women is that changes in social attitudes have made such discriminatory tastes increasingly less acceptable.

Third, the underlying labor market demand shifts that widened wage inequality over the 1980s may have favored women relative to men in certain ways, and thus may have also contributed to a decrease in the unexplained gender gap. Overall, manufacturing employment declined. In addition, there is some evidence that technological change produced within-industry demand shifts that favored white-collar relative to blue-collar workers in general. As noted above, given the traditional male predominance in blue-collar jobs, this shift might be expected to benefit women relative to men, as would increased computer use.

Finally, another factor contributing to the considerable narrowing of the "unexplained" gender wage gap in the 1980s appears to be favorable shifts in the composition of the female labor force. Specifically, we found that, controlling for the measured characteristics mentioned earlier, the women who entered the labor force over this period tended to be those with relatively high (unmeasured) skills. This improved the quality of the female labor force and thus contributed to the narrowing of the gender wage gap. . . .

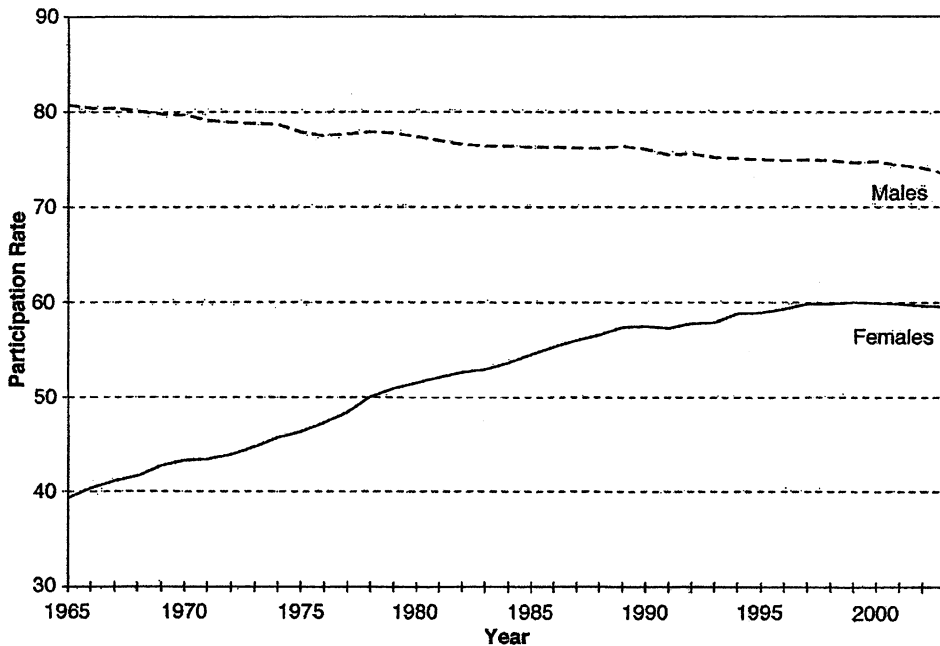


Figure 2. Trends in Female and Male Labor Force Participation Rates, 1965–2003

Explaining the Trends: The 1990s

Why did convergence in female and male wages slow over the 1990s? Again, drawing on our previous work (Blau & Kahn, 2006) we may suggest some tentative answers. We found that human capital trends cannot account for the slowdown: women improved their relative human capital by about the same amount in both the 1980s and the 1990s. In the 1980s this upgrading consisted of rising relative experience while in the 1990s it consisted to a lesser extent of rising relative experience and to a greater extent of increasing educational attainment of women relative to men. Nor did changes in wage structure in the 1990s have a more adverse effect on women than changes in the previous decade—in fact the impact of changing wage structure was actually more negative for women in the 1980s. Slowing convergence in men's and women's occupations and degree of unionization in the 1990s was found to account for some of the slowdown, but only a small portion.

We found that the major reason for the slowdown in wage convergence in the 1990s was the considerably smaller narrowing of the “unexplained” gender pay gap in the 1990s compared to the 1980s. Our reasoning above suggests that this could be due to slower improvement in women's unmeasured qualifications relative to men's in the 1990s than in the 1980s; a smaller decline in discrimination against women in the 1990s than in the 1980s; or less favorable demand shifts for women in the 1990s than in the 1980s. Each of these factors appears to have played a role in explaining the observed trends. In addition, controlling for measured characteristics, female labor force entrants were less skilled during the 1990s, perhaps as a result of the entry of many relatively low-skilled, female single-family heads. Indeed, differences between the two decades in such shifts in labor force composition were found to explain as much as 25% of the apparent slowdown in convergence in the unexplained gender pay gap in the 1990s.

As we noted above, women narrowed the experience gap at a slower pace in the 90s than they did in the 80s. Figure 2 shows the trends in male and female labor force participation that underlie this development. The most striking trend shown in the figure is that the *difference* in the participation rates of men and women has narrowed considerably since the starting year, 1965. This is due to a slow steady decrease in male labor force participation combined with a much sharper and dramatic increase in female labor force participation. The decrease in male participation does not appear to be due very much to changes in gender roles. Rather it primarily reflects the fact that men are retiring at earlier ages and are staying in school longer. Another factor has been the weakening job market for less skilled men (Juhn, 1992).

While the data in the figure begin in the mid-1960s, the large increases in female participation in fact date back to the 1940s. Interestingly, the trend towards rising female labor force participation was strong and consistent until about 1990. After that the line becomes noticeably flatter. Women's participation increased a bit through 1997, with no further increases thereafter.

How do these participation trends relate to the average experience levels of women workers? Unfortunately, it is not possible to figure this out just by looking at participation rates. This is because the labor force participation rate of women can increase for either of two reasons or a combination of both. On the one hand, participation may rise because a lot of new groups of women come into the labor market. This tends to lower the average experience of women workers because there are a lot of new entrants. On the other hand, participation can increase because women stay in the labor force more consistently over a period of time, rather than moving in and out. This works to raise average experience levels of women workers.

Research has shown that during the 1970s the average experience of women did not in-

crease because those two factors counter balanced each other. There were a lot of new entrants and a lot of women staying in more continuously, thus average experience remained about the same (Goldin, 1990). In the 1980s, though, the increase in the labor force participation of women was due to more of them remaining in the labor force more consistently. And, as we have seen, the average experience of women workers rose accordingly. This suggests that the flattening of the trend in female labor force participation shown in the chart caused the gender gap in experience to decline more slowly in the 1990s than in the 1980s.

Before leaving the subject of the participation trends, it is interesting to consider their larger significance. Viewed more broadly, what the trends show is an enormous change in gender roles and a movement away from the traditional family of breadwinner husband and homemaker wife to a family where both husband and wife work outside the home, although not necessarily giving equal weight to each of their careers. Ralph Smith (1979) called this process a "subtle revolution." The trends suggest that this subtle revolution, having accomplished a great deal, may be slowing down now. Is it stopping? Not necessarily. But we have reached a situation where, looking at women in the prime working ages (that is, 25 to 55), over three quarters of them are in the labor force. This means that female labor force participation rates in the United States are very high, although still below the male rates of around 90% in this age group. So it may not be surprising that, of necessity, future participation trends will be less dramatic than past trends.

Prospects for the Future

Although we readily acknowledge that predicting the future is a tricky business, we cautiously offer some thoughts on the prospects for the future. What will happen to the gender pay gap in the coming years? Recent developments make the answer to this

question particularly uncertain. As we have seen, after a period of consistent and sustained narrowing of the gender pay gap over the 1980s, convergence became more fitful in the 1990s. Perhaps what we saw in the 1990s was a mere pause; perhaps we were consolidating the really massive changes that had occurred over the preceding 10 to 20 years—not just in the gender pay gap but also in women's labor force participation and in the occupations in which they work; perhaps the next 20 years will show similar renewed gains on all these fronts. That could very well be. Or we may have reached a point that we are going to stay at for quite a while, a big change from the past but not so much change in the future. It is even possible that under certain circumstances the gender pay gap could begin to widen, returning to levels of an earlier period. While we cannot choose among these options with certainty, it may be instructive to consider each of the factors that we have identified as influencing the gender pay gap and consider the possible future course of each and its likely impact on the pay gap.

One of the factors influencing the trends in the gender pay gap is overall trends in wage inequality. Rising wage inequality, to the extent that it results from increasing returns to skills like work experience that women have less of than men, on average, is expected to widen the gender pay gap, all else equal. In this respect, it has been noted that wage inequality increased less during the 1990s than during the 1980s (Katz & Autor, 1999). If this tapering off in the trend towards rising inequality should continue into the future, the negative effect of this factor on the pace of convergence in the gender pay gap will be small.

On the other hand, to the extent that rising wage inequality was due to demand shifts that favored women relative to men, it may be hypothesized that such shifts, and the relative advantage they may have given women

relative to men, have also tapered off and are likely to be dampened in the future as well. This is suggested by the fact that the shedding of blue collar, manufacturing jobs was particularly pronounced in the 1980s. A closely related development, deunionization, which also disadvantaged men to a greater extent than women, is likely to occur at a slower pace in the future as unionization rates in the private sector have reached single digits, giving little scope for substantial future declines, and public sector unionization remains relatively stable. While demand shifts favoring women may have slowed, so too has the growth in the supply of women to the labor market. If this slower growth in supply continues into the future, it may mean that demand- and supply-side shifts offset each other and thus, taken together, do not have much effect on convergence in the gender pay gap.

Although overall female labor force participation increased modestly in the 1990s, welfare reforms and other government policies spurred an increase in employment among single mothers (see, for example, Meyer & Rosenbaum, 2001). The growth in participation among single heads, who tend on average to be less well educated than other women, could also have slowed wage convergence by shifting the composition of the female labor force toward low-wage women. We did indeed find some evidence consistent with this in our earlier work (Blau & Kahn, 2006), though this factor does not appear to be the main reason for the slowing convergence in the gender wage gap in the 1990s. Thus it seems unlikely that a further entry of single mothers into the labor force in the future will have a large impact on the *aggregate* gender pay gap.

Moving toward the more traditional factors of women's relative qualifications and the possibility of labor market discrimination against them, there is little reason to expect large changes here either. The flattening

of the growth in women's labor force participation rates, if it continues, suggests that large increases in women's work experience and labor force commitment are unlikely, although this statement must be qualified somewhat since, as we have explained, trends in the average experience of women cannot be inferred directly from changes in participation rates. Similarly, now that women comprise the majority of college students, further large gains in the relative educational attainment of women appear unlikely, though there is room for continued reductions in the gender differences in college major and at the graduate level in professional schools and Ph.D. programs in many fields.

Turning to labor market discrimination, now that the most open and egregious forms of sex discrimination have been greatly reduced or eliminated and discrimination increasingly becomes more subtle and possibly even unconscious, future large declines in discrimination in the labor market may become more difficult to attain. In addition, the decreases in statistical discrimination that we hypothesized as occurring in response to women's increasing labor force attachment can be expected to slow as increases in women's attachment also slow. However, there seems room for some further decrease in statistical discrimination as the profound changes in gender roles that have already occurred continue to percolate through the labor market and the larger society and as additional changes continue to occur, albeit at a slower pace than in the past. And it is likely that even subtle barriers do change as women increasingly enter new areas and achieve success at higher levels. Putting this somewhat differently, while the glass ceiling may not have broken completely, it is showing a lot of cracks and is likely to show more and more cracks as time goes on.

Taking all these factors into account, our best guess is that we are going to have fur-

ther changes in the direction of convergence, but most probably at a slower pace. Our own view is that one development that is extremely unlikely is that we will see a reversal of the gains in relative wages and labor force participation women have experienced over the past 25 to 30 years. We do not expect a substantial widening of the male/female pay gap or labor force participation gap to occur. On the other hand, while precisely how much narrowing we will see in the future is an open question, the gender pay gap seems unlikely to vanish in the near term.

For one thing, women continue to confront discrimination in the labor market, and, although its extent seems to be decreasing, it seems unlikely to be completely eliminated soon. In addition, at least some of the remaining pay gap is surely tied to the gender division of labor in the home, both directly through its effect on women's labor force attachment and indirectly through its impact on the strength of statistical discrimination against women. Women still retain primary responsibility for housework and child care in most American families. However, this pattern has been changing as families respond to rising labor market opportunities for women that increase the opportunity cost of such arrangements. Further, policies that facilitate the integration of work and family responsibilities, both voluntary and government-mandated, have become increasingly prevalent in recent years. Employers are likely to continue to expand such policies as they respond to the shifting composition of the work force and a desire to retain employees in whom they have made substantial investments. In the longer run, the increasing availability of such policies will make it easier for women to combine work and family, and also for men to take on a greater share of household tasks.

Finally, while our principal concern has been with the pay of women relative to men,

Table 2. Mean Earnings of Education Groups Relative to High School Graduates, 1974 and 2003 (%)

| <i>Education</i> | 1974 | | 2003 | |
|------------------|------------|--------------|------------|--------------|
| | <i>Men</i> | <i>Women</i> | <i>Men</i> | <i>Women</i> |
| High school: | | | | |
| 1–3 years | 88.9 | 85.3 | 75.9 | 76.6 |
| 4 years | 100.0 | 100.0 | 100.0 | 100.0 |
| College: | | | | |
| 1–3 years | 113.6 | 112.6 | 122.8 | 119.5 |
| 4 or more years | 155.0 | 147.2 | 211.3 | 190.4 |

Data refer to year-round, full-time workers 18 years of age and older. In 2003, median income for 1–3 years of college is computed as a weighted average of the medians for “some college, no degree” and “associate degree.”

Source: 2003: PINC04 Tables of the U.S. Census Bureau Current Population Survey, 2004 Annual Social and Economic Supplement, from http://ferrets.census.gov/macro/032004/perinc/new04_000.htm; 1974: U.S. Census Bureau Historical Income Tables – People, Table P–35, from www.census.gov/hhes/income/histinc/p35.html.

trends in inequality among women show a deterioration in the relative economic status of less educated women that is strikingly parallel to similar trends in the labor market for men (see Table 2). These developments for less educated women serve to underscore the widening gap between more and less skilled Americans of both sexes, as well as to emphasize its broad dimensions.

NOTES

1. Between 1980 and 1990, the average annual increase in the ratio was 1.14 percentage points for annual earnings and .74 percentage points for weekly earnings, while, between 1990 and 2000, it was only .16 percentage points for annual earnings and .42 percentage points for weekly earnings. Relative earnings growth in the early 2000s was more robust: between 2000 and 2003, the average annual increase in the ratio was .75 percentage points for annual earnings and 1.14 percentage points for weekly earnings.

2. One short-term factor could be the recession of 2001 and the relatively high unemployment rates that lingered in its aftermath. The demand for male workers tends to be more cyclically sensitive than that for female workers due to their greater concentration in blue-collar jobs and durable manufacturing industries.

3. If, as is likely, one is unable to completely control for nonwage job characteristics such as fringe

benefits, safety, or job security, then the residual may again not give an accurate estimate of the extent of discrimination against women. We cannot say a priori what the effect of such omissions is. On the one hand, to the extent that men are likely to work in less safe or less secure jobs than women, such analyses may overestimate discrimination. On the other hand, to the extent that men have higher fringe benefit levels, an analysis of wage residuals will understate discrimination. To some degree, these nonwage characteristics can be accounted for by controlling for industry and occupation, although as mentioned, these controls may also reflect exclusionary hiring practices.

4. In addition to gender differences in qualifications and the extent of discrimination, the gender earnings differential may also be affected by the self-selection of women and men into full-time employment and, more generally, into the labor force. In other words, those choosing to participate—or to work full-time—may differ from those outside the labor force or part-time workers in terms of both their measured and unmeasured characteristics. One possibility, for example, is that labor force participants are a positively selected group of those who have received higher wage offers. Similarly, full-time workers may be more highly qualified and more committed to market work. We in fact find that, at a point in time, the gender pay gap is smaller if only full-time workers are considered than if part-timers and nonparticipants are included; we examine the impact of changes in female and male selection into the labor force for trends in the gender pay gap (see below). Other research that has examined the earn-

ings differential for white and black women has found that, if self-selection is not accounted for, the race differential is underestimated; see Neal (2004).

5. The study controls for 19 occupations and 25 industries.

6. This quotation is from the Textbook Site for *Principles of Microeconomics*, 3rd ed., *Additional Topics*, "Using Economics to Explain Gender Pay Gaps," at http://college.hmco.com/economics/taylor/econ/3e/micro/students/add_topics/ch02_genderpay.html, accessed June 28, 2002. The summary is based on materials presented in West's Federal Supplement (1993).

7. See, for example, Sorensen (1990), Kilbourne, England, Farkas, Beron, and Weir (1994), and Macpherson and Hirsch (1995).

REFERENCES

- Aigner, D., & Cain, G. (1977). Statistical theories of discrimination in labor markets. *Industrial and Labor Relations Review*, 30, 175–187.
- Autor, D.H., Katz, L.F., & Krueger, A.B. (1998). Computing inequality: Have computers changed the labor market? *Quarterly Journal of Economics*, 113, 1169–1214.
- Bayard, K., Hellerstein, J., Neumark, D., & Troske, K. (1999). New evidence on sex segregation and sex difference in wages from matched employee-employer data. NBER Working Paper No. 7003, March.
- Becker, G.S. (1957). *The economics of discrimination*, 2 ed. Chicago: University of Chicago Press, 1971; 1st ed.
- Becker, G.S. (1985). Human capital, effort, and the sexual division of labor. *Journal of Labor Economics*, 3, S33–S58.
- Bergmann, B. (1974). Occupational segregation, wages, and profits when employers discriminate by race or sex. *Eastern Economic Journal*, 1, 1–2, 103–110.
- Berman, E., Bound, J., & Griliches, Z. (1994). Changes in the demand of skilled labor within U.S. manufacturing industries: Evidence from the annual survey of manufacturing. *Quarterly Journal of Economics*, 109, 367–397.
- Black, S.E., & Brainerd, E. (2004). The impact of globalization on gender discrimination. *Industrial & Labor Relations Review*, 57, 540–559.
- Black, S.E., & Strahan, P.E. (2001). The division of spoils: Rent-sharing and discrimination in a regulated industry. *American Economic Review*, 91, 814–831.
- Blau, F.D. (1977). *Equal pay in the office*. Lexington, MA: Lexington Books.
- Blau, F.D. (1998). Trends in the well-being of American women, 1970–1995. *Journal of Economic Literature*, 36, 112–165.
- Blau, F.D., Ferber, M.A., & Winkler, A.E. (2006). *The economics of women, men, and work*, 5th ed. Upper Saddle River, NJ: Prentice-Hall.
- Blau, F.D., & Kahn, L.M. (1996). Wage structure and gender earnings differentials: An international comparison. *Economica*, 63, S29–S62.
- Blau, F.D., & Kahn, L.M. (1997). Swimming upstream: Trends in the gender wage differential in the 1980s. *Journal of Labor Economics*, 15, 1–42.
- Blau, F.D., & Kahn, L.M. (2006). The US gender pay gap in the 1990s: Slowing convergence. *Industrial & Labor Relations Review*, 60, 45–66.
- Brown, C., & Corcoran, M. (1997). Sex-based differences in school content and the male/female wage gap. *Journal of Labor Economics*, 15, 431–65.
- Bureau of National Affairs (BNA). (1990). *Daily Labor Report*, no. 235, December 6, pp. A11–A13 and F1–F10.
- Federal Human Resources Week (FEDHR). (2000, April 5). Government to pay \$508 million for sex discrimination at U.S. Information Agency. Vol. 6, p. 47.
- Goldin, C. (1990). *Understanding the gender gap*. New York: Oxford University Press.
- Goldin, C., & Rouse, C. (2000). Orchestrating impartiality: The impact of 'blind' auditions on female musicians. *American Economic Review*, 90, 715–741.
- Groshen, E.L. (1991). The structure of the female/male wage differential: Is it who you are, what you do, or where you work? *Journal of Human Resources*, 26, 457–472.
- Hellerstein, J.K., Neumark, D., & Troske, K. (2002). Market forces and sex discrimination. *Journal of Human Resources*, 37, 353–380.
- Juhn, C. (1992). Decline of male labor market participation: The role of declining market opportunities. *Quarterly Journal of Economics*, 107, 79–121.
- Juhn, C., Murphy, K.M., & Pierce, B. (1991). Accounting for the slowdown in black-white wage convergence. In M. Koster (Ed.), *Workers and their wages* (pp. 107–143). Washington, D.C.: AEI Press.
- Katz, L.F., & Autor, D.H. (1999). Changes in the wage structure and earnings inequality. In C.O. Ashenfelter & D. Card (Eds.), *Handbook of labor*

- economics* (vol. 3A, pp. 1463–1555). Amsterdam: Elsevier.
- Kilbourne, B.S., England, P., Farkas, G., Beron, K. & Weir, D. (1994). Returns to skill, compensating differentials, and gender bias: Effects of occupational characteristics on the wages of white women and men. *American Journal of Sociology*, 100, 689–719.
- King, R. (1997, April 27). Women taking action against many companies. *The Times-Picayune*.
- Lewin, T. (1990, May 16). Partnership awarded to woman in sex bias case. *The New York Times*, pp. A1, A12.
- Leonard, J. (1989). Women and affirmative action. *Journal of Economic Perspectives*, 3, 61–75.
- MacPherson, D.A., & Hirsch, B.T. (1995). Wages and gender composition: Why do women's jobs pay less? *Journal of Labor Economics*, 13, 426–471.
- Meyer, B.D., & Rosenbaum, D.T. (2001). Welfare, the earned income tax credit, and the labor supply of single mothers. *Quarterly Journal of Economics*, 116, 1063–1114.
- Mincer, J., & Polachek, S. (1974). Family investments in human capital: Earnings of women. *Journal of Political Economy*, 82, S76–S108.
- Neal, D. (2004). The measured black-white wage gap among women is too small. *Journal of Political Economy*, 112, S1–S28.
- Neumark, D.M. (1996). Sex discrimination in restaurant hiring: An audit study. *Quarterly Journal of Economics*, 111, 915–941.
- Petersen, T., & Morgan, L.A. (1995). Separate and unequal: Occupation-establishment sex segregation and the gender wage gap. *American Journal of Sociology*, 10, 329–361.
- Ransom, M., & Oaxaca, R.L. (2005). Intrafirm mobility and sex differences in pay. *Industrial and Labor Relations Review*, 58, 219–237.
- Smith, R.E. (1979). The movement of women into the labor force. In R.E. Smith (Ed.), *The subtle revolution: Women at work* (pp. 1–29). Washington, DC: Urban Institute.
- Sorensen, E. (1990). The crowding hypothesis and comparable worth issue. *Journal of Human Resources*, 25, 55–89.
- Taylor, J.B. (2001). *Principles of economics*, 3rd ed. Boston, MA: Houghton Mifflin.
- Waldfoegel, J. (1998). The family gap for young women in the United States and Britain: Can maternity leave make a difference? *Journal of Labor Economics*, 16, 505–545.
- Welch, F. (2000). Growth in women's relative wages and in inequality among men: One phenomenon or two? *American Economic Review*, 90, 444–449.
- West's Federal Supplement*, (1993). Vol. 803, pp. 259–337. St. Paul, MN: West Publishing Company.
- Weinberg, B. (2000). Computer use and the demand for female workers. *Industrial and Labor Relations Review*, 53, 290–308.
- Weinberger, C.J. (1998). Race and gender wage gaps in the market for recent college graduates. *Industrial Relations*, 37, 67–84.
- Wood, R.G., Corcoran, M.E., & Courant, P. (1993). Pay differences among the highly paid: The male-female earnings gap in lawyers' salaries. *Journal of Labor Economics*, 11, 417–441.

PART VIII

The Consequences of Inequality

Theory

92 ■ Thorstein Veblen

The Theory of the Leisure Class

Pecuniary Emulation

The end of acquisition and accumulation is conventionally held to be the consumption of the goods accumulated—whether it is consumption directly by the owner of the goods or by the household attached to him and for this purpose identified with him in theory. This is at least felt to be the economically legitimate end of acquisition, which alone it is incumbent on the theory to take account of. Such consumption may of course be conceived to serve the consumer's physical wants—his physical comfort—or his so-called higher wants—spiritual, æsthetic, intellectual, or what not; the latter class of wants being served indirectly by an expenditure of goods, after the fashion familiar to all economic readers.

But it is only when taken in a sense far removed from its naïve meaning that consumption of goods can be said to afford the incentive from which accumulation invariably proceeds. The motive that lies at the root of ownership is emulation; and the same motive of emulation continues active in the further development of the institution to which it has given rise and in the development of all those features of the social structure which this institution of ownership touches. The possession of wealth confers honor; it is an invidi-

ous distinction. Nothing equally cogent can be said for the consumption of goods, nor for any other conceivable incentive to acquisition, and especially not for any incentive to the accumulation of wealth.

It is of course not to be overlooked that in a community where nearly all goods are private property the necessity of earning a livelihood is a powerful and ever-present incentive for the poorer members of the community. The need of subsistence and of an increase of physical comfort may for a time be the dominant motive of acquisition for those classes who are habitually employed at manual labor, whose subsistence is on a precarious footing, who possess little and ordinarily accumulate little; but it will appear in the course of the discussion that even in the case of these impecunious classes the predominance of the motive of physical want is not so decided as has sometimes been assumed. On the other hand, so far as regards those members and classes of the community who are chiefly concerned in the accumulation of wealth, the incentive of subsistence or of physical comfort never plays a considerable part. Ownership began and grew into a human institution on grounds unrelated to the subsistence minimum. The dominant incentive was from the outset the invidious dis-

inction attaching to wealth, and, save temporarily and by exception, no other motive has usurped the primacy at any later stage of the development. . . .

In any community where goods are held in severalty it is necessary, in order to ensure his own peace of mind, that an individual should possess as large a portion of goods as others with whom he is accustomed to class himself; and it is extremely gratifying to possess something more than others. But as fast as a person makes new acquisitions, and becomes accustomed to the resulting new standard of wealth, the new standard forthwith ceases to afford appreciably greater satisfaction than the earlier standard did. The tendency in any case is constantly to make the present pecuniary standard the point of departure for a fresh increase of wealth; and this in turn gives rise to a new standard of sufficiency and a new pecuniary classification of one's self as compared with one's neighbors. So far as concerns the present question, the end sought by accumulation is to rank high in comparison with the rest of the community in point of pecuniary strength. So long as the comparison is distinctly unfavorable to himself, the normal, average individual will live in chronic dissatisfaction with his present lot; and when he has reached what may be called the normal pecuniary standard of the community, or of his class in the community, this chronic dissatisfaction will give place to a restless straining to place a wider and ever-widening pecuniary interval between himself and this average standard. The invidious comparison can never become so favorable to the individual making it that he would not gladly rate himself still higher relatively to his competitors in the struggle for pecuniary reputability.

In the nature of the case, the desire for wealth can scarcely be satiated in any individual instance, and evidently a satiation of the average or general desire for wealth is out of the question. However widely, or equally, or "fairly," it may be distributed, no general increase of the community's wealth can make

any approach to satiating this need, the ground of which is the desire of everyone to excel everyone else in the accumulation of goods. If, as is sometimes assumed, the incentive to accumulation were the want of subsistence or of physical comfort, then the aggregate economic wants of a community might conceivably be satisfied at some point in the advance of industrial efficiency; but since the struggle is substantially a race for reputability on the basis of an invidious comparison, no approach to a definitive attainment is possible.

What has just been said must not be taken to mean that there are no other incentives to acquisition and accumulation than this desire to excel in pecuniary standing and so gain the esteem and envy of one's fellowmen. The desire for added comfort and security from want is present as a motive at every stage of the process of accumulation in a modern industrial community; although the standard of sufficiency in these respects is in turn greatly affected by the habit of pecuniary emulation. To a great extent this emulation shapes the methods and selects the objects of expenditure for personal comfort and decent livelihood.

Besides this, the power conferred by wealth also affords a motive to accumulation. That propensity for purposeful activity and that repugnance to all futility of effort which belong to man by virtue of his character as an agent do not desert him when he emerges from the naïve communal culture where the dominant note of life is the unanalyzed and undifferentiated solidarity of the individual with the group with which his life is bound up. When he enters upon the predatory stage, where self-seeking in the narrower sense becomes the dominant note, this propensity goes with him still, as the pervasive trait that shapes his scheme of life. The propensity for achievement and the repugnance to futility remain the underlying economic motive. The propensity changes only in the form of its expression and in the proximate objects to which it directs the man's activity. Under the

regime of individual ownership the most available means of visibly achieving a purpose is that afforded by the acquisition and accumulation of goods; and as the self-regarding antithesis between man and man reaches fuller consciousness, the propensity for achievement—the instinct of workmanship—tends more and more to shape itself into a straining to excel others in pecuniary achievement. Relative success, tested by an invidious pecuniary comparison with other men, becomes the conventional end of action. The currently accepted legitimate end of effort becomes the achievement of a favorable comparison with other men; and therefore the repugnance to futility to a good extent coalesces with the incentive of emulation. It acts to accentuate the struggle for pecuniary reputability by visiting with a sharper disapproval all shortcoming and all evidence of shortcoming in point of pecuniary success. Purposeful effort comes to mean, primarily, effort directed to or resulting in a more creditable showing of accumulated wealth. Among the motives which lead men to accumulate wealth, the primacy, both in scope and intensity, therefore, continues to belong to this motive of pecuniary emulation.

In making use of the term “invidious,” it may perhaps be unnecessary to remark, there is no intention to extol or depreciate, or to commend or deplore any of the phenomena which the word is used to characterize. The term is used in a technical sense as describing a comparison of persons with a view to rating and grading them in respect of relative worth or value—in an æsthetic or moral sense—and so awarding and defining the relative degrees of complacency with which they may legitimately be contemplated by themselves and by others. An invidious comparison is a process of valuation of persons in respect of worth.

Conspicuous Leisure

If its working were not disturbed by other economic forces or other features of the em-

ulative process, the immediate effect of such a pecuniary struggle as has just been described in outline would be to make men industrious and frugal. This result actually follows, in some measure, so far as regards the lower classes, whose ordinary means of acquiring goods is productive labor. This is more especially true of the laboring classes in a sedentary community which is at an agricultural stage of industry, in which there is a considerable subdivision of property, and whose laws and customs secure to these classes a more or less definite share of the product of their industry. These lower classes can in any case not avoid labor, and the imputation of labor is therefore not greatly derogatory to them, at least not within their class. Rather, since labor is their recognized and accepted mode of life, they take some emulative pride in a reputation for efficiency in their work, this being often the only line of emulation that is open to them. For those for whom acquisition and emulation is possible only within the field of productive efficiency and thrift, the struggle for pecuniary reputability will in some measure work out in an increase of diligence and parsimony. But certain secondary features of the emulative process, yet to be spoken of, come in to very materially circumscribe and modify emulation in these directions among the pecuniarily inferior classes as well as among the superior class.

But it is otherwise with the superior pecuniary class, with which we are here immediately concerned. For this class also the incentive to diligence and thrift is not absent; but its action is so greatly qualified by the secondary demands of pecuniary emulation, that any inclination in this direction is practically overborne and any incentive to diligence tends to be of no effect. The most imperative of these secondary demands of emulation, as well as the one of widest scope, is the requirement of abstention from productive work. . . . During the predatory culture labor comes to be associated in men's

habits of thought with weakness and subjection to a master. It is therefore a mark of inferiority, and therefore comes to be accounted unworthy of man in his best estate. By virtue of this tradition labor is felt to be debasing, and this tradition has never died out. On the contrary, with the advance of social differentiation it has acquired the axiomatic force due to ancient and unquestioned prescription.

In order to gain and to hold the esteem of men it is not sufficient merely to possess wealth or power. The wealth or power must be put in evidence, for esteem is awarded only on evidence. And not only does the evidence of wealth serve to impress one's importance on others and to keep their sense of importance alive and alert, but it is of scarcely less use in building up and preserving one's self-complacency. In all but the lowest stages of culture the normally constituted man is comforted and upheld in his self-respect by "decent surroundings" and by exemption from "menial offices." Enforced departure from his habitual standard of decency, either in the paraphernalia of life or in the kind and amount of his everyday activity, is felt to be a slight upon his human dignity, even apart from all conscious consideration of the approval or disapproval of his fellows.

The archaic theoretical distinction between the base and the honorable in the manner of a man's life retains very much of its ancient force even today. So much so that there are few of the better class who are not possessed of an instinctive repugnance for the vulgar forms of labor. We have a realizing sense of ceremonial uncleanness attaching in an especial degree to the occupations which are associated in our habits of thought with menial service. It is felt by all persons of refined taste that a spiritual contamination is inseparable from certain offices that are conventionally required of servants. Vulgar surroundings, mean (that is to say, inexpensive) habitations, and vulgarly productive occupations are unhesitatingly condemned and

avoided. They are incompatible with life on a satisfactory spiritual plane—with "high thinking." From the days of the Greek philosophers to the present, a degree of leisure and of exemption from contact with such industrial processes as serve the immediate everyday purposes of human life has ever been recognized by thoughtful men as a prerequisite to a worthy or beautiful, or even a blameless, human life. In itself and in its consequences the life of leisure is beautiful and ennobling in all civilized men's eyes.

This direct, subjective value of leisure and of other evidences of wealth is no doubt in great part secondary and derivative. It is in part a reflex of the utility of leisure as a means of gaining the respect of others, and in part it is the result of a mental substitution. The performance of labor has been accepted as a conventional evidence of inferior force; therefore it comes itself, by a mental shortcut, to be regarded as intrinsically base. . . .

Conspicuous Consumption

In the earlier phases of the predatory culture the only economic differentiation is a broad distinction between an honorable superior class made up of the able-bodied men on the one side, and a base inferior class of laboring women on the other. According to the ideal scheme of life in force at that time it is the office of the men to consume what the women produce. Such consumption as falls to the women is merely incidental to their work; it is a means to their continued labor, and not a consumption directed to their own comfort and fullness of life. Unproductive consumption of goods is honorable, primarily as a mark of prowess and a perquisite of human dignity; secondarily it becomes substantially honorable in itself, especially the consumption of the more desirable things. The consumption of choice articles of food, and frequently also of rare articles of adornment, becomes tabu to the women and

children; and if there is a base (servile) class of men, the tabu holds also for them. With a further advance in culture this tabu may change into simple custom of a more or less rigorous character; but whatever be the theoretical basis of the distinction which is maintained, whether it be a tabu or a larger conventionality, the features of the conventional scheme of consumption do not change easily. When the quasi-peaceable stage of industry is reached, with its fundamental institution of chattel slavery, the general principle, more or less rigorously applied, is that the base, industrious class should consume only what may be necessary to their subsistence. In the nature of things, luxuries and the comforts of life belong to the leisure class. Under the tabu, certain victuals, and more particularly certain beverages, are strictly reserved for the use of the superior class. . . .

As wealth accumulates, the leisure class develops further in function and structure, and there arises a differentiation within the class. There is a more or less elaborate system of rank and grades. This differentiation is furthered by the inheritance of wealth and the consequent inheritance of gentility. With the inheritance of gentility goes the inheritance of obligatory leisure; and gentility of a sufficient potency to entail a life of leisure may be inherited without the complement of wealth required to maintain a dignified leisure. Gentle blood may be transmitted without goods enough to afford a reputedly free consumption at one's ease. Hence results a class of impecunious gentlemen of leisure. These half-caste gentlemen of leisure fall into a system of hierarchical gradations. Those who stand near the higher and the highest grades of the wealthy leisure class, in point of birth, or in point of wealth, or both, outrank the remoter-born and the pecuniarily weaker. These lower grades, especially the impecunious, or marginal, gentlemen of leisure, affiliate themselves by a system of dependence or fealty to the great ones; by so doing they gain an increment of repute, or of the means with which

to lead a life of leisure, from their patron. They become his courtiers or retainers, servants; and being fed and countenanced by their patron they are indices of his rank and vicarious consumers of his superfluous wealth. Many of these affiliated gentlemen of leisure are at the same time lesser men of substance in their own right; so that some of them are scarcely at all, others only partially, to be rated as vicarious consumers. So many of them, however, as make up the retainers and hangers-on of the patron may be classed as vicarious consumers without qualification. Many of these again, and also many of the other aristocracy of less degree, have in turn attached to their persons a more or less comprehensive group of vicarious consumers in the persons of their wives and children, their servants, retainers, etc.

Throughout this graduated scheme of vicarious leisure and vicarious consumption the rule holds that these offices must be performed in some such manner, or under some such circumstance or insignia, as shall point plainly to the master to whom this leisure or consumption pertains, and to whom therefore the resulting increment of good repute of right inures. The consumption and leisure executed by these persons for their master or patron represents an investment on his part with a view to an increase of good fame. . . .

With the disappearance of servitude, the number of vicarious consumers attached to any one gentleman tends, on the whole, to decrease. The like is of course true, and perhaps in a still higher degree, of the number of dependents who perform vicarious leisure for him. In a general way, though not wholly nor consistently, these two groups coincide. The dependent who was first delegated for these duties was the wife, or the chief wife; and, as would be expected, in the later development of the institution, when the number of persons by whom these duties are customarily performed gradually narrows, the wife remains the last. In the higher grades of society a large volume of both these kinds of ser-

vice is required; and here the wife is of course still assisted in the work by a more or less numerous corps of menials. But as we descend the social scale, the point is presently reached where the duties of vicarious leisure and consumption devolve upon the wife alone. In the communities of the Western culture, this point is at present found among the lower middle class.

And here occurs a curious inversion. It is a fact of common observance that in this lower middle class there is no pretense of leisure on the part of the head of the household. Through force of circumstances it has fallen into disuse. But the middle-class wife still carries on the business of vicarious leisure, for the good name of the household and its master. In descending the social scale in any modern industrial community, the primary fact—the conspicuous leisure of the master of the household—disappears at a relatively high point. The head of the middle-class household has been reduced by economic circumstances to turn his hand to gaining a livelihood by occupations which often partake largely of the character of industry, as in the case of the ordinary business man of today. But the derivative fact—the vicarious leisure and consumption rendered by the wife, and the auxiliary vicarious performance of leisure by menials—remains in vogue as a conventionality which the demands of respectability will not suffer to be slighted. It is by no means an uncommon spectacle to find a man applying himself to work with the utmost assiduity, in order that his wife may in due form render for him that degree of vicarious leisure which the common sense of the time demands.

The leisure rendered by the wife in such cases is, of course, not a simple manifestation of idleness or indolence. It almost invariably occurs disguised under some form of work or household duties or social amenities, which prove on analysis to serve little or no ulterior end beyond showing that she does not occupy herself with anything that is gainful or

that is of substantial use. The greater part of the customary round of domestic cares to which the middle-class housewife gives her time and effort is of this character. Not that the results of her attention to household matters, of a decorative and mundificatory character, are not pleasing to the sense of men trained in middle-class proprieties; but the taste to which these effects of household adornment and tidiness appeal is a taste which has been formed under the selective guidance of a canon of propriety that demands just these evidences of wasted effort. The effects are pleasing to us chiefly because we have been taught to find them pleasing. There goes into these domestic duties much solicitude for a proper combination of form and color, and for other ends that are to be classed as æsthetic in the proper sense of the term; and it is not denied that effects having some substantial æsthetic value are sometimes attained. Pretty much all that is here insisted on is that, as regards these amenities of life, the housewife's efforts are under the guidance of traditions that have been shaped by the law of conspicuously wasteful expenditure of time and substance. If beauty or comfort is achieved—and it is a more or less fortuitous circumstance if they are—they must be achieved by means and methods that commend themselves to the great economic law of wasted effort. The more reputable, "presentable" portion of middle-class household paraphernalia are, on the one hand, items of conspicuous consumption, and on the other hand, apparatus for putting in evidence the vicarious leisure rendered by the housewife.

The requirement of vicarious consumption at the hands of the wife continues in force even at a lower point in the pecuniary scale than the requirement of vicarious leisure. At a point below which little if any pretense of wasted effort, in ceremonial cleanness and the like, is observable, and where there is assuredly no conscious attempt at ostensible leisure, decency still requires the wife to

consume some goods conspicuously for the reputability of the household and its head. So that, as the latter-day outcome of this evolution of an archaic institution, the wife, who was at the outset the drudge and chattel of the man, both in fact and in theory—the producer of goods for him to consume—has become the ceremonial consumer of goods which he produces. But she still quite unmistakably remains his chattel in theory; for the habitual rendering of vicarious leisure and consumption is the abiding mark of the un-free servant.

This vicarious consumption practiced by the household of the middle and lower classes cannot be counted as a direct expression of the leisure-class scheme of life, since the household of this pecuniary grade does not belong within the leisure class. It is rather that the leisure-class scheme of life here comes to an expression at the second remove. The leisure class stands at the head of the social structure in point of reputability; and its manner of life and its standards of worth therefore afford the norm of reputability for the community. The observance of these standards, in some degree of approximation, becomes incumbent upon all classes lower in the scale. In modern civilized communities the lines of demarcation between social classes have grown vague and transient, and wherever this happens the norm of reputability imposed by the upper class extends its coercive influence with but slight hindrance down through the social structure to the lowest strata. The result is that the members of each stratum accept as their ideal of decency the scheme of life in vogue in the next higher stratum, and bend their energies to live up to that ideal. On pain of forfeiting their good name and their self-respect in case of failure, they must conform to the accepted code, at least in appearance.

The basis on which good repute in any highly organized industrial community ultimately rests is pecuniary strength; and the means of showing pecuniary strength, and so

of gaining or retaining a good name, are leisure and a conspicuous consumption of goods. Accordingly, both of these methods are in vogue as far down the scale as it remains possible; and in the lower strata in which the two methods are employed, both offices are in great part delegated to the wife and children of the household. Lower still, where any degree of leisure, even ostensible, has become impracticable for the wife, the conspicuous consumption of goods remains and is carried on by the wife and children. The man of the household also can do something in this direction, and indeed, he commonly does; but with a still lower descent into the levels of indigence—along the margin of the slums—the man, and presently also the children, virtually cease to consume valuable goods for appearances, and the woman remains virtually the sole exponent of the household's pecuniary decency. No class of society, not even the most abjectly poor, forgoes all customary conspicuous consumption. The last items of this category of consumption are not given up except under stress of the direst necessity. Very much of squalor and discomfort will be endured before the last trinket or the last pretense of pecuniary decency is put away. There is no class and no country that has yielded so abjectly before the pressure of physical want as to deny themselves all gratification of this higher or spiritual need.

From the foregoing survey of the growth of conspicuous leisure and consumption, it appears that the utility of both alike for the purposes of reputability lies in the element of waste that is common to both. In the one case it is a waste of time and effort, in the other it is a waste of goods. Both are methods of demonstrating the possession of wealth, and the two are conventionally accepted as equivalents. The choice between them is a question of advertising expediency simply, except so far as it may be affected by other standards of propriety, springing from a different source. On grounds of expediency the

preference may be given to the one or the other at different stages of the economic development. The question is, which of the two methods will most effectively reach the persons whose convictions it is desired to affect. Usage has answered this question in different ways under different circumstances.

So long as the community or social group is small enough and compact enough to be effectually reached by common notoriety alone—that is to say, so long as the human environment to which the individual is required to adapt himself in respect of reputability is comprised within his sphere of personal acquaintance and neighborhood gossip—so long the one method is about as effective as the other. Each will therefore serve about equally well during the earlier stages of social growth. But when the differentiation has gone farther and it becomes necessary to reach a wider human environment, consumption begins to hold over leisure as an ordinary means of decency. This is especially true during the later, peaceable economic stage. The means of communication and the mobility of the population now expose the individual to the observation of many persons who have no other means of judging of his reputability than the display of goods (and perhaps of breeding) which he is able to make while he is under their direct observation.

The modern organization of industry works in the same direction also by another line. The exigencies of the modern industrial system frequently place individuals and households in juxtaposition between whom there is little contact in any other sense than that of juxtaposition. One's neighbors, mechanically speaking, often are socially not one's neighbors, or even acquaintances; and still their transient good opinion has a high degree of utility. The only practicable means of impressing one's pecuniary ability on these unsympathetic observers of one's everyday life is an unremitting demonstration of ability to pay. In the modern community there is also a more frequent attendance at large

gatherings of people to whom one's everyday life is unknown; in such places as churches, theaters, ballrooms, hotels, parks, shops, and the like. In order to impress these transient observers, and to retain one's self-complacency under their observation, the signature of one's pecuniary strength should be written in characters which he who runs may read. It is evident, therefore, that the present trend of the development is in the direction of heightening the utility of conspicuous consumption as compared with leisure. . . .

Pecuniary Canons of Taste

The requirements of pecuniary decency have, to a very appreciable extent, influenced the sense of beauty and of utility in articles of use or beauty. Articles are to an extent preferred for use on account of their being conspicuously wasteful; they are felt to be serviceable somewhat in proportion as they are wasteful and ill adapted to their ostensible use. . . .

By habituation to an appreciative perception of the marks of expensiveness in goods, and by habitually identifying beauty with reputability, it comes about that a beautiful article which is not expensive is accounted not beautiful. In this way it has happened, for instance, that some beautiful flowers pass conventionally for offensive weeds; others that can be cultivated with relative ease are accepted and admired by the lower middle class, who can afford no more expensive luxuries of this kind; but these varieties are rejected as vulgar by those people who are better able to pay for expensive flowers and who are educated to a higher schedule of pecuniary beauty in the florist's products; while still other flowers, of no greater intrinsic beauty than these, are cultivated at great cost and call out much admiration from flower-lovers whose tastes have been matured under the critical guidance of a polite environment.

The same variation in matters of taste, from one class of society to another, is visible also as regards many other kinds of consumable

goods, as, for example, is the case with furniture, houses, parks, and gardens. This diversity of views as to what is beautiful in these various classes of goods is not a diversity of the norm according to which the unsophisticated sense of the beautiful works. It is not a constitutional difference of endowments in the æsthetic respect, but rather a difference in the code of reputability which specifies what objects properly lie within the scope of

honorific consumption for the class to which the critic belongs. It is a difference in the traditions of propriety with respect to the kinds of things which may, without derogation to the consumer, be consumed under the head of objects of taste and art. With a certain allowance for variations to be accounted for on other grounds, these traditions are determined, more or less rigidly, by the pecuniary plane of life of the class.

93 ■ Pierre Bourdieu

Distinction

A Social Critique of the Judgement of Taste

The Social Space

The distribution of the different classes (and class fractions) runs from those who are best provided with both economic and cultural capital to those who are most deprived in both respects (see Figures 1 and 2). The members of the professions, who have high incomes and high qualifications, who very often (52.9 percent) originate from the dominant class (professions or senior executives), who receive and consume a large quantity of both material and cultural goods, are opposed in almost all respects to the office workers, who have low qualifications, often originate from the working or middle classes, who receive little and consume little, devoting a high proportion of their time to care maintenance and home

improvement; and they are even more opposed to the skilled or semi-skilled workers, and still more to unskilled workers or farm labourers, who have the lowest incomes, no qualifications, and originate almost exclusively (90.5 percent of farm labourers, 84.5 percent of unskilled workers) from the working classes.¹

The differences stemming from the total volume of capital almost always conceal, both from common awareness and also from “scientific” knowledge, the secondary differences which, within each of the classes defined by overall volume of capital, separate class fractions, defined by different asset structures, i.e., different distributions of their total capital among the different kinds of capital [economic and cultural].

Pierre Bourdieu. *Distinction: A Social Critique of the Judgement of Taste*, pp. 114–115, 128–129, 169–180, 183–197, 199–202, 206–209, 503, 519–521, 523, and 572–576. Copyright © 1984 by the President and Fellows of Harvard University Press and Routledge and Kegan Paul, Ltd. Reprinted by permission of the publishers and of Routledge and Kegan Paul, Ltd. This chapter is based on survey data collected by Pierre Bourdieu and on additional complementary data sources (denoted in the text by C.S. I – C.S. V). For the purpose of conserving space, we have omitted all the special passages, set off in distinct type in the original, that were devoted to “illustrative examples or discussion of ancillary issues” (Bourdieu, *Distinction*, p. xiii).

Once one takes account of the structure of total assets—and not only, as has always been done implicitly, of the dominant kind in a given structure, “birth,” “fortune” or “talents,” as the nineteenth century put it—one has the means of making more precise divisions and also of observing the specific effects of the structure of distribution between the different kinds of capital. This may, for example, be symmetrical (as in the case of the professions, which combine very high income with very high cultural capital) or asymmetrical (in the case of higher-education and secondary teachers or employers, with cultural capital dominant in one case, economic capital in the other). One thus discovers two sets of homologous positions. The fractions whose reproduction depends on economic capital, usually inherited—industrial and commercial employers at the higher level, craftsmen and shopkeepers at the intermediate level—are opposed to the fractions which are least endowed (relatively, of course) with economic capital, and whose reproduction mainly depends on cultural capital—higher-education and secondary teachers at the higher level, primary teachers at the intermediate level. . . .

The Habitus

The mere fact that the social space described here can be presented as a diagram indicates that it is an abstract representation, deliberately constructed, like a map, to give a bird’s-eye view, a point of view on the whole set of points from which ordinary agents (including the sociologist and his reader, in their ordinary behaviour) see the social world. Bringing together in simultaneity, in the scope of a single glance—this is its heuristic value—positions which the agents can never apprehend in their totality and in their multiple relationships, social space is to the practical space of everyday life, with its distances which are kept or signalled, and neighbours who may be more remote than strangers, what geometrical space is to the “travelling

space” (*espace hodologique*) of ordinary experience, with its gaps and discontinuities.

But the most crucial thing to note is that the question of this space is raised within the space itself—that the agents have points of view on this objective space which depend on their position within it and in which their will to transform or conserve it is often expressed. Thus many of the words which sociology uses to designate the classes it constructs are borrowed from ordinary usage, where they serve to express the (generally polemical) view that one group has of another. As if carried away by their quest for greater objectivity, sociologists almost always forget that the “objects” they classify produce not only objectively classifiable practices but also classifying operations that are no less objective and are themselves classifiable. The division into classes performed by sociology leads to the common root of the classifiable practices which agents produce and of the classificatory judgements they make of other agents’ practices and their own. The habitus is both the generative principle of objectively classifiable judgements and the system of classification (*principium divisionis*) of these practices. It is in the relationship between the two capacities which define the habitus, the capacity to produce classifiable practices and works, and the capacity to differentiate and appreciate these practices and products (taste), that the represented social world, i.e., the space of lifestyles, is constituted.

The relationship that is actually established between the pertinent characteristics of economic and social condition (capital volume and composition, in both synchronic and diachronic aspects) and the distinctive features associated with the corresponding position in the universe of life-styles only becomes intelligible when the habitus is constructed as the generative formula which makes it possible to account both for the classifiable practices and products and for the judgements, themselves classified, which make these practices and works into a system of distinctive signs. When one speaks of the aristocratic asceticism of

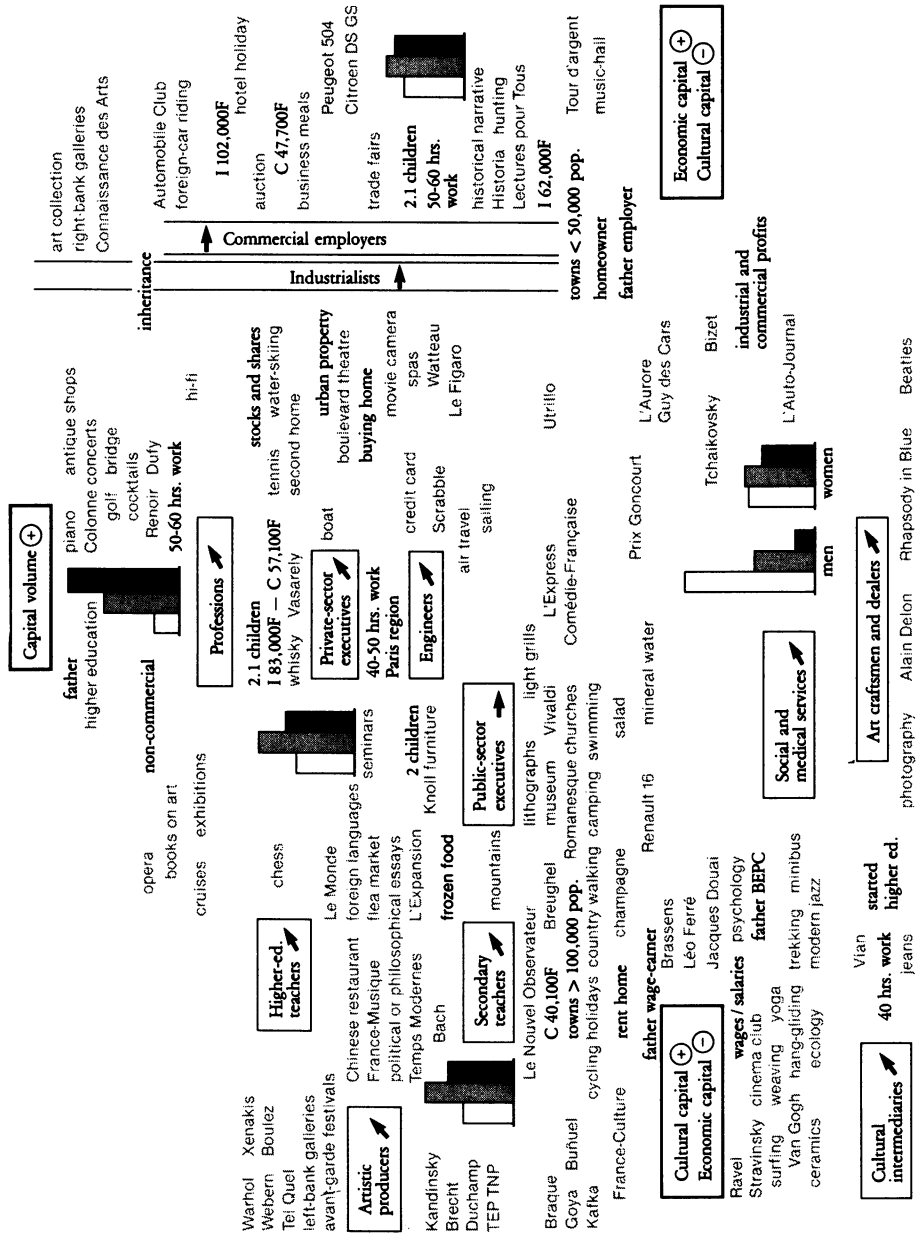


Figure 1. The Space of Social Positions (shown in black)

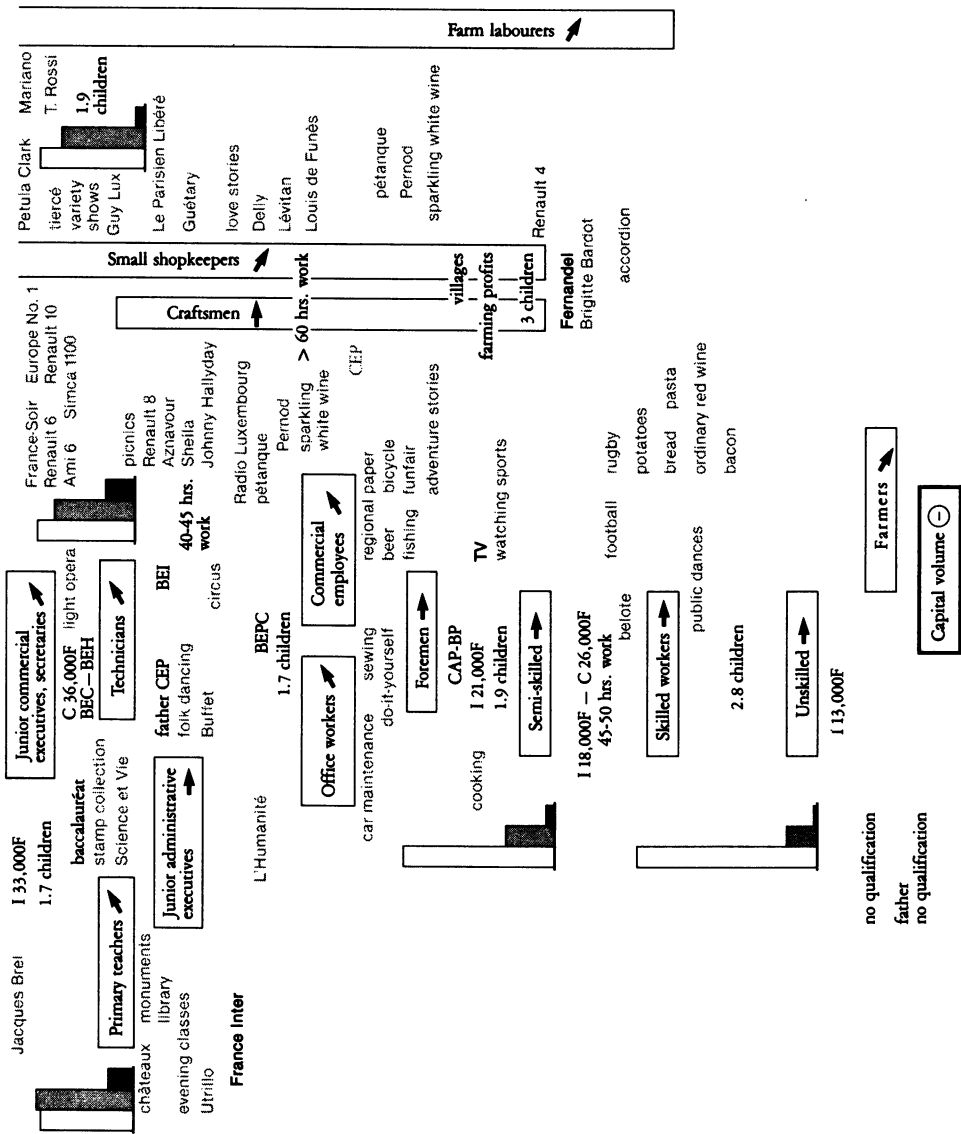


Figure 2. The Space of Life-Styles (shown in grey)

teachers or the pretension of the petite bourgeoisie, one is not only describing these groups by one, or even the most important, of their properties, but also endeavouring to name the principle which generates all their properties and all their judgements of their, or other people's, properties. The habitus is necessity internalized and converted into a disposition that generates meaningful practices and meaning-giving perceptions; it is a general, transposable disposition which carries out a systematic, universal application—beyond the limits of what has been directly learnt—of the necessity inherent in the learning conditions. That is why an agent's whole set of practices (or those of a whole set of agents produced by similar conditions) are both systematic, inasmuch as they are the product of the application of identical (or interchangeable) schemes, and systematically distinct from the practices constituting another life-style.

Because different conditions of existence produce different habitus—systems of generative schemes applicable, by simple transfer, to the most varied areas of practice—the practices engendered by the different habitus appear as systematic configurations of properties expressing the differences objectively inscribed in conditions of existence in the form of systems of differential deviations which, when perceived by agents endowed with the schemes of perception and appreciation necessary in order to identify, interpret and evaluate their pertinent features, function as life-styles (see Figure 3).²

The habitus is not only a structuring structure, which organizes practices and the perception of practices, but also a structured structure: the principle of division into logical classes which organizes the perception of the social world is itself the product of internalization of the division into social classes. Each class condition is defined, simultaneously, by its intrinsic properties and by the relational properties which it derives from its position in the system of class conditions, which is also a system of differences, differential positions,

i.e., by everything which distinguishes it from what it is not and especially from everything it is opposed to; social identity is defined and asserted through difference. This means that inevitably inscribed within the dispositions of the habitus is the whole structure of the system of conditions, as it presents itself in the experience of a life-condition occupying a particular position within that structure. The most fundamental oppositions in the structure (high/low, rich/poor etc.) tend to establish themselves as the fundamental structuring principles of practices and the perception of practices. As a system of practice-generating schemes which expresses systematically the necessity and freedom inherent in its class condition and the difference constituting that position, the habitus apprehends differences between conditions, which it grasps in the form of differences between classified, classifying practices (products of other habitus), in accordance with principles of differentiation which, being themselves the product of these differences, are objectively attuned to them and therefore tend to perceive them as natural.

While it must be reasserted, against all forms of mechanism, that ordinary experience of the social world is a cognition, it is equally important to realize—contrary to the illusion of the spontaneous generation of consciousness which so many theories of the “awakening of class consciousness:” (*prise de conscience*) amount to—that primary cognition is misrecognition, recognition of an order which is also established in the mind. Life-styles are thus the systematic products of habitus, which, perceived in their mutual relations through the schemes of the habitus, become sign systems that are socially qualified (as “distinguished,” “vulgar” etc.). The dialectic of conditions and habitus is the basis of an alchemy which transforms the distribution of capital, the balance-sheet of a power relation, into a system of perceived differences, distinctive properties, that is, a distribution of symbolic capital, legitimate capital, whose objective truth is misrecognized.

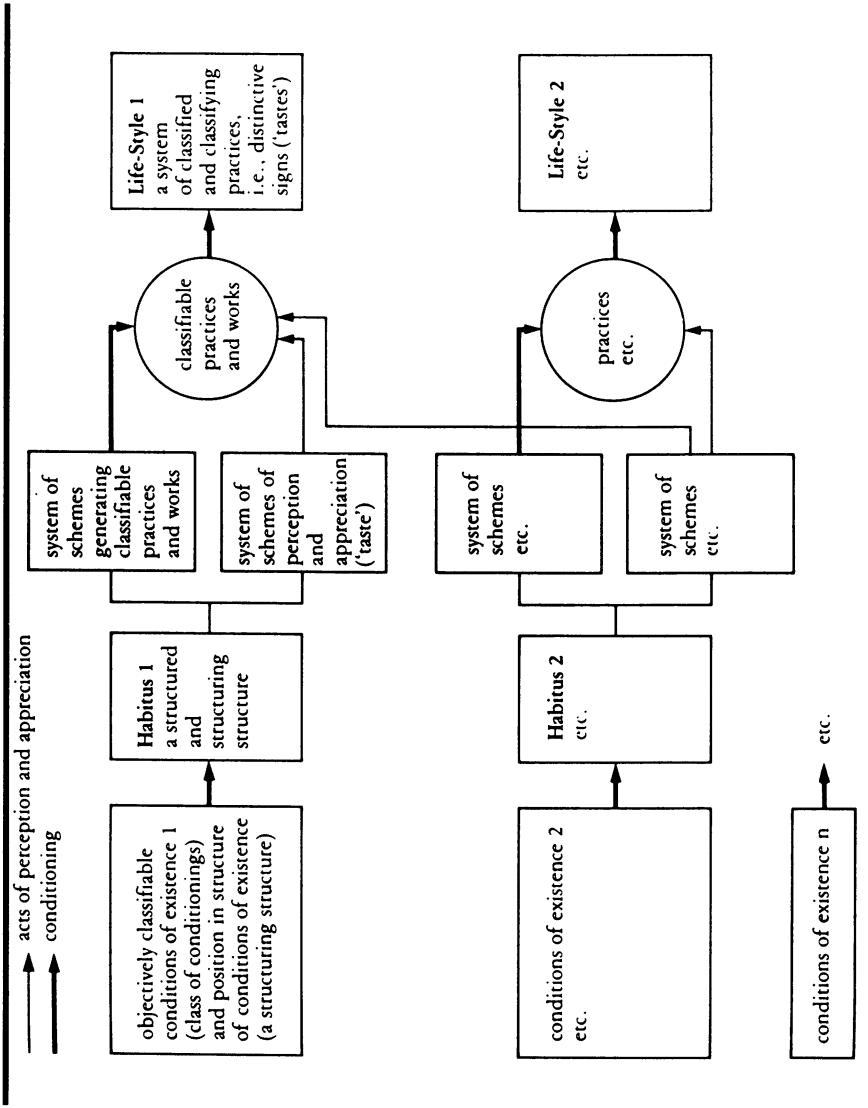


Figure 3. Conditions of Existence, Habitus, and Life-Style

As structured products (*opus operatum*) which a structuring structure (*modus operandi*) produces through retranslations according to the specific logic of the different *fields*, all the practices and products of a given agent are objectively harmonized among themselves, without any deliberate pursuit of coherence, and objectively orchestrated, without any conscious concertation, with those of all members of the same class. The habitus continuously generates practical metaphors, that is to say, transfers (of which the transfer of motor habits is only one example) or, more precisely, systematic transpositions required by the particular conditions in which the habitus is “put into practice” (so that, for example, the ascetic ethos which might be expected always to express itself in saving may, in a given context, express itself in a particular way of using credit). The practices of the same agent, and, more generally, the practices of all agents of the same class, owe the stylistic affinity which makes each of them a metaphor of any of the others to the fact that they are the product of transfers of the same schemes of action from one field to another. An obvious paradigm would be the disposition called “handwriting”, a singular way of tracing letters which always produces the same writing, i.e., graphic forms which, in spite of all the differences of size, material or colour due to the surface (paper or blackboard) or the instrument (pen or chalk)—in spite, therefore, of the different use of muscles—present an immediately perceptible family resemblance, like all the features of style or manner whereby a painter or writer can be recognized as infallibly as a man by his walk.

Systematicity is found in the *opus operatum* because it is in the *modus operandi*.³ It is found in all the properties—and property—with which individuals and groups surround themselves, houses, furniture, paintings, books, cars, spirits, cigarettes, perfume, clothes, and in the practices in which they manifest their distinction, sports, games, entertainments, only because it is in the syn-

thetic unity of the habitus, the unifying, generative principle of all practices. Taste, the propensity and capacity to appropriate (materially or symbolically) a given class of classified, classifying objects or practices, is the generative formula of life-style, a unitary set of distinctive preferences which express the same expressive intention in the specific logic of each of the symbolic subspaces, furniture, clothing, language or body hexis. Each dimension of life-style “symbolizes with” the others, in Leibniz’s phrase, and symbolizes them. An old cabinetmaker’s world view, the way he manages his budget, his time or his body, his use of language and choice of clothing are fully present in his ethic of scrupulous, impeccable craftsmanship and in the aesthetic of work for work’s sake which leads him to measure the beauty of his products by the care and patience that have gone into them.

The system of matching properties, which includes people—one speaks of a “well-matched couple”, and friends like to say they have the same tastes—is organized by taste, a system of classificatory schemes which may only very partially become conscious although, as one rises in the social hierarchy, life-style is increasingly a matter of what Weber calls the “stylization of life”. Taste is the basis of the mutual adjustment of all the features associated with a person, which the old aesthetic recommended for the sake of the mutual reinforcement they give one another; the countless pieces of information a person consciously or unconsciously imparts endlessly underline and confirm one another, offering the alert observer the same pleasure an art-lover derives from the symmetries and correspondences produced by a harmonious distribution of redundancies. The over-determination that results from these redundancies is felt the more strongly because the different features which have to be isolated for observation or measurement strongly interpenetrate in ordinary perception; each item of information imparted in practice (e.g., a judgement of a painting) is contaminated—and, if it deviates from the

probable feature, corrected—by the effect of the whole set of features previously or simultaneously perceived. That is why a survey which tends to isolate features—for example, by dissociating the things said from the way they are said—and detach them from the system of correlative features tends to minimize the deviation, on each point, between the classes, especially that between the *petit bourgeois* and the *bourgeois*. In the ordinary situations of *bourgeois* life, banalities about art, literature or cinema are inseparable from the steady tone, the slow, casual diction, the distant or self-assured smile, the measured gesture, the well-tailored suit and the *bourgeois* salon of the person who pronounces them.

Taste is the practical operator of the transmutation of things into distinct and distinctive signs, of continuous distributions into discontinuous oppositions; it raises the differences inscribed in the physical order of bodies to the symbolic order of significant distinctions. It transforms objectively classified practices, in which a class condition signifies itself (through taste), into classifying practices, that is, into a symbolic expression of class position, by perceiving them in their mutual relations and in terms of social classificatory schemes. Taste is thus the source of the system of distinctive features which cannot fail to be perceived as a systematic expression of a particular class of conditions of existence, i.e., as a distinctive life-style, by anyone who possesses practical knowledge of the relationships between distinctive signs and positions in the distributions—between the universe of objective properties, which is brought to light by scientific construction, and the no less objective universe of life-styles, which exists as such for and through ordinary experience.

This classificatory system, which is the product of the internalization of the structure of social space, in the form in which it impinges through the experience of a particular position in that space, is, within the limits of economic possibilities and impossibilities (which it tends to reproduce in its own logic),

the generator of practices adjusted to the regularities inherent in a condition. It continuously transforms necessities into strategies, constraints into preferences, and, without any mechanical determination, it generates the set of “choices” constituting life-styles, which derive their meaning, i.e., their value, from their position in a system of oppositions and correlations.⁴ It is a virtue made of necessity which continuously transforms necessity into virtue by inducing “choices” which correspond to the condition of which it is the product. As can be seen whenever a change in social position puts the *habitus* into new conditions, so that its specific efficacy can be isolated, it is taste—the taste of necessity or the taste of luxury—and not high or low income which commands the practices objectively adjusted to these resources. Through taste, an agent has what he likes because he likes what he has, that is, the properties actually given to him in the distributions and legitimately assigned to him in the classifications.⁵

The Homology Between the Spaces

Bearing in mind all that precedes, in particular the fact that the generative schemes of the *habitus* are applied, by simple transfer, to the most dissimilar areas of practice, one can immediately understand that the practices or goods associated with the different classes in the different areas of practice are organized in accordance with structures of opposition which are homologous to one another because they are all homologous to the structure of objective oppositions between class conditions. Without presuming to demonstrate here in a few pages what the whole of the rest of this work will endeavour to establish—but lest the reader fail to see the wood for the trees of detailed analysis—I shall merely indicate, very schematically, how the two major organizing principles of the social space govern the structure and modification of the space of cultural consumption, and, more generally, the whole universe of life-styles.

In cultural consumption, the main opposition, by overall capital value, is between the practices designated by their rarity as distinguished, those of the fractions richest in both economic and cultural capital, and the practices socially identified as vulgar because they are both easy and common, those of the fractions poorest in both these respects. In the intermediate position are the practices which are perceived as pretentious, because of the manifest discrepancy between ambition and possibilities. In opposition to the dominated condition, characterized, from the point of view of the dominant, by the combination of forced poverty and unjustified laxity, the dominant aesthetic—of which the work of art and the aesthetic disposition are the most complete embodiments—proposes the combination of ease and asceticism, i.e., self-imposed austerity, restraint, reserve, which are affirmed in that absolute manifestation of excellence, relaxation in tension.

This fundamental opposition is specified according to capital composition. Through the mediation of the means of appropriation available to them, exclusively or principally cultural on the one hand, mainly economic on the other, and the different forms of relation to works of art which result from them, the different fractions of the dominant class are oriented towards cultural practices so different in their style and object and sometimes so antagonistic (those of “artists” and “bourgeois”)⁶ that it is easy to forget that they are variants of the same fundamental relationship to necessity and to those who remain subject to it, and that each pursues the exclusive appropriation of legitimate cultural goods and the associated symbolic profits. Whereas the dominant fractions of the dominant class (the “bourgeoisie”) demand of art a high degree of denial of the social world and incline towards a hedonistic aesthetic of ease and facility, symbolized by boulevard theatre or Impressionist painting, the dominated fractions (the “intellectuals” and “artists”) have affinities with the ascetic as-

pect of aesthetics and are inclined to support all artistic revolutions conducted in the name of purity and purification, refusal of ostentation and the bourgeois taste for ornament; and the dispositions towards the social world which they owe to their status as poor relations incline them to welcome a pessimistic representation of the social world.

While it is clear that art offers it the greatest scope, there is no area of practice in which the intention of purifying, refining and sublimating facile impulses and primary needs cannot assert itself, or in which the stylization of life, i.e., the primacy of form over function, which leads to the denial of function, does not produce the same effects. In language, it gives the opposition between popular outspokenness and the highly censored language of the bourgeois, between the expressionist pursuit of the picturesque or the rhetorical effect and the choice of restraint and false simplicity (*litotes*). The same economy of means is found in body language: here too, agitation and haste, grimaces and gesticulation are opposed to slowness—“the slow gestures, the slow glance” of nobility, according to Nietzsche⁷—to the restraint and impassivity which signify elevation. Even the field of primary tastes is organized according to the fundamental opposition, with the antithesis between quantity and quality, belly and palate, matter and manners, substance and form.

Form and Substance

The fact that in the realm of food the main opposition broadly corresponds to differences in income has masked the secondary opposition which exists, both within the middle classes and within the dominant class, between the fractions richer in cultural capital and less rich in economic capital and those whose assets are structured in the opposite way. Observers tend to see a simple effect of income in the fact that, as one rises in the social hierarchy, the proportion of income spent on food diminishes, or that, within the food budget, the proportion spent on heavy, fatty,

fattening foods, which are also cheap—pasta, potatoes, beans, bacon, pork—declines (C.S. III), as does that spent on wine, whereas an increasing proportion is spent on leaner, lighter (more digestible), non-fattening foods (beef, veal, mutton, lamb, and especially fresh fruit and vegetables).⁸ Because the real principle of preferences is taste, a virtue made of necessity, the theory which makes consumption a simple function of income has all the appearances to support it, since income plays an important part in determining distance from necessity. However, it cannot account for cases in which the same income is associated with totally different consumption patterns. Thus, foremen remain attached to “popular” taste although they earn more than clerical and commercial employees, whose taste differs radically from that of manual workers and is closer to that of teachers.

For a real explanation of the variations which J. F. Engel’s law merely records, one has to take account of all the characteristics of social condition which are (statistically) associated from earliest childhood with possession of high or low income and which tend to shape tastes adjusted to these conditions.⁹ The true basis of the differences found in the area of consumption, and far beyond it, is the opposition between the tastes of luxury (or freedom) and the tastes of necessity. The former are the tastes of individuals who are the product of material conditions of existence defined by distance from necessity, by the freedoms or facilities stemming from possession of capital; the latter express, precisely in their adjustment, the necessities of which they are the product. Thus it is possible to deduce popular tastes for the foods that are simultaneously most “filling” and most economical¹⁰ from the necessity of reproducing labour power at the lowest cost which is forced on the proletariat as its very definition. The idea of taste, typically bourgeois, since it presupposes absolute freedom of choice, is so closely associated with the idea of freedom that many people find it

hard to grasp the paradoxes of the taste of necessity. Some simply sweep it aside, making practice a direct product of economic necessity (workers eat beans because they cannot afford anything else), failing to realize that necessity can only be fulfilled, most of the time, because the agents are inclined to fulfil it, because they have a taste for what they are anyway condemned to. Others turn it into a taste of freedom, forgetting the conditionings of which it is the product, and so reduce it to pathological or morbid preference for (basic) essentials, a sort of congenital coarseness, the pretext for a class racism which associates the populace with everything heavy, thick and fat.¹¹ Taste is *amor fati*, the choice of destiny, but a forced choice, produced by conditions of existence which rule out all alternatives as mere daydreams and leave no choice but the taste for the necessary.

The taste of necessity can only be the basis of a life-style “in-itself”, which is defined as such only negatively, by an absence, by the relationship of privation between itself and the other life-styles. For some, there are elective emblems, for others stigmata which they bear in their very bodies. “As the chosen people bore in their features the sign that they were the property of Jehovah, so the division of labour brands the manufacturing worker as the property of capital”.¹² The brand which Marx speaks of is nothing other than life-style, through which the most deprived immediately betray themselves, even in their use of spare time; in so doing they inevitably serve as a foil to every distinction and contribute, purely negatively, to the dialectic of pretension and distinction which fuels the incessant changing of taste. Not content with lacking virtually all the knowledge or manners which are valued in the markets of academic examination or polite conversation nor with only possessing skills which have no value there, they are the people “who don’t know how to live”, who sacrifice most to material foods, and to the heaviest, grossest and most fattening of them, bread, potatoes, fats,

and the most vulgar, such as wine; who spend least on clothing and cosmetics, appearance and beauty; those who “don’t know how to relax”, “who always have to be doing something”, who set off in their Renault 5 or Simca 1000 to join the great traffic jams of the holiday exodus, who picnic beside major roads, cram their tents into overcrowded campsites, fling themselves into the prefabricated leisure activities designed for them by the engineers of cultural mass production; those who by all these uninspired “choices” confirm class racism, if it needed to be confirmed, in its conviction that they only get what they deserve.

The art of eating and drinking remains one of the few areas in which the working classes explicitly challenge the legitimate art of living. In the face of the new ethic of sobriety for the sake of slimness, which is most recognized at the highest levels of the social hierarchy, peasants and especially industrial workers maintain an ethic of convivial indulgence. A bon vivant is not just someone who enjoys eating and drinking; he is someone capable of entering into the generous and familiar—that is, both simple and free—relationship that is encouraged and symbolized by eating and drinking together, in a conviviality which sweeps away restraints and reticence.

The boundary marking the break with the popular relation to food runs, without any doubt, between the manual workers and the clerical and commercial employees (C.S. II). Clerical workers spend less on food than skilled manual workers, both in absolute terms (9,376 francs as against 10,347 francs) and in relative terms (34.2 percent as against 38.3 percent); they consume less bread, pork, pork products (*charcuterie*), milk, cheese, rabbit, poultry, dried vegetables and fats, and, within a smaller food budget, spend as much on meat—beef, veal, mutton and lamb—and slightly more on fish, fresh fruit and aperitifs. These changes in the structure of spending on food are accompanied by increased spending on health and beauty care and clothing, and a

slight increase in spending on cultural and leisure activities. When it is noted that the reduced spending on food, especially on the most earthly, earthy, down-to-earth foods, is accompanied by a lower birth-rate, it is reasonable to suppose that it constitutes one aspect of an overall transformation of the relationship to the world. The “modest” taste which can defer its gratifications is opposed to the spontaneous materialism of the working classes, who refuse to participate in the Benthamite calculation of pleasures and pains, benefits and costs (e.g., for health and beauty). In other words, these two relations to the “fruits of the earth” are grounded in two dispositions towards the future which are themselves related in circular causality to two objective futures. Against the imaginary anthropology of economics, which has never shrunk from formulating universal laws of “temporal preference”, it has to be pointed out that the propensity to subordinate present desires to future desires depends on the extent to which this sacrifice is “reasonable”, that is, on the likelihood, in any case, of obtaining future satisfactions superior to those sacrificed.¹³

Among the economic conditions of the propensity to sacrifice immediate satisfactions to expected satisfactions one must include the probability of these future satisfactions which is inscribed in the present condition. There is still a sort of economic calculation in the unwillingness to subject existence to economic calculation. The hedonism which seizes day by day the rare satisfactions (“good times”) of the immediate present is the only philosophy conceivable to those who “have no future” and, in any case, little to expect from the future.¹⁴ It becomes clearer why the practical materialism which is particularly manifested in the relation to food is one of the most fundamental components of the popular ethos and even the popular ethic. The being-in-the-present which is affirmed in the readiness to take advantage of the good times and take time as it comes is, in itself, an affirmation of solidarity with

Table 1. Yearly Spending by Teachers, Professionals, and Industrial and Commercial Employers, 1972

| Type of spending | Teachers (higher and secondary) | | Professionals | | Industrial and commercial employers | |
|---------------------------|------------------------------------|------------|---------------|------------|-------------------------------------|------------|
| | Francs | % of total | Francs | % of total | Francs | % of total |
| Food ^a | 9,969 | 24.4 | 13,956 | 24.4 | 16,578 | 37.4 |
| Presentation ^b | 4,912 | 12.0 | 12,680 | 22.2 | 5,616 | 12.7 |
| Culture ^c | 1,753 | 4.3 | 1,298 | 2.3 | 574 | 1.3 |

Source: C.S. II (1972).

a. Includes restaurant or canteen meals.

b. Clothes, shoes, repairs and cleaning, toiletries, hairdressing, domestic servants.

c. Books, newspapers and magazines, stationery, records, sport, toys, music, entertainments.

others (who are often the only present guarantee against the threats of the future), inasmuch as this temporal immanentism is a recognition of the limits which define the condition. This is why the sobriety of the petit bourgeois is felt as a break: in abstaining from having a good time and from having it with others, the would-be petit bourgeois betrays his ambition of escaping from the common present, when, that is, he does not construct his whole self-image around the opposition between his home and the café, abstinence and intemperance, in other words, between individual salvation and collective solidarities.

Three Styles of Distinction

The basic opposition between the tastes of luxury and the tastes of necessity is specified in as many oppositions as there are different ways of asserting one's distinction vis-à-vis the working class and its primary needs, or—which amounts to the same thing—different powers whereby necessity can be kept at a distance. Thus, within the dominant class, one can, for the sake of simplicity, distinguish three structures of the consumption distributed under three items: food, culture and presentation (clothing, beauty care, toiletries, domestic servants). These structures take strictly opposite forms—like the structures of their capital—among the teachers as against the industrial and commercial em-

ployers (see Table 1). Whereas the latter have exceptionally high expenditure on food (37 percent of the budget), low cultural costs and medium spending on presentation and representation, the former, whose total spending is lower on average, have low expenditure on food (relatively less than manual workers), limited expenditure on presentation (though their expenditure on health is one of the highest) and relatively high expenditure on culture (books, papers, entertainments, sport, toys, music, radio and record-player). Opposed to both these groups are the members of the professions, who devote the same proportion of their budget to food as the teachers (24.4 percent), but out of much greater total expenditure (57,122 francs as against 40,884 francs), and who spend much more on presentation and representation than all other fractions, especially if the costs of domestic service are included, whereas their cultural expenditure is lower than that of the teachers (or even the engineers and senior executives, who are situated between the teachers and the professionals, though nearer the latter, for almost all items).

The system of differences becomes clearer when one looks more closely at the patterns of spending on food. In this respect the industrial and commercial employers differ markedly from the professionals, and a fortiori from the teachers, by virtue of the importance they give to cereal-based products

(especially cakes and pastries), wine, meat preserves (foie gras, etc.) and game, and their relatively low spending on meat, fresh fruit and vegetables. The teachers, whose food purchases are almost identically structured to those of office workers, spend more than all other fractions on bread, milk products, sugar, fruit preserves and non-alcoholic drinks, less on wine and spirits and distinctly less than the professions on expensive products such as meat—especially the most expensive meats, such as mutton and lamb—and fresh fruit and vegetables. The members of the professions are mainly distinguished by the high proportion of their spending which goes on expensive products, particularly meat (18.3 percent of their food budget), and especially the most expensive meat (veal, lamb, mutton), fresh fruit and vegetables, fish and shellfish, cheese and aperitifs.¹⁵

Thus, when one moves from the manual workers to the industrial and commercial employers, through foremen, craftsmen and small shopkeepers, economic constraints tend to relax without any fundamental change in the pattern of spending (see Table 2). The opposition between the two extremes is here established between the poor and the rich (*nouveau riche*), between *la bouffe* and *la grande bouffe*;¹⁶ the food consumed is increasingly rich (both in cost and in calories) and increasingly heavy (game, foie gras). By contrast, the taste of the professionals or senior executives defines the popular taste, by negation, as the taste for the heavy, the fat and the coarse, by tending towards the light, the refined and the delicate (see Figure 4). The disappearance of economic constraints is accompanied by a strengthening of the social censorships which forbid coarseness and fatness, in favour of slimness and distinction. The taste for rare, aristocratic foods points to a traditional cuisine, rich in expensive or rare products (fresh vegetables, meat). Finally, the teachers, richer in cultural capital than in economic capital, and therefore inclined to ascetic consumption in all areas, pursue originality at the lowest

economic cost and go in for exoticism (Italian, Chinese cooking etc.)¹⁷ and culinary populism (peasant dishes). They are thus almost consciously opposed to the (new) rich with their rich food, the buyers and sellers of *grosse bouffe*, the “fat cats,”¹⁸ gross in body and mind, who have the economic means to flaunt, with an arrogance perceived as “vulgar”, a life-style which remains very close to that of the working classes as regards economic and cultural consumption.

Unpretentious or Uncouth?

It is clear that tastes in food cannot be considered in complete independence of the other dimensions of the relationship to the world, to others and to one's own body, through which the practical philosophy of each class is enacted. To demonstrate this, one would have to make a systematic comparison of the working-class and bourgeois ways of treating food, of serving, presenting and offering it, which are infinitely more revelatory than even the nature of the products involved (especially since most surveys of consumption ignore differences in quality). The analysis is a difficult one, because each life-style can only really be constructed in relation to the other, which is its objective and subjective negation, so that the meaning of behaviour is totally reversed depending on which point of view is adopted and on whether the common words which have to be used to name the conduct (e.g., “manners”) are invested with popular or bourgeois connotations.

Plain speaking, plain eating: the working-class meal is characterized by plenty (which does not exclude restrictions and limits) and above all by freedom. “Elastic” and “abundant” dishes are brought to the table—soups or sauces, pasta or potatoes (almost always included among the vegetables)—and served with a ladle or spoon, to avoid too much measuring and counting, in contrast to everything that has to be cut and divided, such as roasts.¹⁹ This impression of abundance, which is the norm on special occasions, and always

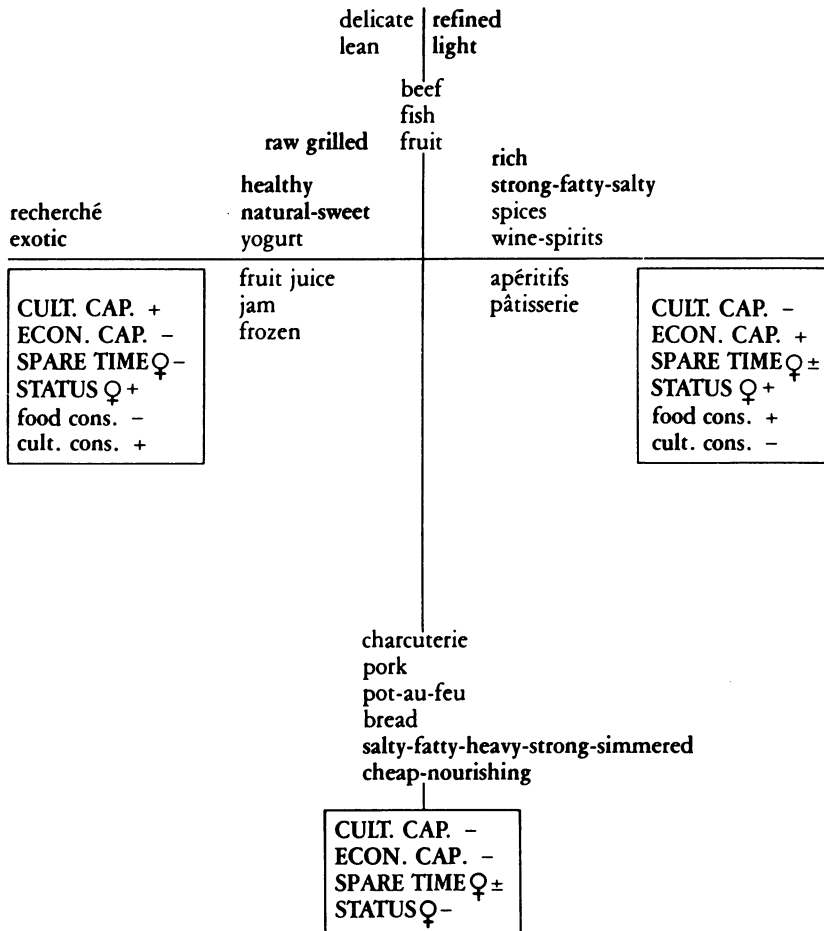


Figure 4. The Food Space

applies, so far as is possible, for the men, whose plates are filled twice (a privilege which marks a boy's accession to manhood), is often balanced, on ordinary occasions, by restrictions which generally apply to the women, who will share one portion between two, or eat the leftovers of the previous day; a girl's accession to womanhood is marked by doing without. It is part of men's status to eat and to eat well (and also to drink well); it is particularly insisted that they should eat, on the grounds that "it won't keep", and there is something suspect about a refusal. On Sundays, while the women are on their feet, busily serving, clearing the table, washing up, the men remain seated, still eating and drinking.

These strongly marked differences of social status (associated with sex and age) are accompanied by no practical differentiation (such as the bourgeois division between the dining room and the kitchen, where the servants eat and sometimes the children), and strict sequencing of the meal tends to be ignored. Everything may be put on the table at much the same time (which also saves walking), so that the women may have reached the dessert, and also the children, who will take their plates and watch television, while the men are still eating the main dish and the "lad", who has arrived late, is swallowing his soup.

This freedom, which may be perceived as disorder or slovenliness, is adapted to its

Table 2. Annual Household Expenditures on Food: Fractions of the Dominant Class, 1972

| Type of Food | Teachers (higher and secondary) | | Senior executives | | Professions | | Engineers | | Industrial and commer- cial employers | |
|---|---------------------------------------|-----------------------------|----------------------|-----------------------------|--------------|-----------------------------|--------------|-----------------------------|---|-----------------------------|
| | Average exp. | | Average exp. | | Average exp. | | Average exp. | | Average exp. | |
| | Frans | As % of all food exp. | Frans | As % of all food exp. | Frans | As % of all food exp. | Frans | As % of all food exp. | Frans | As % of all food exp. |
| Average number persons per household | 3.11 | | 3.6 | | 3.5 | | 3.6 | | 3.6 | |
| Average total household expenditure (frans) | 40,844 | | 52,156 | | 57,122 | | 49,822 | | 44,359 | |
| Average total household expenditure on food (frans) | 9,969 | | 13,158 | | 13,956 | | 12,666 | | 16,578 | |
| Expenditure on food as % of total expenditure | 24.4 | | 25.2 | | 24.4 | | 25.4 | | 37.4 | |
| | Average exp. | | Average exp. | | Average exp. | | Average exp. | | Average exp. | |
| | Frans | As % of all food exp. | Frans | As % of all food exp. | Frans | As % of all food exp. | Frans | As % of all food exp. | Frans | As % of all food exp. |
| Cereals | 865 | 8.7 | 993 | 7.5 | 1,011 | 7.2 | 951 | 7.2 | 1,535 | 9.2 |
| bread | 322 | 3.2 | 347 | 2.6 | 326 | 2.3 | 312 | 2.5 | 454 | 2.5 |
| cakes, pastries | 452 | 4.5 | 552 | 4.1 | 548 | 4.0 | 539 | 4.2 | 989 | 5.6 |
| rusk | 16 | 0.2 | 27 | 0.2 | 33 | 0.2 | 28 | 0.2 | 29 | 0.1 |
| rice | 35 | 0.3 | 32 | 0.2 | 62 | 0.4 | 41 | 0.3 | 33 | 0.1 |
| flour | 40 | 0.4 | 35 | 0.2 | 41 | 0.3 | 31 | 0.2 | 28 | 0.1 |
| Vegetables | 766 | 7.7 | 1,015 | 7.7 | 1,100 | 7.9 | 899 | 7.1 | 1,222 | 7.4 |
| potatoes | 81 | 0.8 | 94 | 0.7 | 95 | 0.7 | 98 | 0.7 | 152 | 0.8 |
| fresh vegetables | 555 | 5.6 | 729 | 5.5 | 811 | 5.8 | 647 | 5.1 | 915 | 5.1 |
| dried or canned | 131 | 1.3 | 191 | 1.4 | 216 | 1.5 | 154 | 1.2 | 153 | 0.8 |
| Fruit | 632 | 6.3 | 871 | 6.6 | 990 | 7.2 | 864 | 6.8 | 877 | 5.2 |
| fresh fruit | 295 | 2.9 | 405 | 3.1 | 586 | 4.2 | 424 | 3.3 | 547 | 3.1 |
| citrus fruit, bananas | 236 | 2.4 | 343 | 2.6 | 303 | 2.2 | 324 | 2.5 | 256 | 1.4 |
| dried | 102 | 1.0 | 122 | 0.9 | 98 | 0.7 | 116 | 0.9 | 72 | 0.4 |
| Butcher's meat | 1,556 | 15.6 | 2,358 | 18.0 | 2,552 | 18.3 | 2,073 | 16.4 | 2,323 | 14.0 |
| beef | 814 | 8.1 | 1,291 | 9.8 | 1,212 | 8.7 | 1,144 | 9.0 | 1,273 | 7.2 |
| veal | 335 | 3.4 | 452 | 3.4 | 630 | 4.5 | 402 | 3.1 | 377 | 2.3 |
| mutton, lamb | 156 | 1.6 | 315 | 2.3 | 438 | 3.2 | 242 | 1.9 | 390 | 2.2 |
| horse | 31 | 0.3 | 49 | 0.3 | 31 | 0.2 | 37 | 0.3 | 94 | 0.5 |
| pork (fresh) | 221 | 2.2 | 251 | 1.7 | 239 | 1.7 | 247 | 1.9 | 187 | 1.3 |

| | | | | | | | | | | |
|-----------------------------|-----|-----|-------|------|-------|------|-------|------|-------|------|
| Pork products | 634 | 6.3 | 741 | 5.6 | 774 | 5.5 | 705 | 5.6 | 812 | 4.9 |
| Meat preserves | 336 | 3.4 | 350 | 2.6 | 233 | 1.7 | 310 | 2.4 | 1,362 | 8.0 |
| Fish, shellfish | 336 | 3.4 | 503 | 3.8 | 719 | 5.1 | 396 | 3.1 | 588 | 3.5 |
| Poultry | 235 | 2.3 | 311 | 2.4 | 399 | 2.8 | 310 | 2.4 | 333 | 2.0 |
| Rabbit, game | 36 | 0.3 | 97 | 0.7 | 148 | 1.1 | 89 | 0.7 | 289 | 1.7 |
| Eggs | 149 | 1.4 | 172 | 1.3 | 190 | 1.4 | 178 | 1.4 | 185 | 1.1 |
| Milk | 299 | 3.0 | 271 | 2.0 | 249 | 1.8 | 287 | 2.3 | 309 | 1.9 |
| Cheese, yogurt | 692 | 6.9 | 776 | 5.9 | 843 | 6.0 | 785 | 6.1 | 1,090 | 6.5 |
| Fats | 399 | 4.0 | 564 | 4.3 | 525 | 3.8 | 504 | 4.0 | 551 | 3.3 |
| butter | 320 | 3.2 | 408 | 3.1 | 379 | 2.7 | 371 | 2.9 | 405 | 2.4 |
| oil | 66 | 0.6 | 136 | 1.0 | 132 | 1.0 | 103 | 0.8 | 112 | 0.6 |
| margarine | 12 | 0.1 | 17 | 0.1 | 12 | 0.1 | 29 | 0.2 | 19 | 0.1 |
| lard | 1 | 0 | 2 | 0 | 1 | 0 | 1 | 0 | 13 | 0.1 |
| Sugar, confectionery, cocoa | 304 | 3.0 | 395 | 3.0 | 265 | 1.9 | 327 | 2.6 | 407 | 2.4 |
| Alcohol | 711 | 7.1 | 1,365 | 10.3 | 1,329 | 9.5 | 937 | 7.4 | 2,218 | 13.4 |
| wine | 457 | 4.6 | 869 | 6.6 | 899 | 6.4 | 392 | 3.1 | 1,881 | 11.8 |
| beer | 82 | 0.8 | 91 | 0.7 | 40 | 0.3 | 184 | 1.4 | 93 | 0.5 |
| cider | 13 | 0.1 | 12 | 0 | 0 | 0 | 8 | 0 | 5 | 0 |
| apéritifs, liqueurs etc. | 157 | 1.6 | 391 | 3.0 | 389 | 2.8 | 352 | 2.8 | 237 | 1.4 |
| Non-alcoholic drinks | 344 | 3.4 | 342 | 2.6 | 267 | 1.9 | 295 | 2.3 | 327 | 2.0 |
| Coffee, tea | 152 | 1.5 | 215 | 1.5 | 291 | 2.1 | 178 | 1.4 | 298 | 1.8 |
| Restaurant meals | 829 | 8.3 | 1,863 | 13.0 | 1,562 | 11.2 | 1,372 | 10.8 | 1,179 | 7.1 |
| Canteen meals | 745 | 7.5 | 562 | 4.0 | 221 | 1.6 | 773 | 6.1 | 299 | 1.8 |
| Miscellaneous | 264 | 2.6 | 379 | 2.7 | 258 | 1.8 | 432 | 3.4 | 324 | 1.9 |

Source: C.S. II (1972).

function. Firstly, it is labour-saving, which is seen as an advantage. Because men take no part in housework, not least because the women would not allow it—it would be a dishonour to see men step outside their rôle—every economy of effort is welcome. Thus, when the coffee is served, a single spoon may be passed around to stir it. But these short cuts are only permissible because one is and feels at home, among the family, where ceremony would be an affectation. For example, to save washing up, the dessert may be handed out on improvised plates torn from the cake-box (with a joke about “taking the liberty”, to mark the transgression), and the neighbour invited in for a meal will also receive his piece of cardboard (offering a plate would exclude him) as a sign of familiarity. Similarly, the plates are not changed between dishes. The soup plate, wiped with bread, can be used right through the meal. The hostess will certainly offer to “change the plates”, pushing back her chair with one hand and reaching with the other for the plate next to her, but everyone will protest (“It all gets mixed up inside you”) and if she were to insist it would look as if she wanted to show off her crockery (which she is allowed to if it is a new present) or to treat her guests as strangers, as is sometimes deliberately done to intruders or “scroungers” who never return the invitation. These unwanted guests may be frozen out by changing their plates despite their protests, not laughing at their jokes, or scolding the children for their behaviour (“No, no, *we* don’t mind”, say the guests; “They ought to know better by now”, the parents respond). The common root of all these “liberties” is no doubt the sense that at least there will not be self-imposed controls, constraints and restrictions—especially not in eating, a primary need and a compensation—and especially not in the heart of domestic life, the one realm of freedom, when everywhere else, and at all other times, necessity prevails.

In opposition to the free-and-easy working-class meal, the bourgeoisie is concerned to eat

with all due form. Form is first of all a matter of rhythm, which implies expectations, pauses, restraints; waiting until the last person served has started to eat, taking modest helpings, not appearing over-eager. A strict sequence is observed and all coexistence of dishes which the sequence separates, fish and meat, cheese and dessert, is excluded: for example, before the dessert is served, everything left on the table, even the salt-cellar, is removed, and the crumbs are swept up. This extension of rigorous rules into everyday life (the bourgeois male shaves and dresses first thing every morning, and not just to “go out”), refusing the division between home and the exterior, the quotidian and the extraquotidian, is not explained solely by the presence of strangers—servants and guests—in the familiar family world. It is the expression of a habitus of order, restraint and propriety which may not be abdicated. The relation to food—the primary need and pleasure—is only one dimension of the bourgeois relation to the social world. The opposition between the immediate and the deferred, the easy and the difficult, substance (or function) and form, which is exposed in a particularly striking fashion in bourgeois ways of eating, is the basis of all aestheticization of practice and every aesthetic. Through all the forms and formalisms imposed on the immediate appetite, what is demanded—and inculcated—is not only a disposition to discipline food consumption by a conventional structuring which is also a gentle, indirect, invisible censorship (quite different from enforced privations) and which is an element in an art of living (correct eating, for example, is a way of paying homage to one’s hosts and to the mistress of the house, a tribute to her care and effort). It is also a whole relationship to animal nature, to primary needs and the populace who indulge them without restraint; it is a way of denying the meaning and primary function of consumption, which are essentially common, by making the meal a social ceremony, an affirmation of ethical tone and

aesthetic refinement. The manner of presenting and consuming the food, the organization of the meal and setting of the places, strictly differentiated according to the sequence of dishes and arranged to please the eye, the presentation of the dishes, considered as much in terms of shape and colour (like works of art) as of their consumable substance, the etiquette governing posture and gesture, ways of serving oneself and others, of using the different utensils, the seating plan, strictly but discreetly hierarchical, the censorship of all bodily manifestations of the act or pleasure of eating (such as noise or haste), the very refinement of the things consumed, with quality more important than quantity—this whole commitment to stylization tends to shift the emphasis from substance and function to form and manner, and so to deny the crudely material reality of the act of eating and of the things consumed, or, which amounts to the same thing, the basely material vulgarity of those who indulge in the immediate satisfactions of food and drink.²⁰

Given the basic opposition between form and substance, one could re-generate each of the oppositions between the two antagonistic approaches to the treatment of food and the act of eating. In one case, food is claimed as a material reality, a nourishing substance which sustains the body and gives strength (hence the emphasis on heavy, fatty, strong foods, of which the paradigm is pork—fatty and salty—the antithesis of fish—light, lean and bland); in the other, the priority given to form (the shape of the body, for example) and social form, formality, puts the pursuit of strength and substance in the background and identifies true freedom with the elective asceticism of a self-imposed rule. And it could be shown that two antagonistic world views, two worlds, two representations of human excellence are contained in this matrix. Substance—or matter—is what is substantial, not only “filling” but also real, as opposed to all appearances, all the fine words and empty gestures that “butter no parsnips”

and are, as the phrase goes, purely symbolic; reality, as against sham, imitation, window-dressing; the little eating-house with its marble-topped tables and paper napkins where you get an honest square meal and aren’t “paying for the wallpaper” as in fancy restaurants; being, as against seeming, nature and the natural, simplicity (pot-luck, “take it as it comes”, “no standing on ceremony”), as against embarrassment, mincing and posturing, airs and graces, which are always suspected of being a substitute for substance, i.e., for sincerity, for feeling, for what is felt and proved in actions; it is the free-speech and language of the heart which make the true “nice guy”, blunt, straightforward, unbending, honest, genuine, “straight down the line” and “straight as a die”, as opposed to everything that is pure form, done only for form’s sake; it is freedom and the refusal of complications, as opposed to respect for all the forms and formalities spontaneously perceived as instruments of distinction and power. On these moralities, these world views, there is no neutral viewpoint; what for some is shameless and slovenly, for others is straightforward, unpretentious; familiarity is for some the most absolute form of recognition, the abdication of all distance, a trusting openness, a relation of equal to equal; for others, who shun familiarity, it is an unseemly liberty.

The popular realism which inclines working people to reduce practices to the reality of their function, to do what they do, and be what they are (“That’s the way I am”), without “kidding themselves” (“That’s the way it is”), and the practical materialism which inclines them to censor the expression of feelings or to divert emotion into violence or oaths, are the near-perfect antithesis of the aesthetic disavowal which, by a sort of essential hypocrisy (seen, for example, in the opposition between pornography and eroticism) masks the interest in function by the primacy given to form, so that what people do, they do as if they were not doing it.

The Visible and the Invisible

But food—which the working classes place on the side of being and substance, whereas the bourgeoisie, refusing the distinction between inside and outside or “at home” and “for others”, the quotidian and the extra-quotidian, introduces into it the categories of form and appearance—is itself related to clothing as inside to outside, the domestic to the public, being to seeming. And the inversion of the places of food and clothing in the contrast between the spending patterns of the working classes, who give priority to being, and the middle classes, where the concern for “seeming” arises, is the sign of a reversal of the whole world view. The working classes make a realistic or, one might say, functionalist use of clothing. Looking for substance and function rather than form, they seek “value for money” and choose what will “last”. Ignoring the bourgeois concern to introduce formality and formal dress into the domestic world, the place for freedom—an apron and slippers (for women), bare chest or a vest (for men)—they scarcely mark the distinction between top clothes, visible, intended to be seen, and underclothes, invisible or hidden—unlike the middle classes, who have a degree of anxiety about external appearances, both sartorial and cosmetic, at least outside and at work (to which middle-class women more often have access).

Thus, despite the limits of the data available, one finds in men’s clothing (which is much more socially marked, at the level of what can be grasped by statistics on purchases, than women’s clothing) the equivalent of the major oppositions found in food consumption. In the first dimension of the space, the division again runs between the office workers and the manual workers and is marked particularly by the opposition between grey or white overalls and blue dungarees or boiler-suits, between town shoes and the more relaxed moccasins, kickers or sneakers (not to mention dressing-gowns, which clerical workers buy 3.5 times more often

than manual workers). The increased quantity and quality of all purchases of men’s clothing is summed up in the opposition between the suit, the prerogative of the senior executive, and the blue overall, the distinctive mark of the farmer and industrial worker (it is virtually unknown in other groups, except craftsmen); or between the overcoat, always much rarer among men than women, but much more frequent among senior executives than the other classes, and the fur-lined jacket or lumber jacket, mainly worn by agricultural and industrial workers. In between are the junior executives, who now scarcely ever wear working clothes but fairly often buy suits.

Among women, who, in all categories (except farmers and farm labourers), spend more than men (especially in the junior and senior executive, professional and other high-income categories), the number of purchases increases as one moves up the social hierarchy; the difference is greatest for suits and costumes—expensive garments—and smaller for dresses and especially skirts and jackets. The topcoat, which is increasingly frequent among women at higher social levels, is opposed to the “all-purpose” raincoat, in the same way as overcoat and lumber jacket are opposed for men. The use of the smock and the apron, which in the working classes is virtually the housewife’s uniform, increases as one moves down the hierarchy (in contrast to the dressing-gown, which is virtually unknown among peasants and industrial workers).

The interest the different classes have in self-presentation, the attention they devote to it, their awareness of the profits it gives and the investment of time, effort, sacrifice and care which they actually put into it are proportionate to the chances of material or symbolic profit they can reasonably expect from it. More precisely, they depend on the existence of a labour market in which physical appearance may be valorized in the performance of the job itself or in professional relations; and on the differential chances of access to this market and the sectors of this

market in which beauty and deportment most strongly contribute to occupational value. A first indication of this correspondence between the propensity to cosmetic investments and the chances of profit may be seen in the gap, for all forms of beauty care, between those who work and those who do not (which must also vary according to the nature of the job and the work environment). It can be understood in terms of this logic why working-class women, who are less likely to have a job and much less likely to enter one of the occupations which most strictly demand conformity to the dominant norms of beauty, are less aware than all others of the “market” value of beauty and much less inclined to invest time and effort, sacrifices and money in cultivating their bodies.

It is quite different with the women of the petite bourgeoisie, especially the new petite bourgeoisie, in the occupations involving presentation and representation, which often impose a uniform (*tenue*) intended, among other things, to abolish all traces of heterodox taste, and which always demand what is called *tenue*, in the sense of “dignity of conduct and correctness of manners”, implying, according to the dictionary, “a refusal to give way to vulgarity or facility”. (In the specialized “charm schools” which train hostesses, the working-class girls who select themselves on the basis of “natural” beauty undergo a radical transformation in their way of walking, sitting, laughing, smiling, talking, dressing, making-up etc.) Women of the petite bourgeoisie who have sufficient interests in the market in which physical properties can function as capital to recognize the dominant image of the body unconditionally without possessing, at least in their own eyes (and no doubt objectively) enough body capital to obtain the highest profits, are, here too, at the site of greatest tension.

The self-assurance given by the certain knowledge of one’s own value, especially that of one’s body or speech, is in fact very closely linked to the position occupied in social

space (and also, of course, to trajectory). Thus, the proportion of women who consider themselves below average in beauty, or who think they look older than they are, falls very rapidly as one moves up the social hierarchy. Similarly, the ratings women give themselves for the different parts of their bodies tend to rise with social position, and this despite the fact that the implicit demands rise too. It is not surprising that petit-bourgeois women—who are almost as dissatisfied with their bodies as working-class women (they are the ones who most often wish they looked different and who are most discontented with various parts of their bodies), while being more aware of the usefulness of beauty and more often recognizing the dominant ideal of physical excellence—devote such great investments, of self-denial and especially of time, to improving their appearance and are such unconditional believers in all forms of cosmetic voluntarism (e.g., plastic surgery).

As for the women of the dominant class, they derive a double assurance from their bodies. Believing, like petit-bourgeois women, in the value of beauty and the value of the effort to be beautiful, and so associating aesthetic value and moral value, they feel superior both in the intrinsic, natural beauty of their bodies and in the art of self-embellishment and everything they call *tenue*, a moral and aesthetic virtue which defines “nature” negatively as sloppiness. Beauty can thus be simultaneously a gift of nature and a conquest of merit, as much opposed to the abdications of vulgarity as to ugliness.

Thus, the experience par excellence of the “alienated body”, embarrassment, and the opposite experience, ease, are clearly unequally probable for members of the petite bourgeoisie and the bourgeoisie, who grant the same recognition to the same representation of the legitimate body and legitimate deportment, but are unequally able to achieve it. The chances of experiencing one’s own body as a vessel of grace, a continuous

miracle, are that much greater when bodily capacity is commensurate with recognition; and, conversely, the probability of experiencing the body with unease, embarrassment, timidity grows with the disparity between the ideal body and the real body, the dream body and the “looking-glass self” reflected in the reactions of others (the same laws are also true of speech).

Although it is not a petit-bourgeois monopoly, the petit-bourgeois experience of the world starts out from timidity, the embarrassment of someone who is uneasy in his body and his language and who, instead of being “as one body with them,” observes them from outside, through other people’s eyes, watching, checking, correcting himself, and who, by his desperate attempts to reappropriate an alienated being-for-others, exposes himself to appropriation, giving himself away as much by hyper-correction as by clumsiness. The timidity which, despite itself, realizes the objectified body, which lets itself be trapped in the destiny proposed by collective perception and statement (nicknames etc.), is betrayed by a body that is subject to the representation of others even in its passive, unconscious reactions (one feels oneself blushing). By contrast, ease, a sort of indifference to the objectifying gaze of others which neutralizes its powers, presupposes the self-assurance which comes from the certainty of being able to objectify that objectification, appropriate that appropriation, of being capable of imposing the norms of appearance of one’s own body, in short, of commanding all the powers which, even when they reside in the body and apparently borrow its most specific weapons, such as “presence” or charm, are essentially irreducible to it. This is the real meaning of the findings of the experiment by W. D. Dannenmaier and F. J. Thumin, in which the subjects, when asked to assess the height of familiar persons from memory, tended to overestimate most of the height of those who

had most authority or prestige in their eyes.²¹ It would seem that the logic whereby the “great” are perceived as physically greater than they are applies very generally, and that authority of whatever sort contains a power of seduction which it would be naive to reduce to the effect of self-interested servility. That is why political contestation has always made use of caricature, a distortion of the bodily image intended to break the charm and hold up to ridicule one of the principles of the effect of authority imposition.

Charm and charisma in fact designate the power, which certain people have, to impose their own self-image as the objective and collective image of their body and being; to persuade others, as in love or faith, to abdicate their generic power of objectification and delegate it to the person who should be its object, who thereby becomes an absolute subject, without an exterior (being his own Other), fully justified in existing, legitimated. The charismatic leader manages to be for the group what he is for himself, instead of being for himself, like those dominated in the symbolic struggle, what he is for others. He “makes” the opinion which makes him; he constitutes himself as an absolute by a manipulation of symbolic power which is constitutive of his power since it enables him to produce and impose his own objectification.

The Universes of Stylistic Possibles

Thus, the spaces defined by preferences in food, clothing or cosmetics are organized according to the same fundamental structure, that of the social space determined by volume and composition of capital. Fully to construct the space of life-styles within which cultural practices are defined, one would first have to establish, for each class and class fraction, that is, for each of the configurations of capital, the generative formula of the habitus which retranslates the necessi-

ties and facilities characteristic of that class of (relatively) homogeneous conditions of existence into a particular life-style. One would then have to determine how the dispositions of the habitus are specified, for each of the major areas of practice, by implementing one of the stylistic possibilities offered by each field (the field of sport, or music, or food, decoration, politics, language etc.). By superimposing these homologous spaces one would obtain a rigorous representation of the space of life-styles, making it possible to characterize each of the distinctive features (e.g., wearing a cap or playing the piano) in the two respects in which it is objectively defined, that is, on the one hand by reference to the set of features constituting the area in question (e.g., the system of hairstyles), and on the other hand by reference to the set of features constituting a particular life-style (e.g., the working-class life-style), within which its social significance is determined.

SOURCES

The survey on which the work was based was carried out in 1963, after a preliminary survey by extended interview and ethnographic observation, on a sample of 692 subjects (both sexes) in Paris, Lille, and a small provincial town. To obtain a sample large enough to make it possible to analyse variations in practices and opinions in relation to sufficiently homogeneous social units, a complementary survey was carried out in 1967-68, bringing the total number of subjects to 1,217. Because the survey measured relatively stable dispositions, this time-lag does not seem to have affected the responses (except perhaps for the question on singers, an area of culture where fashions change more rapidly).

The following complementary sources (C.S.) were also used:

I. The 1966 survey on "businessmen and senior executives" was carried out by SOFRES (Société française d'enquêtes par sondages) on behalf of the Centre d'études des supports de publicité (CESP). The sample consisted of 2,257 persons aged 15 and over, each living in a household the head of which was a large industrial or commercial employer, a member of the professions, a senior executive,

an engineer, or a secondary or higher-education teacher. The questionnaire included a set of questions on reading habits and the previous few days' reading of daily, weekly and monthly newspapers and magazines, use of radio and TV, standard of living, household equipment, life-style (holidays, sport, consumption), professional life (conferences, travel, business meals), cultural practices and the principal basic data (educational level, income, population of place of residence etc.). I had access to the whole set of distributions by the socio-occupational category of the head of household or individual.

II. The regular survey by INSEE (Institut national de la statistique et des études économiques) on household living conditions and expenditure was based in 1972 on a representative sample of 13,000 households. It consists of a survey by questionnaire dealing with the characteristics of the household (composition, ages, occupation of the head), accommodation and facilities, major expenditure (clothing, fuel etc.), periodic expenditure (rent, service charges etc.), combined with analysis of account books for current expenditure, left with each household for a week, collected and checked by the interviewer. This survey makes it possible to assess the whole range of expenditure (except certain major and infrequent items such as air travel, removal expenses etc.), as well as items of consumption not preceded by purchase (food in the case of farmers, items drawn from their stocks by craftsmen and shopkeepers), which are evaluated at their retail price to allow comparison with the other categories of households. This explains why consumption is considerably higher than income in the case of farmers and small businessmen (categories which are always particularly prone to under-declare their income). For the overall findings, see G. Bigata and B. Bouvier, "Les conditions de vie des ménages en 1972", *Collections de l'INSEE*, ser. M, no. 32 (February 1972). The data presented here come from secondary analysis of tables by narrow categories produced at my request.

The other surveys consulted are simply listed below. These studies, almost always devoted to a particular area of cultural activity, are generally based on relatively limited samples. They mostly use a classification which groups the occupations into five categories: (1) *agriculteurs* (farmers and farm labourers); (2) *ouvriers* (industrial manual workers); (3) industrial and commercial employers; (4) clerical workers and junior executives; (5) senior executives and the "liberal professions".

III. INSEE, "La consommation alimentaire des Français", *Collections de l'INSEE*. (The regular INSEE surveys on eating habits.)

IV. SOFRES, *Les habitudes de table des Français* (Paris, 1972).

V. SOFRES, *Les Français et la gastronomie*, July 1977. (Sample of 1,000.)

NOTES

1. The gaps are more clear-cut and certainly more visible as regards education than income, because information on incomes (based on tax declarations) is much less reliable than information on qualifications. This is especially true of industrial and commercial employers (who, in the CESP survey—C.S. I—provided, along with doctors, the highest rate of non-response to the questions about income), craftsmen, shopkeepers, and farmers.

2. It follows from this that the relationship between conditions of existence and practices or the meaning of practices is not to be understood in terms either of the logic of mechanism or of the logic of consciousness.

3. In contrast to the atomistic approach of social psychology, which breaks the unity of practice to establish partial "laws" claiming to account for the products of practice, the *opus operatum*, the aim is to establish general laws reproducing the laws of production, the *modus operandi*.

4. Economic theory, which treats economic agents as interchangeable actors, paradoxically fails to take account of the economic dispositions, and is thereby prevented from really explaining the systems of preferences which define incommensurable and independent subjective use-values.

5. An ethic, which seeks to impose the principles of an ethos (i.e., the forced choices of a social condition) as a universal norm, is another, more subtle way of succumbing to *amor fati*, of being content with what one is and has. Such is the basis of the felt contradiction between ethics and revolutionary intent.

6. "Bourgeois" is used here as shorthand for "dominant fractions of the dominant class", and "intellectual" or "artist" functions in the same way for "dominated fractions of the dominant class".

7. F. Nietzsche, *Der Wille zur Macht* (Stuttgart, Alfred Kröner, 1964), no. 943, p. 630.

8. Bananas are the only fruit for which manual workers and farm workers have higher annual per capita spending (FF 23.26 and FF 25.20) than all other classes, especially the senior executives, who

spend most on apples (FF 31.60 as against FF 21.00 for manual workers), whereas the rich, expensive fruits—grapes, peaches and nuts—are mainly eaten by professionals and industrial and commercial employers (FF 29.04 for grapes, 19.09 for peaches and 17.33 for nuts, as against FF 6.74, 11.78 and 4.90, respectively, for manual workers).

9. This whole paragraph is based on secondary analysis of the tables from the 1972 INSEE survey on household expenditure on 39 items by socio-occupational category (C.S. II).

10. A fuller translation of the original text would include: "les *nourritures* à la fois les plus *nourrissantes* et les plus *économiques*" (the double tautology showing the reduction to pure economic function.)" (Translator's note.)

11. In the French: "le gros et le gras, gros rouge, gros sabots, gros travaux, gros rire, grosses blagues, gros bon sens, plaisanteries grasses"—cheap red wine, clogs (i.e., obviousness), heavy work, belly laughs, crude common sense, crude jokes (translator).

12. K. Marx, *Capital*, I (Harmondsworth, Penguin, 1976), 482 (translator).

13. One fine example, taken from Böhm-Bawerk, will demonstrate this essentialism: "We must now consider a *second* phenomenon of human experience—one that is heavily fraught with consequence. That is the fact that we feel less concerned about future sensations of joy and sorrow simply because they do lie in the future, and the lessening of our concern is in proportion to the remoteness of that future. Consequently we accord to goods which are intended to serve future ends a value which falls short of the true intensity of their future marginal utility. *We systematically undervalue our future wants and also the means which serve to satisfy them.*" E. Böhm-Bawerk, *Capital and Interest*, II (South Holland, Ill., 1959), 268, quoted by G. L. Stigler and G. S. Becker, "De gustibus non est disputandum", *American Economic Review*, 67 (March 1977), 76–90.

14. We may assume that the deep-seated relation to the future (and also to one's own person—which is valued more at higher levels of the social hierarchy) is reflected in the small proportion of manual workers who say that "there is a new life after death" (15 percent, compared with 18 percent of craftsmen and shopkeepers, office workers and middle managers, and 32 percent of senior executives).

15. The oppositions are much less clear-cut in the middle classes, although homologous differences are found between primary teachers and office workers on the one hand and shopkeepers on the other.

16. *La bouffe*: “grub”, “nosh”; *grande bouffe*: “blow-out” (translator).

17. The preference for foreign restaurants—Italian, Chinese, Japanese and, to a lesser extent, Russian—rises with level in the social hierarchy. The only exceptions are Spanish restaurants, which are associated with a more popular form of tourism, and North African restaurants, which are most favoured by junior executives (C.S. IV).

18. *Les gros*: the rich; *grosse bouffe*: bulk food (cf. *grossiste*: wholesaler, and English “grocer”). See also note 17 above (translator).

19. One could similarly contrast the bowl, which is generously filled and held two-handed for unpretentious drinking, and the cup, into which a little is poured, and more later (“Would you care for a little more coffee?”), and which is held between two fingers and sipped from.

20. Formality is a way of denying the truth of the social world and of social relations. Just as popular “functionalism” is refused as regards food, so too there is a refusal of the realistic vision which leads the working classes to accept social exchanges for what they are (and, for example, to say, without cynicism, of someone who has done a favour or rendered a service, “She knows I’ll pay her back”). Suppressing avowal of the calculation which pervades social relations, there is a striving to see presents, received or given, as “pure” testimonies of friendship, respect, affection, and equally “pure” manifestations of generosity and moral worth.

21. W. D. Dannenmaier and F. J. Thumin, “Authority Status as a Factor in Perceptual Distortion of Size”, *Journal of Social Psychology*, 63 (1964), 361–365.

94 ■ Melvin L. Kohn

Job Complexity and Adult Personality

There have been several distinct approaches to the study of work, each of them emphasizing some aspect that bears on a particular theoretical concern of the investigator. Rarely, though, has that concern been the effect of work on personality. Sociologists have learned much about social stratification and mobility, for example, by focusing on the dimension of work most pertinent to the stratificational system, the status of the job (Blau and Duncan, 1967; Duncan, Featherman, and Duncan, 1972; Sewell, Hauser, and Featherman, 1976). But however important status may be for studies of mobility, it would be unwise to assume—as is often done—that the status of a job is equally pertinent for personality. In terms of impact on personality, job status serves mainly as a

gross indicator of the job’s location in the hierarchical organization of the economic and social system. The status of the job is closely linked to such structural conditions of work as how complex it is, how closely it is supervised, and what sorts of pressures it entails. It is these structural realities, not status as such, that affect personality (Kohn and Schooler, 1973).

It is also indisputable that economists have learned much about the functioning of the economic system by focusing on the extrinsic rewards the job confers—in particular, income. Just as with occupational status, though, it would be incorrect to assume that because income is important for an understanding of the economic system *qua* system, income is also the most significant

aspect of the job in terms of the meaning of work to the worker or the impact of work on his sense of self and orientation to the rest of the world (Whyte, 1955; Kohn and Schooler, 1973).

Organizational theorists, both the Weberian sociologists and those more applied scholars who call themselves administrative scientists, have, by studying formal organizational structure, undoubtedly contributed much to our understanding of how organizations function (Blau and Schoenherr, 1971). But the very strength of their approach—its systematic attention to how formal organizations function as systems, regardless of the personalities of those who play the various organizational roles—means that they largely ignore the effect of organizational structure on the individual worker and his work. When they do pay attention to the individual worker, their interest rarely goes beyond his role as worker.

The human-relations-in-industry approach, in deliberate juxtaposition to the formal organizational approach, focuses on informal, interpersonal relationships and the symbolic systems that emerge out of such relationships (Whyte, 1961; 1969). Scholars using this approach supply a needed corrective to the formal organizational perspective. But they sometimes seem unaware that people not only relate to one another on the job; they also work. Moreover, this perspective has been concerned almost exclusively with the implications of work for on-the-job behavior, paying little attention to the effects of the job on other realms of life.

Occupational psychologists come close to understanding the relationship between work and personality, but there are two major limitations to their studies. First, many of them misinterpret Kurt Lewin (and, I would add, W. I. Thomas) by dealing exclusively with how people perceive their work while neglecting the actual conditions under which that work is per-

formed. (This criticism applies as well to most sociological studies of alienation in work.) Thus, they measure boredom rather than routinization, interest in the work rather than its substantive complexity, and “alienation in work” rather than actual working conditions. Such an approach ignores the possibilities that there can be a gap between the conditions to which a person is subjected and his awareness of those conditions; that the existence or nonexistence of such a gap is itself problematic and may be structurally determined; and that conditions felt by the worker to be benign can have deleterious consequences, while conditions felt to be onerous can have beneficial consequences. The second limitation is the preoccupation of most occupational psychologists with job satisfaction, as if this were the only psychological consequence of work. I am less disturbed by the technical difficulties in measuring job satisfaction—a notoriously slippery concept—than I am by the assumption that work has no psychological ramifications beyond the time and place during and within which it occurs. Work affects people’s values, self-conceptions, orientation to social reality, even their intellectual functioning. Job satisfaction is only one, and far from the most important, psychological consequence of work.

The research that comes closest of all to dealing straightforwardly with work and its consequences for personality employs an old tradition of sociological study—case studies of occupations. Practitioners of this art have sometimes done a magnificent job of depicting the reality of work as it impinges on the worker. Unfortunately, though, their studies cannot determine which aspects of work are most pertinent for which aspects of psychological functioning. W. Fred Cottrell’s (1940) classic study of railroaders, for example, pointed out a multitude of ways that the job conditions of men who operate trains differ from those of men

in many other occupations—including the unpredictability of working hours, geographical mobility, precision of timing, outsider status in the home community, and unusual recruitment and promotion practices. Since all these conditions are tied together in one occupational package, it is not possible to disentangle the psychological concomitants of each. More recent comparative studies face similar interpretative problems. For example, Robert Blauner's (1964) study of alienation among blue-collar workers in four industries, chosen to represent four technological levels, showed that differences in working conditions are systematically associated with the stage of technological development of the industry. But these differences, too, come in packages: Printing differs from automobile manufacture not only in technology and in the skill levels of workers but also in pace of the work, closeness of supervision, freedom of physical movement, and a multitude of other conditions. One cannot tell which, if any, of these interlocked occupational conditions are conducive to alienation.

Disentangling Occupational Conditions

Disentangling occupational conditions to assess their psychological impact requires a mode of research different from that employed in studies of particular occupations and particular industries. Carmi Schooler and I have dealt with the problem by shifting the focus from named occupations—carpenter, surgeon, or flight engineer—to dimensions of occupation (Kohn, 1969; Kohn and Schooler, 1969; 1973). Our strategy has been to secure a large and representative sample of employed men, who necessarily work in many occupations and many industries.¹ We have inventoried the men's job conditions and then differentiated the psychological concomitants of each

facet of occupation by statistical analysis. In our most recent research, my colleagues and I have done comparable analyses for employed women (Miller et al., 1979).

Even though occupational conditions are intercorrelated, they are not perfectly intercorrelated. Thus, substantively complex jobs are likely also to be time-pressured; but there are enough jobs that are substantively complex yet not time-pressured, and enough that are substantively simple yet time-pressured, for us to examine the relationship between substantive complexity and, say, receptiveness or resistance to change, while statistically controlling time pressure. We can also look for statistical interaction between the two, asking whether the impact of substantive complexity on, let us say, stance toward change is different for men who are more time-pressured and for men who are less time-pressured. And, at the same time, we can statistically control many other occupational conditions, as well as important nonoccupational variables, for example, education, which usually precedes and is often a prerequisite for the job.

In all, we have indexed more than fifty separable dimensions of occupation, including such diverse aspects of work experience as the substantive complexity of work, the routinization or diversity of the flow of work, relationships with co-workers and with supervisors, pace of work and control thereof, physical and environmental conditions, job pressures and uncertainties, union membership and participation, bureaucratization, job protections, and fringe benefits. (For complete information, see Kohn, 1969, pp. 236, 244–253.) These indices provide the basis for a broad descriptive picture of the principal facets of occupations, as experienced by men in all types of industries and at all levels of the civilian economy.

I must admit that this approach—based as it is on a sample survey of all men employed in civilian occupations—is not optimum for

securing some kinds of job information. Men may have only limited information about certain aspects of their jobs, such as the overall structure of the organization in which they work. Moreover, a sample of men scattered across many occupations and many work places does not contain enough people in any occupation or any workplace to trace out interpersonal networks and belief systems. Similarly, the method is not well adapted for studying the industrial and technological context in which the job is embedded. The method is most useful for studying the immediate conditions of a man's own job—what he does, who determines how he does it, in what physical and social circumstances he works, to what risks and rewards he is subject.

We found that nearly all of the more than fifty occupational conditions that we had inventoried are correlated with at least some of the several aspects of values, self-conception, social orientation, and intellectual functioning that we had measured (Kohn and Schooler, 1973).² But most of these statistical relationships reflect the interrelatedness of occupational conditions with one another and with education. Only twelve of the occupational conditions we studied appear to have any substantial relationship to men's psychological functioning when education and all other pertinent occupational conditions are statistically controlled. Few though they are, these twelve occupational conditions are sufficient to define the "structural imperatives of the job," in that they identify a man's position in the organizational structure, his opportunities for occupational self-direction, the principal job pressures to which he is subject, and the principal uncertainties built into his job.³ These job conditions are "structural" in two senses: they are built into the structure of the job and they are largely determined by the job's location in the structures of the economy and the society.

Substantive Complexity

Because of its theoretical and empirical importance, I devote the remainder of this [chapter] to one of the twelve structural imperatives of the job, the substantive complexity of work.⁴ By the substantive complexity of work, I mean the degree to which the work, in its very substance, requires thought and independent judgment. Substantively complex work by its very nature requires making many decisions that must take into account ill-defined or apparently conflicting contingencies. Although, in general, work with data or with people is likely to be more complex than work with things, this is not always the case, and an index of the overall complexity of work should reflect its degree of complexity in each of these three types of activity. Work with things can vary in complexity from ditch digging to sculpting; similarly, work with people can vary in complexity from receiving simple directions or orders to giving legal advice; and work with data can vary from reading instructions to synthesizing abstract conceptual systems. Thus, the index of substantive complexity that we have generally employed is based on the degree of complexity of the person's work with things, with data, and with people; our appraisal of the overall complexity of his work, regardless of whether he works primarily with things, with data, or with people; and estimates of the amount of time he spends working at each type of activity (Kohn and Schooler, 1973; Kohn, 1976). The several components receive weightings based on a factor analysis.⁵

I focus on substantive complexity for two reasons. The first is that I conceive substantive complexity to be central to the experience of work. The other structural imperatives of the job—even closeness of supervision and routinization—set the conditions under which work is done but they do not characterize the work itself. The substantive com-

plexity of work, by contrast, is at the heart of the experience of work. More than any other occupational condition, it gives meaning to this experience.

The second reason for my preoccupation with substantive complexity is empirical. Our analyses show that substantive complexity of work is strongly related to a wide range of psychological variables. The substantive complexity of work is of course correlated with job satisfaction and with occupational commitment. It also bears on many facets of off-the-job psychological functioning, ranging from valuation of self-direction to self-esteem to authoritarian conservatism to intellectual flexibility. It is even related to the intellectual demands of men's leisure-time pursuits. Moreover, these correlations remain statistically significant and large enough to be meaningful even when education and all other pertinent dimensions of occupation are statistically controlled. Thus, the substantive complexity of work has a strong, independent relationship to many facets of psychological functioning, a relationship stronger than that of any other dimension of occupation we have studied. This is true for men, and our most recent analyses show it to be equally true for employed women (Miller et al., 1979).

The Direction of Causal Effects

There is evidence that the substantive complexity of work is not only correlated with, but has a causal impact on, psychological functioning. The evidence of causal directionality is of two types. The more extensive but less definitive evidence comes from our analyses of cross-sectional data, derived from a large sample of men employed in civilian occupations. Social scientists have long recognized that one cannot make inferences about the direction of causal effects from cross-sectional data unless some of the described phenomena clearly preceded oth-

ers in their time of occurrence. But where one can realistically assume reciprocity—that *a* affects *b* and *b* also affects *a*—it is possible to assess the magnitude of these reciprocal effects, using econometric techniques for solving simultaneous equations. The simplest of these, which we used, is called two-stage least squares.⁶ With this technique, we have assessed the relationships between the substantive complexity of work and many facets of psychological functioning: occupational commitment, job satisfaction, valuation of self-direction or of conformity to external authority, anxiety, self-esteem, receptiveness or resistance to change, standards of morality, authoritarian conservatism, intellectual flexibility, the intellectuality of leisure-time activities, and three types of alienation—powerlessness, self-estrangement, and normlessness.

Our findings indicate that the substantive complexity of men's work affects all these facets of psychological functioning, independent of the selection processes that draw men into particular fields of work and independent of their efforts to mold their jobs to fit their needs, values, and capacities. Moreover, the substantive complexity of work in every instance affects psychological functioning more—often, much more—than the particular facet of psychological functioning affects the substantive complexity of work. This evidence is not definitive—only longitudinal studies can provide definitive evidence—but it does establish a strong *prima facie* case that the substantive complexity of work has a real and meaningful effect on a wide range of psychological phenomena.

More definitive, albeit less extensive, evidence comes from a follow-up study we conducted with a representative subsample of men in the original study ten years after the initial survey (Kohn and Schooler, 1978). Analyses of longitudinal data require the development of "measurement models" that separate unreliability of measurement from

change in the phenomena studied. The essence of the method employed in constructing these models is the use of multiple indicators for each principal concept, inferring from the covariation of the indicators the degree to which each reflects the underlying concept that they are all hypothesized to reflect and the degree to which each reflects anything else, which for measurement purposes is considered to be error (see Jöreskog, 1969, and other papers cited in Kohn and Schooler, 1978). These models permit us to take into account that errors in the measurement of any indicator at the time of the initial survey may well be correlated with errors in the measurement of that same indicator at the time of the follow-up survey. Disregarding such correlated errors in the indicators might make the underlying concept seem more stable or less stable than it really is—thereby distorting any causal analysis in which that index is employed. . . .

We chose intellectual flexibility as the first aspect of psychological functioning to be assessed because it offers us the greatest challenge—intellectual flexibility obviously affects recruitment into substantively complex jobs, and there is every reason to expect it to be resistant to change. Still, intellectual flexibility—though not much studied by sociologists—is so important a part of psychological functioning that we must not unthinkingly assume it to be entirely the product of genetics and early life experience. Rather, we should empirically test the possibility that intellectual flexibility may be responsive to adult occupational experience.

Our index of intellectual flexibility is meant to reflect men's actual intellectual performance in the interview situation. We used a variety of indicators—including the men's answers to seemingly simple but highly revealing cognitive problems, their handling of perceptual and projective tests, their propensity to agree when asked "agree-disagree" questions, and the impression they made on

the interviewer during a long session that required a great deal of thought and reflection. None of these indicators is believed to be completely valid; but we do believe that all the indicators reflect, to some substantial degree, men's flexibility in coping with an intellectually demanding situation.

The stability of intellectual flexibility, thus measured, is remarkably high over time: The correlation between men's intellectual flexibility at the time of the original study and their intellectual flexibility ten years later, shorn of measurement error, is 0.93. It would be erroneous to assume, though, that the high over-time stability of intellectual flexibility means that it is unaffected by adult experience; it might even be that this stability reflects unchanging life circumstances. In fact, we find that the effect of the substantive complexity of work on intellectual flexibility is striking—on the order of one-fourth as great as that of the men's ten-year-earlier levels of intellectual flexibility. This effect is essentially contemporaneous: The path from the substantive complexity of the job held at the time of the initial survey to intellectual flexibility at the time of the follow-up survey ten years later is small and statistically nonsignificant, while the path from the substantive complexity of the current job to current intellectual flexibility is much more substantial and is statistically significant.⁷

The reciprocal effect of intellectual flexibility on substantive complexity is still more impressive than the effect of substantive complexity on intellectual flexibility. This effect is entirely lagged, that is, it is the men's intellectual flexibility at the time of the initial survey that significantly affects the substantive complexity of their current jobs, and not their current intellectual flexibility. The longitudinal analysis thus demonstrates something that no cross-sectional analysis could show—that, over time, the relationship between substantive complexity and intellectual flexibility is truly reciprocal. The

effect of substantive complexity on intellectual flexibility is more immediate: current job demands affect current thinking processes. Intellectual flexibility, by contrast, has a time-lagged effect on substantive complexity: current intellectual flexibility has scant effect on current job demands, but considerable effect on the future course of one's career. Cross-sectional analyses portray only part of this process, making it seem that the relationship between the substantive complexity of work and intellectual functioning were mainly unidirectional, with work affecting intellectual functioning but not the reverse. Longitudinal analysis portrays a more intricate and more interesting, truly reciprocal process.

The data thus demonstrate, beyond reasonable doubt, what heretofore could be stated only as a plausible thesis buttressed by presumptive evidence—that the substantive complexity of work both considerably affects, and is considerably affected by, intellectual flexibility.

My colleagues and I have recently completed two further analyses that extend these conclusions. A much more extensive longitudinal analysis of job conditions and intellectual flexibility (Kohn and Schooler, 1981) confirms that the substantive complexity of work affects intellectual flexibility not only when prior levels of intellectual flexibility and pertinent aspects of social background are taken into account but also when all other structural imperatives of the job are taken into account as well. We further find that substantive complexity is not the only job condition that affects intellectual flexibility; several other job conditions that stimulate and challenge the individual are conducive to intellectual flexibility. But, clearly, substantive complexity plays a key role, not only because it has such a great effect on intellectual flexibility, but also because it provides the principal mechanism through which other job conditions affect intellectual functioning.

In another analysis (Miller et al., 1979), we found that the substantive complexity of work is as important for women's psychological functioning as it is for men's. In particular, a causal analysis using measurement models similar to those described above, but limited to cross-sectional data, shows the contemporaneous effect of substantive complexity on intellectual flexibility to be at least as great for employed women as for employed men.

These findings come down solidly in support of those who hold that occupational conditions affect personality and in opposition to those who believe that the relationship between occupational conditions and personality results solely from selective recruitment and job molding. Admittedly, personality has great importance in determining who go into what types of jobs and how they perform those jobs; in fact, our analyses underline the importance of these processes. But that has never been seriously at issue. What has been disputed is whether the reverse phenomenon—of job conditions molding personality—also occurs. The evidence of our longitudinal analysis supports the position that it does occur.

In particular, this analysis adds to and helps specify the growing evidence that the structure of the environment has an important effect on cognitive development (Rosenbaum, 1976) and that cognitive processes do not become impervious to environmental influence after adolescence or early adulthood but continue to show "plasticity" throughout the life span (Baltes, 1968; Horn and Donaldson, 1976; Baltes and Schaie, 1976). Our findings reinforce this conclusion by showing that intellectual flexibility continues to be responsive to experience well into midcareer. In fact, it appears that the remarkable stability of intellectual flexibility reflects, at least in part, stability in people's life circumstances. Intellectual flexibility is ever responsive to changes in the substantive complexity of

people's work; for most people, though, the substantive complexity of work does not fluctuate markedly.

This analysis demonstrates as well the importance of intellectual flexibility for substantive complexity. I think it noteworthy that this effect appears to be lagged rather than contemporaneous. The implication is that the structure of most jobs does not permit any considerable variation in the substantive complexity of the work: job conditions are not readily modified to suit the needs or capacities of the individual worker. But over a long enough time—certainly over a period as long as ten years—many men either modify their jobs or move on to other jobs more consonant with their intellectual functioning. Thus, the long-term effects of intellectual flexibility on substantive complexity are considerable, even though the contemporaneous effects appear to be negligible.

Our models, of course, deal mainly with the events of midcareer or later. I think it reasonable to assume that men's intellectual flexibility in childhood, adolescence, and early adulthood have had a considerable effect on their educational attainments, and our data show that educational attainment is very important for the substantive complexity of the early jobs in men's careers. Since the substantive complexity of early jobs is a primary determinant of the substantive complexity of later jobs, it seems safe to infer that intellectual flexibility's long-term, indirect effects on the substantive complexity of later jobs has been even greater than our analysis depicts.

The reciprocal relationship between substantive complexity and intellectual flexibility implies an internal dynamic by which relatively small differences in substantive complexity at early stages of a career may become magnified into larger differences in both substantive complexity and intellectual flexibility later in the career. If two men of equivalent intellectual flexibility were to start

their careers in jobs differing in substantive complexity, the man in the more complex job would be likely to outstrip the other in further intellectual growth. This, in time, might lead to his attaining jobs of greater complexity, further affecting his intellectual growth. Meantime, the man in the less complex job would develop intellectually at a slower pace, perhaps not at all, and in the extreme case might even decline in his intellectual functioning. As a result, small differences in the substantive complexity of early jobs might lead to increasing differences in intellectual development. . . .

Processes of Learning and Generalization

Why does substantive complexity have such wide-spread ramifications for personality? This question is still largely unanswered. Our findings suggest that the simplest type of learning-generalization process is operating here: that there is a direct translation of the lessons of the job to outside-the-job realities, rather than some indirect process, such as reaction formation or compensation (Breer and Locke, 1965). Thus, men who do complex work come to exercise their intellectual prowess not only on the job but also in their nonoccupational lives. They become more open to new experience. They come to value self-direction more highly. They even come to engage in more intellectually demanding leisure-time activities. In short, the lessons of work are directly carried over to nonoccupational realms.

Several alternative interpretations have been advanced, but none has been adequately tested. One such interpretation is that doing more (or less) substantively complex work leads to a stronger (or weaker) sense of control over the forces that affect one's life, and that this sense of controlling or not controlling those forces influences one's self-concept, in particular, one's self-esteem.

But why is it necessary to accord so strategic a role in the causal chain to “locus of control”? Why not say, more directly, that people who do substantively complex work come to think of themselves as capable of doing difficult and challenging tasks—and thus deserve respect in everyone’s eyes, including their own? Similarly, people who do substantively simple work come to think of themselves as capable of nothing more than this simple-minded stuff.

I think the most reasonable hypothesis is the most straightforward: that in an industrial society, where work is central to people’s lives, what people do in their work directly affects their values, their conceptions of self, and their orientation to the world around them—“I do, therefore I am.” Hence, doing substantively complex work tends to increase one’s respect for one’s own capacities, one’s valuation of self-direction, one’s intellectuality (even in leisure-time pursuits), and one’s sense that the problems one encounters in the world are manageable. There is no need to posit that locus of control or any other aspect of values and orientation plays a necessary intermediary role in this process; the substantive complexity of work can directly affect all aspects of people’s conceptions of reality. . . .

Still another possible interpretation is that occupational conditions do not affect all people similarly, but differently, depending on the individual’s own needs, values, and abilities—the so-called “fit” hypothesis. Our analyses do not support this interpretation. We repeatedly find, for example, that the substantive complexity of work has much the same effects regardless of whether men value intrinsic or extrinsic aspects of their work more highly (Kohn and Schooler, 1973; Kohn, 1976). But these analyses have been broad-gauge, and more detailed analyses of pertinent subpopulations might require us to modify our views. Moreover, our analyses of this issue have thus far not been longitudinal.

But since our data demonstrate that people’s values and even their abilities are affected by their job conditions, it is clear that only longitudinal assessments can be conclusive.

Our studies have led me to conclude that the intrinsic meaning and psychological impact of a job result not just from the status or income or interpersonal relationships that the job provides but also—and especially—from the meaningful challenges the work itself poses (or fails to pose). The most important challenge is that of mastering complex tasks, that is, the substantive complexity of the work. Our data indicate that substantive complexity affects people’s psychological functioning regardless of their needs, values, and personal capacities and regardless of their social class (but, of course, the type of work one does is intimately related to one’s social class; so, too, are one’s values). What matters most about work, in short, is not any of its attendant rewards or social experiences, but the work itself.

Moreover, the relationship between work and psychological functioning is quintessentially reciprocal. There is an ongoing process, throughout all of adult life, whereby the occupational conditions encountered by the individual both mold his psychological processes and in turn are molded by his personality and behavior. No theory of adult personality development that fails to take account of the ongoing effects of occupational (and, presumably, other social) conditions can be regarded as realistic. By the same token, no social psychology of occupations that fails to take account of the ongoing effects of individual psychological functioning can be regarded as realistic.

NOTES

1. Our primary source of data is a sample survey of 3,101 men, representative of all men employed in civilian occupations in the United States. These men were interviewed for us by the National Opinion Research Center (NORC) in the spring and summer of

1964. For more detailed information on sample and research design, see Kohn, 1969, pp. 235–264. In 1974 NORC reinterviewed a representative subsample of these men for us; this time, the wives (and, where applicable, one of the children) were interviewed, too. For detailed information on the follow-up study, see Kohn and Schooler (1978) and Kohn (1977).

2. Our principal indices of psychological functioning measure subjective reactions to the job itself (that is, job satisfaction and occupational commitment), valuation of self-direction or of conformity to external authority (both for oneself and for one's children), self-conception (self-confidence, self-deprecation, fatalism, anxiety, and idea conformity), social orientation (authoritarian conservatism, criteria of morality, trustfulness, and receptiveness or resistance to change), alienation (powerlessness, self-estrangement, and normlessness), and intellectual functioning (intellectual flexibility, intellectuality of leisure-time activities). For detailed information about our definitions of these concepts and our methods of indexing them, see Kohn, 1969, pp. 47–58, 73–84, 265–269; Kohn and Schooler, 1973, pp. 99–101; Kohn, 1976, pp. 114–118.

3. Specifically, these twelve crucial occupational conditions are: (1) ownership/nonownership; (2) bureaucratization; (3) position in the supervisory hierarchy; (4) closeness of supervision; (5) routinization of the work; (6) substantive complexity of the work; (7) frequency of time-pressure; (8) heaviness of work; (9) dirtiness of work; (10) the likelihood, in this field, of there occurring a sudden and dramatic change in a person's income, reputation, or position; (11) the probability, in this line of work, of being held responsible for things outside one's control; and (12) the risk of loss of one's job or business.

4. The concept "substantive complexity" has been the subject of much research that goes considerably beyond the issues addressed in this essay. Many writers have adopted the concept and used it for such diverse purposes as reinterpreting the status-attainment model (Spaeth, 1976), proposing a new method of classifying the occupational structure of the U.S. economy (Temme, 1975), reassessing the psychological effects of complex role sets (Coser, 1975), interpreting the effects of fathers' occupational experiences on their sons' occupational choices (Mortimer, 1974; 1976), and searching out the sources of powerlessness (Tudor, 1972).

5. To validate this index, which is specifically tailored to each respondent's description of his own job, we have compared it to assessments of the average level of complexity of work with things, with data, and with people for the entire occupation, made by trained occupational analysts for the *Dictionary of Occupational Titles* (United States Department of Labor, 1965). The multiple correlation between our index of substantive complexity and the independently coded *Dictionary* ratings is 0.78—sufficiently high to assure us that our appraisals of substantive complexity accurately reflect the reality of people's work.

6. The two-stage least squares technique is described in detail by Kohn and Schooler (1973) and the references cited therein. This method attempts to "purge" each variable of the effects of all others with which it is reciprocally related by estimating from other pertinent data what each individual's score on that variable would have been if the other variables had not had an opportunity to affect it. These estimated scores are then used as independent variables in the (second stage) multiple-regression equations.

7. Concretely, the time-lagged path (that is, from substantive complexity in 1964 to intellectual flexibility in 1974) is 0.05 and the contemporaneous path is 0.18. A path of 0.18 might not under ordinary circumstances be considered striking; but a continuing effect of this magnitude on so stable a phenomenon as intellectual flexibility is impressive, because the cumulative impact will be much greater than the immediate effect at any one time. Continuing effects, even small-to-moderate continuing effects, on highly stable phenomena become magnified in importance. The effect of the substantive complexity of work on intellectual flexibility is especially noteworthy when we take into account that we are dealing with men who are at least ten years into their occupational careers.

REFERENCES

- Baltes, Paul B. 1968. Longitudinal and cross-sectional sequences in the study of age and generation effects. *Human Development*, 11:145–171.
- Baltes, Paul B., and K. Warner Schaie. 1976. On the plasticity of intelligence in adulthood and old age. *American Psychologist*, 31:720–725.
- Blau, Peter M., and Otis Duncan. 1967. *The American occupational structure*. New York: Wiley.
- Blau, Peter M., and Richard A. Schoenherr. 1971. *The structure of organizations*. New York: Basic Books.

- Blauner, Robert. 1964. *Alienation and freedom: the factory worker and his industry*. Chicago: University of Chicago Press.
- Breer, Paul E., and Edwin A. Locke. 1965. *Task experience as a source of attitudes*. Homewood, Ill.: Dorsey.
- Coser, Rose Laub. 1975. The complexity of roles as a seedbed of individual autonomy. In *The idea of social structure: papers in honor of Robert K. Merton*, ed. L. A. Coser. New York: Harcourt Brace Jovanovich.
- Cottrell, W. Fred. 1940. *The railroader*. Stanford: Stanford University Press.
- Duncan, Otis D., David L. Featherman, and Beverly Duncan. 1972. *Socioeconomic background and achievement*. New York: Seminar Press.
- Horn, John L., and Gary Donaldson. 1976. On the myth of intellectual decline in adulthood. *American Psychologist*, 31:701–719.
- Jöreskog, Karl G. 1969. A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, 34:183–202.
- Kohn, Melvin L. 1977. Reassessment, 1977. In *Class and conformity: a study in values*. 2nd ed. Chicago: University of Chicago Press.
- . 1976. Occupational structure and alienation. *American Journal of Sociology* 82: 111–130.
- . 1969. *Class and conformity: a study in values*. Homewood, Ill.: Dorsey. (2nd ed., Chicago: University of Chicago Press, 1977.)
- Kohn, Melvin L., and Carmi Schooler. 1981. Job conditions and intellectual flexibility: a longitudinal assessment of their reciprocal effects. In *Factor analysis and measurement in sociological research: a multidimensional perspective*, ed. E. F. Borgatta and D. J. Jackson. Beverly Hills: Sage Publications.
- . 1978. The reciprocal effects of the substantive complexity of work and intellectual flexibility: a longitudinal assessment. *American Journal of Sociology*, 84:24–52.
- . 1973. Occupational experience and psychological functioning: an assessment of reciprocal effects. *American Sociological Review*, 38:97–118.
- . 1969. Class, occupation, and orientation. *American Sociological Review*, 34:659–678.
- Miller, Joanne, Carmi Schooler, Melvin L. Kohn, and Karen A. Miller. 1979. Women and work: the psychological effects of occupational conditions. *American Journal of Sociology*, 85:66–94.
- Mortimer, Jeylan T. 1976. Social class, work, and the family: some implications of the father's occupation for familial relationships and sons' career decisions. *Journal of Marriage and the Family*, 38:241–256.
- . 1974. Patterns of intergenerational occupational movements: a smallest-space analysis. *American Journal of Sociology*, 79:1278–1299.
- Rosenbaum, James E. 1976. *Making inequality: the hidden curriculum of high school tracking*. New York: Wiley.
- Sewell, William H., Robert M. Hauser, and David L. Featherman, eds. 1976. *Schooling and achievement in American society*. New York: Academic Press.
- Spaeth, Joe L. 1976. Cognitive complexity: a dimension underlying the socioeconomic achievement process. In *Schooling and achievement in American society*, ed. William H. Sewell, Robert M. Hauser, and David L. Featherman. New York: Academic Press.
- Temme, Lloyd V. 1975. *Occupation: meanings and measures*. Washington, D.C.: Bureau of Social Science Research.
- Tudor, Bill. 1972. A specification of relationships between job complexity and powerlessness. *American Sociological Review*, 37:596–604.
- United States Department of Labor. 1965. *Dictionary of occupational titles*. Washington, DC: United States Government Printing Office. 3rd ed.
- Whyte, William F. 1969. *Organizational behavior: theory and application*. Homewood, Ill.: Richard D. Irwin.
- . 1961. *Men at work*. Homewood, Ill.: Dorsey.
- . 1955. *Money and motivation: an analysis of incentives in industry*. New York: Harper.

Health

95 ■ John Mullahy, Stephanie Robert, and Barbara Wolfe

Health, Income, and Inequality

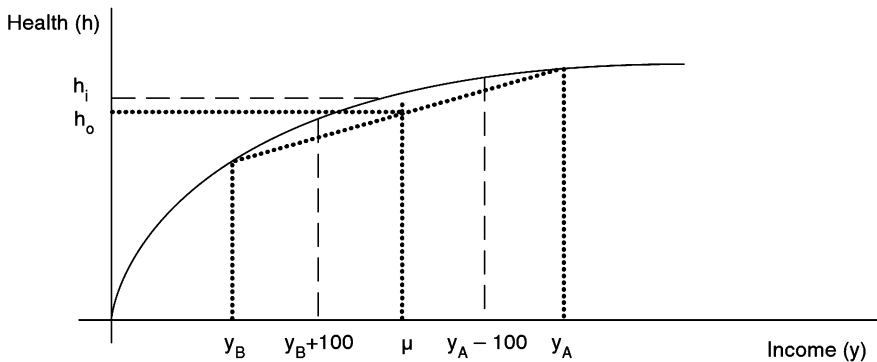
The idea that income is associated with health goes back a long way in the literature. Perhaps the most influential work affecting contemporary work in this area was by Samuel Preston (1975), who observed, in comparing mortality rates across countries, that the impact of additional income on health (as measured by mortality) is greater on those with low income than on those with higher income. This effect is illustrated in Figure 1, which shows the health of individuals on the vertical axis, plotted against the income of the individual's family group on the horizontal axis. The concave shape of the curve conveys the idea that because a dollar transferred from the rich to the poor improves the health of the poor person more than it decreases the health of the rich person, this transfer increases the average level of health of the members of a community.

This is called the *absolute income hypothesis*, and in its simplest form it argues that if all that matters to health at the level of an individual is income, a community with more equal income will tend to have better average health than a community with more inequality when comparing two communities with equal average income. In an international context, Angus Deaton (2001) points out that, according to the absolute income hy-

pothesis, redistribution could improve health even if average income was not increased, and that redistribution from rich to poor countries would in principle improve worldwide average health.

A related concept is the *absolute deprivation or poverty hypothesis*. According to this hypothesis, those with the lowest incomes face poorer health and a greater risk of mortality owing to a variety of factors associated with extreme poverty, such as inadequate nutrition, lack of quality health care, exposure to a variety of physical hazards, and heightened stress. In this hypothesis, a dollar redistributed from rich to poor would improve the health of the poor and improve the average health of the entire population. The difference between the absolute income hypothesis and the absolute deprivation or poverty hypothesis is that in the former greater income improves the health of all persons, although in a nonlinear way, while in the latter only those who have low incomes are expected to see health improvements as a result of an increase in income.

The *relative income hypothesis* focuses on an individual's income relative to that of others in that individual's "group" rather than on his or her absolute income. According to this hypothesis, if the income of everyone but



Source: Wagstaff and Van Doorslaer (2000, 546, fig. 2).

Figure 1. Implications of the Convexity of the Health-Income Relationship

one in a group increases, that one person's health is expected to deteriorate. A related hypothesis is the *relative position* or *relative deprivation hypothesis*. According to this hypothesis, it is an individual's relative rank in society that is tied to health outcomes. This hypothesis encompasses the relative income hypothesis, but it extends the concept of relative position to measures of rank other than income, such as occupational rank or educational rank. These hypotheses are consistent with some research in the United States and the United Kingdom that demonstrates that the association between socioeconomic position and health occurs at all levels of the socioeconomic hierarchy, with even those in the highest socioeconomic groups having better health than those just below them in the socioeconomic hierarchy, which is referred to as a "gradient effect" of socioeconomic position on health (Adler et al. 1994; Marmot et al. 1991).

The relative income and relative position hypotheses imply that it is not just the conditions experienced by those in absolute poverty that lead to poor health. Rather, there are psychosocial and other factors that remain unevenly distributed all the way up the income scale that perpetuate income inequalities in health. Perceptions of being relatively deprived ("keeping up with the Joneses"), stress, and other more psycho-

social and behavioral than material factors may play a role in perpetuating income inequalities in health at the upper income levels. Health effects at the upper end of the income distribution may reflect relative position, whereas health effects at the lower end may reflect absolute deprivation (Adler and Newman 2002).

The hypothesis that focuses most directly on the tie between health inequality and income inequality is the *income inequality hypothesis*. According to the strong version of this hypothesis, societies with greater inequality produce worse health among their citizens, holding constant the average income of societies. Although these arguments were first made in comparisons of income inequality between countries, research has more recently examined whether regions, states, counties, and cities with greater income inequality have worse health than their more equal counterparts.

A Brief Review of the Evidence

Although all of these hypotheses initially seem testable, controversy arises because they are not always mutually exclusive. For example, tests of the income inequality hypothesis at the aggregate level may be empirically consistent with tests of the absolute poverty or relative income hypotheses at the individual

level. Such overlap continues to obscure our understanding of the relationship between income, income inequality, and health. Without a clearer conceptual and empirical understanding of these relationships, program and policy recommendations about how to reduce income-related inequalities in health remain elusive and hotly debated.

In general, two empirical approaches have been taken to examine these hypotheses. Research examining the absolute deprivation, relative income, and relative position hypotheses has usually examined individual-level data on income and health or mortality to examine the existence and shape of the income-health relationship among individuals. Recent research has extended examination of these hypotheses by including aggregate measures of *community* socioeconomic level, such as the percentage of persons in poverty or community median family income, along with individual-level income data. In contrast, research examining the income inequality hypothesis has employed aggregate data either exclusively or at least at the level of measuring income inequality. We divide our brief review of the literature into these two general types of studies, focusing particularly on the recent research testing the income inequality hypothesis.

Income and Health

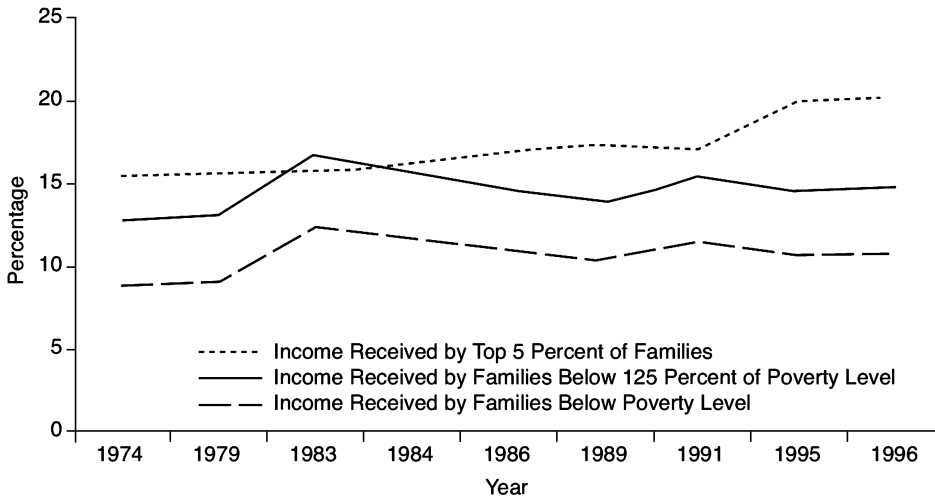
Voluminous empirical studies and reviews demonstrate a robust association between income and morbidity and mortality, using various measures of both income and health across samples and at various time points (Adler et al. 1993, 1994; Antonovsky 1967; Feinstein 1993; Williams and Collins 1995; Robert and House 2000a). To date, most of the evidence demonstrates a nonlinear rather than linear gradient relationship between income and mortality (Backlund, Sorlie, and Johnson 1996; McDonough et al. 1997; Ettner 1996) and between income and morbidity (House et al. 1990, 1994; Mirowsky and Hu 1996).

Although most research on income and health is cross-sectional, there is some evi-

dence that there may be widening socioeconomic inequalities in health in the United Kingdom (Black et al. 1982), the United States (Pappas et al. 1993), and other developed countries (Evans, Barer, and Marmor 1994; Marmot, Kogevinas, and Elston 1987). However, Deaton (2001, 60) suggests that the rapid increases in income inequality in the 1970s and 1980s in Britain and the United States “have not been associated with any slowdown in the rate of mortality decline.” To date, datasets have been of insufficiently long duration or too incomplete in terms of longitudinal information on both income and health to explore the relationship between income, income inequality, and health over time.

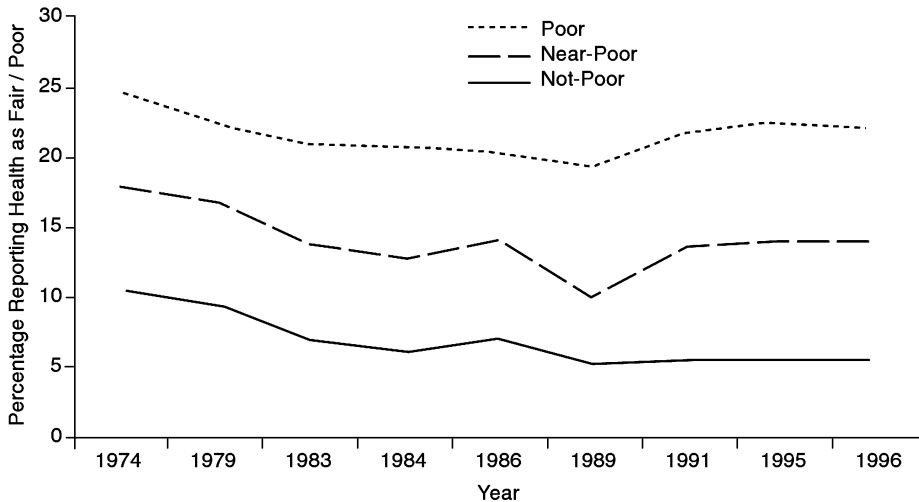
Figure 2 reports on trends in family income from 1974 to 1996. It illustrates the growing share of income received by the top 5 percent of families and shows the proportion of U.S. families below the official poverty line and the proportion who are near-poor (those below 125 percent of the poverty line). The growing share of the top 5 percent is indicative of growing inequality; the proportion poor and near-poor shows a more complicated picture—an increase from 1974 to 1983, a decline through 1989, followed by a smaller increase to 1991, with little change in the next few years.

Figure 3 reports the proportion of the population who reported poor or fair health by income group over these same years. Consistent with both the absolute income and deprivation hypotheses and the relative income and deprivation hypotheses, the health of the poor is always worse than that of the nonpoor, and the health of the near-poor is between the two. The trends over time show a continual improvement in the health of the nonpoor, but not among the poor and near-poor: in both cases, after some improvement over the period 1974 to 1989, there was some deterioration in health from 1991 to 1996. Caution should be used in viewing these trends, however, since the way in which these health sta-



Source: U.S. Department of Commerce (2000, tables 745, 760).

Figure 2. Trends in Income of U.S. Families, 1974 to 1996



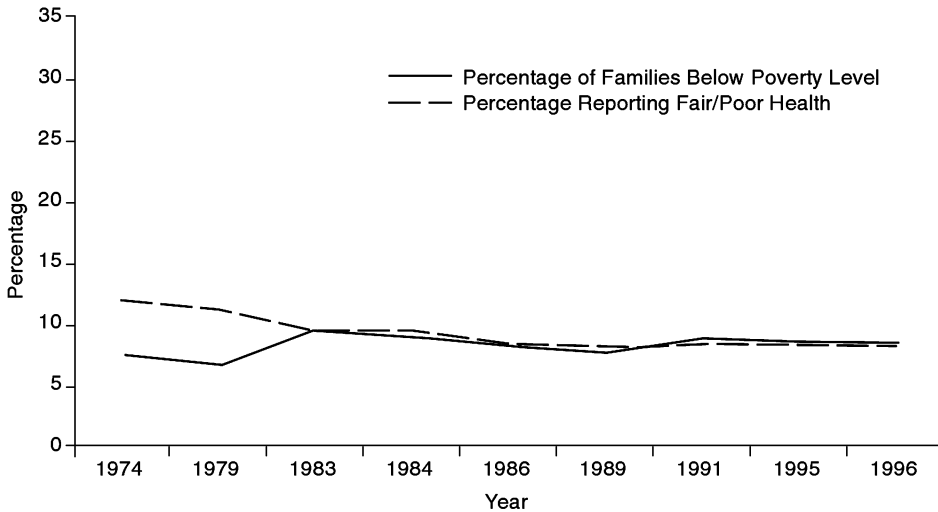
Source: National Center for Health Statistics (1998-1999).

Figure 3. Trends in Self-Reported Health of U.S. Families by Family Income, 1974 to 1996

tistics are reported changed over this time period; hence the “jump” may be due to the reporting change. The increase after 1991, however, cannot be explained by the change in reporting, since there were no further changes.

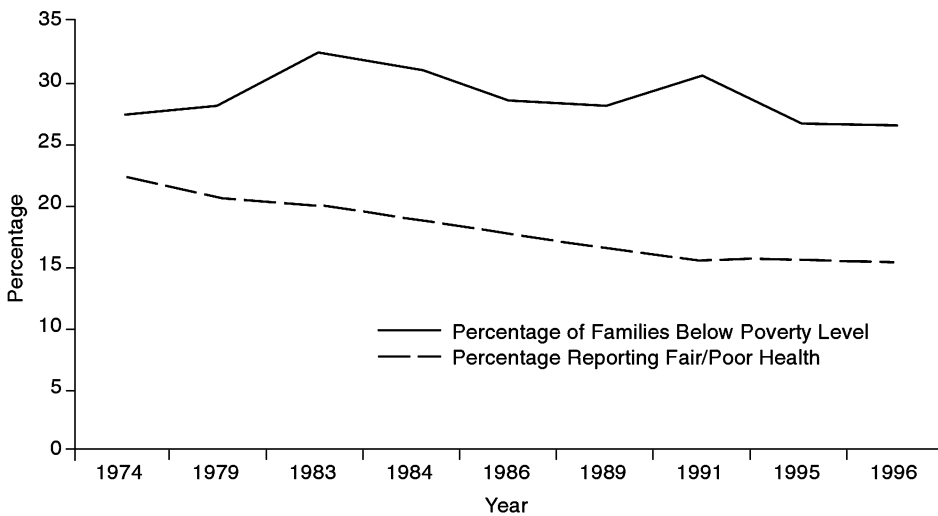
Health status also differs systematically by race, and increasing attention is being paid to racial inequalities in health in the United

States. There is a strong but far from complete overlap between racial and income inequalities in health (Williams 2002; Hayward et al. 2000; Williams and Collins 1995; Sorlie, Backlund, and Keller 1995; Ren and Amick 1996). Although this overlap is often acknowledged, it is not explicitly examined in much research on income and health.



Source: U.S. Department of Commerce (2000) National Center for Health Statistics (1981 to 1999).

Figure 4. Comparing Self-Reported Health and Poverty Status of U.S. Families: White Families, 1974 to 1996



Source: U.S. Department of Commerce (2000) National Center for Health Statistics (1981 to 1999).

Figure 5. Comparing Self-Reported Health and Poverty Status of U.S. Families: African American Families, 1974 to 1996

Figures 4 and 5 present the trend in poverty and in poor and fair health for white and African American respondents. Figure 5 highlights the well-known higher poverty rates among African Americans and the higher proportion of African Americans with poor or fair health relative to the white pop-

ulation. The trends in health are of interest, for they are consistent with a tie between poverty and overall health for whites but not for African Americans. After 1983, the two lines track closely for whites. For African Americans, Figure 5 shows generally improving health, while the poverty rate fluctuates

more. In this case, the figure does not illustrate or suggest a link between the two, although if the trend lines began in 1983, both would show some steady decline and hence improvement in health and reduction of poverty. Clearly, more attention should be paid to measuring trends in the relationship between income, race, and health and mortality over time in the United States.

Some recent research examines not only the relationship between individual or family income and health but also the impact of the neighborhood or community income level on individual health (see reviews in Robert 1999; Bond Huie 2001; Pickett and Pearl 2001; Robert and House 2000b; Diez-Roux 2000, 2001). This research finds evidence that living in communities with a higher proportion of poverty households, or with overall lower income (for example, low median family income), is associated with poor health and mortality over and above the effects of individual or family income in the United States (Haan, Kaplan, and Camacho 1987; Waitzman and Smith 1998a; Robert 1998; Diez-Roux et al. 1997; Anderson et al. 1997; Blakely, Lochner, and Kawachi 2002). Living in poorer communities may be detrimental to the health of all residents, regardless of their own income.

Most of this multilevel research, however, finds that individual- or family-level socioeconomic characteristics are more significant correlates of health than are community variables (Robert 1998; Brooks-Gunn, Duncan, and Aber 1997; Elliott et al. 1996).

Combining multilevel research on the individual and community socioeconomic context and health with research on race, income, and health may be particularly important. For example, because the United States is residentially segregated by race (Jargowsky 1997; Massey and Denton 1993; Wilson 1987), African American and non-Hispanic white people of the same family income level live in very different neighborhoods. In a study by Stephanie Robert and Kum Yi Lee

(2002), the disadvantaged health status of African American older adults that remains after controlling for individual SES is further explained after considering community socioeconomic characteristics. The higher number of chronic conditions experienced by African American older adults is explained only after considering both individual-level and community-level socioeconomic characteristics. Moreover, Robert and Lee demonstrate that African American older adults may even have slightly better self-rated health than non-African American older adults once the individual SES and the community socioeconomic context are controlled.

A number of researchers have also demonstrated the necessity of exploring more explicitly how the community racial context and individual race and ethnicity are related to community and individual income and health (Deaton and Lubotsky 2003; Collins and Williams 1999; Robert and House 2000b; Haan, Kaplan, and Camacho 1987; LeClere, Rogers, and Peters 1997, 1998; Robert 2002; Nazroo 2003; Krieger 2003; Acevedo-Garcia et al. 2003; Miller and Paxson 2001). What role do racial distribution, concentration, and segregation play in determining or explaining community-, family-, and individual-level income inequalities in health? We turn to these questions.

Income Inequality and Health

Whereas the evidence for a relationship between individual income and health is strong and relatively consistent, the evidence for a relationship between aggregate measures of income inequality and health is weak and controversial. There has been much focus in recent years on the latter research, with heated and excited debate about the importance to health and well-being of living in more or less unequal places. Ironically, the focus on this latter, more speculative, research may have unwittingly diverted attention from the former, more consistent, research.

The literature regarding the income inequality hypothesis must refer to Richard Wilkinson, given the influence of his writings on the field. In various writings over the last decade (see, for example, Wilkinson 1992, 1994, 1996), he has provided evidence of a relationship between income inequality in a country and life expectancy, both at a point in time and over time. However, other researchers have raised questions about this evidence, calling into doubt the reliability of the data Wilkinson uses to measure income inequality (Gravelle, Wildman, and Sutton 2002; Mellor and Milyo 2001). Using two more recent datasets considered to be of superior quality, two teams of researchers, Ken Judge, Jo-Ann Mulligan, and Michaela Benzeval (1998) and Hugh Gravelle, John Wildman, and Matthew Sutton (2002), find no significant relationship between income inequality and life expectancy across developed countries.

The idea that income inequality per se may influence health has caught the attention of many researchers, and a growing body of literature tests this hypothesis not only at the level of countries but across regions, states, counties, and cities within nations. For example, using states as the unit of comparison, George Kaplan and his colleagues (1996) and Bruce Kennedy, Ichiro Kawachi, and Deborah Prothrow-Stith (1996) find a significant relationship between several mortality rates and income inequality. Kennedy and his colleagues (1998) also find similar associations between self-rated health and income inequality. Kaplan and his colleagues find a tie between mortality rates and inequality in household income at the state level in the United States, even when median incomes are included. Kennedy and his colleagues (1996) use an alternative index of inequality, the “Robin Hood Index,” also using state data for the United States. In their analysis of a variety of mortality rates, inequality is a statistically

significant predictor. A follow-up study by Kawachi and Kennedy (1997) explores additional measures of inequality and mortality with similar results: a statistically significant link between measures of income inequality at the state level and health as measured by mortality rates. Examining 283 metropolitan areas as units of analysis rather than states, John Lynch, George Kaplan, and Elsie Pamuk (1998) similarly look at the relationships between multiple measures of inequality and mortality. The inequality measures predict mortality over and above other community-level factors, such as per capita income and the proportion of the population with incomes under 200 percent of the poverty level. Using county and tract-level data, Felicia LeClere and Mah-Jabeen Soobader (2000) find independent income inequality effects primarily at the county level, but only for some specific subgroups by age, race, and gender. In these studies, various controls were added for median income, for proportion below the poverty line or near poverty, per capita income, and average household size.

Although most researchers accept this raw association between income inequality and health, even after controlling for median income levels or poverty, debate arises primarily over potential *explanations* for this relationship. Lynch and his colleagues (2000) describe three types of explanations for the association between income inequality and health: the individual income interpretation, the psychosocial environment interpretation, and the neomaterial interpretation. According to the *individual income interpretation*, aggregate associations between income inequality and health simply reflect the nonlinear (convex) relation between income and health at the individual level. According to the *psychosocial environment interpretation*, perceptions of inequality produce instabilities in social capital—perceived mistrust, lack of social cohesion, and so on—that

work through behaviors and psychosocial responses that affect an individual's biological responses and ultimate health. The *neomaterial interpretation* suggests that places with the greatest income inequality are places that also have inequalities in human, physical, health, and social infrastructure that ultimately affect health.

The individual income interpretation suggests that, theoretically, if the relationship between individual income and health is nonlinear, then there should be an aggregate-level association between income inequality and health (Ecob and Davey Smith 1999; Gravelle 1998; Preston 1975; Rodgers 1979). Empirically, this has been tested in several studies that control for individual income while examining the association between income inequality and health. The association between income inequality and health or mortality sometimes remains after controlling for individual SES (Waitzman and Smith 1998b; Soobader and LeClere 1999; Fiscella and Franks 2000; Kennedy et al. 1998; LeClere and Soobader 2000; Lochner et al. 2001), and sometimes it does not (Fiscella and Franks 1997, 2000; Mellor and Milyo 1999; LeClere and Soobader 2000; Blakely, Lochner, and Kawachi 2002; Daly et al. 1998). For example, Tony Blakely, Kimberly Lochner, and Ichiro Kawachi (2002) use individual-level data from the 1996 and 1998 Current Population Survey in conjunction with 1990 Gini coefficient measures at the metropolitan area (MA) levels to look at self-rated health. They find no significant association between MA-level income inequality and health self-rated as poor or fair after controlling for sex, age, race, individual-level household income, and MA-level household income. In a recent review, Subramanian, Blakely, and Kawachi (2003) suggest that the only consistent associations between income inequality and health in multilevel studies may be seen at the state level in the United States, rather than at ei-

ther smaller levels of aggregation within the United States or between or within other countries.

It seems crucial to explore further the extent to which the relationship between measures of inequality at the aggregate level and health reflects the nonlinear association between individual income and health. However, even if the aggregate relationship were fully explained by the individual-level relationship, the interpretation of this finding would still be open for debate. Should the aggregate relationship then be seen as an artifact of the individual-level data, and therefore not meaningful, or does income inequality at more macro levels play a more causal role in producing or reproducing income inequalities in health?

Much of the heated debate in the literature on income inequality and health seems to stem from reactions to the so-called psychosocial environment interpretation. A number of researchers have explored the possibility that inequality per se produces a social environment that ultimately affects an individual's health (see, for example, Wilkinson 1996). For example, Kawachi and his colleagues (1997) suggest that high income inequality within states is associated with low social cohesion and disinvestment in social capital, which then affect mortality. Examples of this would be the relative income and deprivation hypotheses, which suggest that an individual's perceived relative income or position may affect health through psychosocial pathways (Adler et al. 1994).

Proponents of the neomaterial interpretation believe that more consideration should be given to factors that exist outside of, or as precursors to, psychosocial pathways. For example, Lynch and his colleagues (2000) point out that there are structural, political, and economic processes that generate inequality, and that these processes exist before individuals experience or perceive their effects. Therefore, focusing on the structural,

political, and economic processes that both produce and result from inequality may be more appropriate than attending primarily to individual perceptions or experiences that occur downstream from these processes.

Examples of research exploring neomaterial interpretations of the link between income inequality and health include research that explores the potential intermediating role of public investment in, for instance, health care, public safety, education, and environmental quality. To the extent that such investments are concentrated in higher-income areas, or in less unequal areas, then an individual living in such areas gains more access to such goods and services (Kaplan et al. 1996; Lynch, Kaplan, and Pamuk 1998; Davey Smith 1996). The neomaterial interpretation implies that greater income inequality in society leads to greater differences in resources across communities that are associated with greater differences in health across communities. This finding highlights deprivation, but in this view deprivation at the individual level is compounded by deprivation at the community level as socioeconomic segregation and greater income inequality make the indirect costs of acquisition of health care and a healthy lifestyle more expensive to those with the lowest incomes. This suggests interaction effects between income inequality and individual income level. A number of studies suggest that income inequality may be a significant determinant of health, but only for those with the lowest incomes (Mellor and Milyo 1999; Soobader and LeClere 1999).

Critiquing the Literature on Income Inequality and Health

To summarize our primary understanding of the literature on income, income inequality, and health: there is a strong relationship between individual income and

health that persists, and perhaps strengthens, over time. The shape of the relationship between income and health appears to be nonlinear, such that increases in income among those at the low end of the distribution should have greater positive consequences for their health than similar increases in income would have for the health of those at the top end of the income distribution. Much attention has been paid recently to the possibility that income inequality per se is related to health. The results of this research are still tenuous, though deserving of further study. However, it appears that the focus on this more tenuous research has unwittingly diverted attention from the former research on individual income and health. We conclude that (1) more research is needed to examine the relationship between income inequality and health, although (2) the greater advances in our understanding of the relationship between income and health will come from research that refocuses attention on the relationship between individual income level and health. Nevertheless, even if the primary relationship between income and health should be examined at the individual level, mediators of this relationship may very well operate at more aggregate levels such as households or neighborhoods.

REFERENCES

- Acevedo-Garcia, Dolores, Kimberly A. Lochner, Theresa L. Osypuk, and Sukanya V. Subramanian. 2003. "Future Directions in Residential Segregation and Health Research: A Multilevel Approach." *American Journal of Public Health* 93(2): 215–21.
- Adler, Nancy E., Thomas Boyce, Margaret A. Chesney, Sheldon Cohen, Susan Folkman, Robert L. Kahn, and S. Leonard Syme. 1994. "Socioeconomic Status and Health: The Challenge of the Gradient." *American Psychologist* 49(1): 15–24.
- Adler, Nancy E., W. Thomas Boyce, Margaret A. Chesney, Susan Folkman, and S. Leonard Syme. 1993. "Socioeconomic Inequalities in Health: No

- Easy Solution." *Journal of the American Medical Association* 269: 3140–45.
- Adler, Nancy E., and Katherine Newman. 2002. "Socioeconomic Disparities in Health: Pathways and Policies." *Health Affairs* 21(2): 60–76.
- Anderson, Roger T., Paul Sorlie, Eric Backlund, Norman Johnson, and George A. Kaplan. 1997. "Mortality Effects of Community Socioeconomic Status." *Epidemiology* 8: 42–47.
- Antonovsky, Aaron. 1967. "Social Class, Life Expectancy, and Overall Mortality." *Milbank Memorial Fund Quarterly* 45: 31–73.
- Backlund, Eric, Paul D. Sorlie, and Norman J. Johnson. 1996. "The Shape of the Relationship Between Income and Mortality in the United States: Evidence from the National Longitudinal Mortality Study." *Annals of Epidemiology* 6(1): 12–20.
- Black, Douglas, J. N. Morris, Cyril Smith, and Peter Townsend. 1982. *Inequalities in Health: The Black Report*. New York: Penguin.
- Blakely, Tony A., Kimberly Lochner, and Ichiro Kawachi. 2002. "Metropolitan-Area Income Inequality and Self-Rated Health: A Multilevel Study." *Social Science and Medicine* 54(1): 65–77.
- Bond Huie, Stephanie A. 2001. "The Concept of Neighborhood in Health and Mortality Research." *Sociological Spectrum* 21: 341–58.
- Brooks-Gunn, Jeanne, Greg J. Duncan, and J. Lawrence Aber, eds. 1997. *Neighborhood Poverty: Context and Consequences for Children*, vol. 1. New York: Russell Sage Foundation.
- Collins, Chiquita A., and David R. Williams. 1999. "Segregation and Mortality: The Deadly Effects of Racism?" *Sociological Forum* 14(3): 495–523.
- Daly, Mary C., Greg J. Duncan, George A. Kaplan, and John W. Lynch. 1998. "Macro-to-Micro Links in the Relation Between Income Inequality and Mortality." *Milbank Memorial Fund Quarterly* 76(3): 303–4, 315–39.
- Davey Smith, George. 1996. "Income Inequality and Mortality: Why Are They Related?" *British Medical Journal* 313: 987–88.
- Deaton, Angus. 2001. "Health, Inequality, and Economic Development." Working paper 8318. Cambridge, Mass.: National Bureau of Economic Research (June).
- Deaton, Angus, and Darren Lubotsky. 2003. "Mortality, Inequality, and Race in American Cities and States." *Social Science and Medicine* 56(6): 1139–53.
- Diez-Roux, Ana V. 2000. "Multilevel Analysis in Public Health Research." *Annual Review of Public Health* 21: 171–92.
- _____. 2001. "Investigating Neighborhood and Area Effects on Health." *American Journal of Public Health* 91: 1783–89.
- Diez-Roux, Ana V., F. Javier Nieto, Carles Muntaner, Herman A. Tyroler, George W. Comstock, Elad Shahar, Leslie S. Cooper, Robert L. Watson, and Moyses Szklo. 1997. "Neighborhood Environments and Coronary Heart Disease: A Multilevel Analysis." *American Journal of Epidemiology* 146(1): 48–63.
- Ecob, Russell, and George Davey Smith. 1999. "Income and Health: What Is the Nature of the Relationship?" *Social Science and Medicine* 48(5): 693–705.
- Elliott, Delbert S., William J. Wilson, David Huizinga, Robert J. Sampson, Amanda Elliott, and Bruce Rankin. 1996. "The Effects of Neighborhood Disadvantage on Adolescent Development." *Journal of Research in Crime and Delinquency* 33: 389–426.
- Ettner, Susan L. 1996. "New Evidence on the Relationship Between Income and Health." *Journal of Health Economics* 15(1): 67–86.
- Evans, Robert G., Morris L. Barer, and Theodore R. Marmor, eds. 1994. *Why Are Some People Healthy and Others Not? The Determinants of Health of Populations*. New York: Aldine de Gruyter.
- Feinstein, Jonathan S. 1993. "The Relationship Between Socioeconomic Status and Health: A Review of the Literature." *Milbank Memorial Fund Quarterly* 71: 279–322.
- Fiscella, Kevin, and Peter Franks. 2000. "Individual Income, Income Inequality, Health, and Mortality: What Are the Relationships?" *Health Services Research* 35(1): 307–18.
- Gravelle, Hugh. 1998. "How Much of the Relationship Between Population Mortality and Unequal Distribution of Income Is a Statistical Artifact?" *British Medical Journal* 316: 382–85.
- Gravelle, Hugh, John Wildman, and Matthew Sutton. 2002. "Income, Income Inequality, and Health: What Can We Learn from the Aggregate Data?" *Social Science and Medicine* 54: 577–89.
- Haan, Mary, George A. Kaplan, and Terry Camacho. 1987. "Poverty and Health: Prospective Evidence from the Alameda County Study." *American Journal of Epidemiology* 125(6): 989–98.
- Hayward, Mark D., Toni P. Miles, Eileen M. Crimmins, and Yu Yang. 2000. "The Significance of Socioeconomic Status in Explaining the Racial Gap in Chronic Health Conditions." *American Sociological Review* 65: 910–30.

- House, James S., Ronald C. Kessler, A. Regula Herzog, Richard P. Mero, Ann M. Kinney, and Martha J. Breslow. 1990. "Age, Socioeconomic Status, and Health." *Milbank Memorial Fund Quarterly* 68(3): 383–411.
- House, James S., James M. Lepkowski, Ann M. Kinney, Richard P. Mero, Ronald C. Kessler, and A. Regula Herzog. 1994. "The Social Stratification of Aging and Health." *Journal of Health and Social Behavior* 35: 213–34.
- Jargowsky, Paul A. 1997. *Poverty and Place: Ghettos, Barrios, and the American City*. New York: Russell Sage Foundation.
- Judge, Ken, Jo-Ann Mulligan, and Michaela Benzeval. 1998. "Income Inequality and Population Health." *Social Science and Medicine* 46(4–5): 567–79.
- Kaplan, George A., Elsie R. Pamuk, John W. Lynch, Richard D. Cohen, and Jennifer L. Balfour. 1996. "Inequality in Income and Mortality in the United States: Analysis of Mortality and Potential Pathways." *British Medical Journal* 312: 999–1003.
- Kawachi, Ichiro, and Bruce P. Kennedy. 1997. "The Relationship of Income Inequality to Mortality: Does the Choice of Indicator Matter?" *Social Science and Medicine* 45: 1121–27.
- Kawachi, Ichiro, Bruce P. Kennedy, Kimberley Lochner, and Deborah Prothrow-Stith. 1997. "Social Capital, Income Inequality, and Mortality." *American Journal of Public Health* 87(9): 1491–98.
- Kennedy, Bruce P., Ichiro Kawachi, Roberta Glass, and Deborah Prothrow-Stith. 1998. "Income Distribution, Socioeconomic Status, and Self-rated Health in the United States: Multilevel Analysis." *British Medical Journal* 317: 917–21.
- Kennedy, Bruce P., Ichiro Kawachi, and Deborah Prothrow-Stith. 1996. "Income Distribution and Mortality: Cross-sectional Ecological Study of the Robin Hood Index in the United States." *British Medical Journal* 312: 1004–7.
- Krieger, Nancy. 2003. "Does Racism Harm Health? Did Child Abuse Exist Before 1962? On Explicit Questions, Critical Science, and Current Controversies: An Ecosocial Perspective." *American Journal of Public Health* 93(2): 194–99.
- LeClere, Felicia B., and Mah-Jabeen Soobader. 2000. "The Effect of Income Inequality on the Health of Selected U.S. Demographic Groups." *American Journal of Public Health* 90(12): 1892–97.
- Lynch, John W., George A. Kaplan, and Elsie R. Pamuk. 1998. "Income Inequality and Mortality in Metropolitan Areas of the United States." *American Journal of Public Health* 88: 1074–80.
- Lynch, John W., George Davey Smith, George A. Kaplan, and James S. House. 2000. "Income Inequality and Mortality: Importance to Health of Individual Income, Psychosocial Environment, or Material Conditions." *British Medical Journal* 320(7243): 1200–4.
- Marmot, Michael G., Manolis Kogevinas, and Mary Ann Elston. 1987. "Social/Economic Status and Disease." *Annual Review of Public Health* 8: 111–35.
- Marmot, Michael G., G. Davey Smith, Stephen A. Stansfeld, C. Patel, F. North, J. Head, I. White, E. J. Brunner, and A. Feeney. 1991. "Health Inequalities Among British Civil Servants: The Whitehall II Study." *Lancet* 337(8754):1387–93.
- Massey, Douglas S. 1990. "American Apartheid: Segregation and the Making of the Underclass." *American Journal of Sociology* 96: 329–58.
- Massey, Douglas S., and Nancy A. Denton. 1993. *American Apartheid: Segregation and the Making of the Underclass*. Cambridge, Mass.: Harvard University Press.
- McDonough, Peggy, Greg J. Duncan, David Williams, and James House. 1997. "Income Dynamics and Adult Mortality in the United States, 1972 Through 1989." *American Journal of Public Health* 87(9): 1476–83.
- Mellor, Jennifer M., and Jeffrey D. Milyo. 1999. "Income Inequality and Individual Health: Evidence from the Current Population Survey." Working paper 8. Boston: Robert Wood Johnson Foundation Scholars in Health Policy Research Program.
- _____. 2001. "Reexamining the Evidence of an Ecological Association Between Income Inequality and Health." *Journal of Health Politics, Policy, and Law* 26(3): 487–522.
- Miller, Douglas, and Christina Paxson. 2001. "Relative Income, Race, and Mortality." Working paper. Princeton, N.J.: Princeton University, Center for Health and Wellbeing.
- Mirowsky, John, and Paul Nongzhuang Hu. 1996. "Physical Impairment and the Diminishing Effects of Income." *Social Forces* 74: 1073–96.
- National Center for Health Statistics. Various years. *Health, United States*. Washington: U.S. Government Printing Office.
- Nazroo, James Y. 2003. "The Structuring of Ethnic Inequalities in Health: Economic Position, Racial Discrimination, and Racism." *American Journal of Public Health* 93(2): 277–84.

- Pappas, Gregory, Susan Queen, Wilbur Hadden, and Gail Fisher. 1993. "The Increasing Disparity in Mortality Between Socioeconomic Groups in the United States, 1960 and 1986." *New England Journal of Medicine* 329: 103–9.
- Pickett, Kate E., and Michelle Pearl. 2001. "Multi-level Analyses of Neighborhood Socioeconomic Context and Health Outcomes: A Critical Review." *Journal of Epidemiology and Community Health* 55: 111–22.
- Preston, Samuel. 1975. "The Changing Relation Between Mortality and Level of Development." *Population Studies* 29: 231–48.
- Ren, Xinhua S., and Benjamin Amick. 1996. "Race and Self-assessed Health Status: The Role of Socioeconomic Factors in the United States." *Journal of Epidemiology and Community Health* 50: 269–73.
- Robert, Stephanie A. 1998. "Community-Level Socioeconomic Status Effects on Adult Health." *Journal of Health and Social Behavior* 39: 18–37.
- _____. 1999. "Socioeconomic Position and Health: The Independent Contribution of Community Context." *Annual Review of Sociology* 25: 489–516.
- Robert, Stephanie A., and James S. House. 2000a. "Socioeconomic Inequalities in Health: An Enduring Sociological Problem." In *Handbook of Medical Sociology*, edited by Chloe E. Bird, Peter Conrad, and Allen M. Fremont. New York: Prentice-Hall.
- _____. 2000b. "Socioeconomic Inequalities in Health: Integrating Individual-, Community-, and Societal-Level Theory and Research." In *Handbook of Social Studies in Health and Medicine*, edited by Gary L. Albrecht, Ray Fitzpatrick, and Susan C. Scrimshaw. London: Sage Publications.
- Robert, Stephanie A., and Kum Yi Lee. 2002. "Explaining Race Differences in Health Among Older Adults: The Contribution of Community Socioeconomic Context." *Research on Aging* 24(6): 654–83.
- Rodgers, Gerry B. 1979. "Income and Inequality as Determinants of Mortality: An International Cross-section Analysis." *Population Studies* 33: 343–51.
- Soobader, Mah-Jabeen, and Felicia B. LeClere. 1999. "Aggregation and the Measurement of Income Inequality: Effects on Morbidity." *Social Science and Medicine* 48: 733–44.
- Sorlie, Paul D., Eric Backlund, and Jacob B. Keller. 1995. "U.S. Mortality by Economic, Demographic, and Social Characteristics: The National Longitudinal Mortality Study." *American Journal of Public Health* 85: 949–56.
- Subramanian, Sukanya V., Tony Blakely, and Ichiro Kawachi. 2003. "Income Inequality as a Public Health Concern: Where Do We Stand?" *Health Services Research* 38: 153–67.
- U.S. Department of Commerce. U.S. Census Bureau. 2000. *Statistical Abstract of the United States: 2000*. Washington: U.S. Government Printing Office.
- Wagstaff, Adam, and Eddy van Doorslaer. 2000. "Income Inequality and Health: What Does the Literature Tell Us?" *Annual Review of Public Health* 21: 543–67.
- Waitzman, Norman J., and Ken R. Smith. 1998a. "Phantom of the Area: Poverty-Area Residence and Mortality in the United States." *American Journal of Public Health* 88(6): 973–76.
- _____. 1998b. "Separate but Lethal: The Effects of Economic Segregation on Mortality in Metropolitan America." *Milbank Memorial Fund Quarterly* 76(3): 341–73.
- Wilkinson, Richard G. 1992. "Income Distribution and Life Expectancy." *British Medical Journal* 304: 165–68.
- _____. 1994. "The Epidemiological Transition: From Material Scarcity to Social Disadvantage?" *Daedalus* 123: 61–77.
- _____. 1996. *Unhealthy Societies: The Affliction of Inequality*. London: Routledge.
- Williams, David R. 2002. "Racial/Ethnic Variations in Women's Health: The Social Embeddedness of Health." *American Journal of Public Health* 92(4): 588–97.
- Williams, David R., and Chiquita Collins. 1995. "U.S. Socioeconomic and Racial Differences in Health: Patterns and Explanations." *Annual Review of Sociology* 21: 349–86.
- Wilson, William Julius. 1987. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press.

Culture

96 ■ Tak Wing Chan and John H. Goldthorpe

The Social Stratification of Theatre, Dance, and Cinema Attendance

In current sociological literature the relationship between social inequality and patterns of cultural taste and consumption is the subject of a large and complex debate. This paper starts by outlining the leading positions that have been taken up in this debate, and then presents some illustrative findings from a research programme in which the authors are currently engaged. These findings are used chiefly as the basis for a critical evaluation of the rival positions that have been set out, although they may also be of interest in their own right. In addition, attention is drawn to the methodology on which the paper rests. Although this may appear somewhat daunting, at least to readers who lack a statistical background, it is, we would argue, only by following such an approach that the issues with which we are concerned can be adequately addressed.

In the sociological literature referred to, it is possible to identify three main lines of argument. In their essentials, these arguments can be stated as follows, although each has variant forms.

1. *The homology argument*: this claims that social stratification—that is, the prevail-

ing structure of inequality within a society—and cultural stratification map onto each other very closely. Individuals in higher social strata are those who prefer and predominantly consume ‘high’ or ‘elite’ culture, and individuals in lower social strata are those who prefer and predominantly consume ‘popular’ or ‘mass’ culture—with, usually, various intermediate situations being also recognized. In some versions of the argument (e.g. Bourdieu, 1984) it is further claimed that the arrogation of ‘distinction’ in cultural taste and, conversely, processes of ‘aesthetic distancing’ are actively used by members of dominant social classes as means of symbolically demonstrating and confirming their superiority.

2. *The individualization argument*: this seeks in effect to relegate the homology argument to the past. In modern, relatively affluent and highly commercialized societies, it is held that differences in cultural taste and consumption are rapidly losing any clear grounding in social stratification: age, gender, ethnicity or sexuality, for example, all can, and do, serve as alternative social bases

of cultural differentiation. And in more radical forms of the argument (e.g. Bauman, 1988; Featherstone, 1987), emphasis is placed on the growing ability of individuals to free themselves from social conditioning and influence of any kind and to choose and form their own distinctive identities and lifestyles—patterns of cultural consumption included.

3. *The omnivore–univore argument*: this in effect challenges both the homology and individualization arguments (see especially Peterson & Kern, 1996; Peterson & Simkus, 1992). As against the latter, it sees cultural differentiation as still mapping closely onto social stratification; but, as against the former, it does not see this mapping as being on ‘elite-to-mass’ lines. Rather, it claims that the cultural consumption of individuals in higher social strata differs from that of individuals in lower social strata in that it is *greater and much wider in its range*. It comprises not only more ‘high-brow’ culture, but also more ‘middle-brow’ and more ‘low-brow’ culture as well, while the consumption of individuals in lower social strata tends to be largely restricted to more popular cultural forms. Thus, the crucial distinction is not between elite and mass but rather between cultural omnivores and cultural univores.

In other work (Chan & Goldthorpe, 2005), we seek to evaluate these three arguments by analysing data on musical consumption. Far stronger, even if still somewhat qualified, support is found for the omnivore–univore argument than for either the homology or individualization argument, as explained further below. However, it has to be noted that the omnivore–univore argument was initially developed with specific reference to musical consumption (Peterson & Simkus, 1992); and while, in this context, it has in

fact received support additional to that of the present authors (see e.g. Coulangeon, 2003; Van Eijck, 2001), its wider application has been rather little studied. The present paper, therefore, aims to examine how well it fares in another domain of cultural consumption, that of theatre, dance and cinema. This differs from music in several ways, but perhaps most obviously for present purposes in that consumption generally requires attendance at some venue and, with the main exception of video film (which is discounted in what follows), there is no equivalent to the home consumption via various media that is of major importance in the domain of music.

Data and Concepts

As in the authors’ work on musical consumption, data is used from the Arts in England survey which was carried out by the Office for National Statistics on behalf of Arts Council England. The survey was based on a stratified probability sample of individuals aged over 16 and living in private households in England in 2001. In all, 6,042 interviews were completed, representing a response rate of 64 per cent (for further details, see Skelton *et al.*, 2002). However, here attention is restricted to respondents aged 20 to 64. With this limitation, and after deleting all cases with missing values on variables of interest here, there is an effective sample of $N = 3,819$.

As regards theatre and cinema attendance, attention is concentrated on the results obtained in the survey from six questions. Respondents were asked whether or not, in the last 12 months, they had attended a performance of a play/drama, a musical, a pantomime, a ballet, some other form of dance (including contemporary dance, African People’s dance and South Asian dance), or had seen a film at a cinema (or other venue rather than at home). It should be stressed that interest in these results lies not primarily in the answers given to each question taken separately, but rather in the possibility that the

answers to the six questions *taken together* can reveal *patterns* of theatre and cinema consumption and in turn serve to identify *types of consumer* in this domain. As will be seen, this interest is reflected in the way in which the data are analysed.

As regards social stratification, we believe it important to do more than treat this through some single classification or scale of an essentially *ad hoc* kind—such as, say, the Market Research Society categories (AB, C1, C2, D) or the old Registrar General's Social Classes. Therefore information collected in the Arts in England survey on respondents' employment and occupation is drawn on in order to allocate them by both *social class* and *social status*, which are viewed, following a long-established tradition in sociology, as *conceptually separate forms of stratification*.

The class structure overall, and likewise individuals' positions within it, are seen as being defined in a quite objective way by economic relations or, more precisely, by relations in labour markets and production units. To allocate individuals by class, the new National Statistics Socio-Economic Classification (NS-SEC) is used—in its seven-category version—which is specifically designed to capture differences in employment relations (Rose & Pevalin, 2003). The categories of the classification are shown in Table 4. In contrast, the status order is seen as reflecting inter-subjective assessments of individuals' social superiority, equality and inferiority as expressed most directly in relations of social intimacy. Such relations, where present among members of different social groupings, imply a basic equality of status and, where absent, a recognition of inequality. To allocate individuals by status, a 31-category occupationally based scale developed by the authors from analyses of patterns of close friendship in contemporary British society is used (Chan & Goldthorpe, 2004). The closer together any two categories in the scale are, the more similar, occupationally, are their members' friends; the

further apart they are, the less similar are their members' friends.

In any society, the positions of individuals within the class structure and the status order will tend to be correlated—but not perfectly so. Instances of discrepancy between class and status position are always likely to occur. For example, the authors' own results indicate that in present-day Britain salaried professionals and associate professionals tend to have higher status than do salaried managers, and especially managers in manufacturing, construction or transport, despite holding similar class positions as defined in terms of employment relations; or again that routine wage workers in services, especially personal services, tend to have higher status than even skilled manual workers. Since class and status are only imperfectly correlated, it is possible to ask whether it is the one or the other that exerts the greater influence on individuals' experience and action across different areas of social life. There is, for example, evidence that class is the dominant influence so far as individuals' economic life-chances are concerned—i.e. in determining their degree of economic security and their prospects—and also in shaping their political orientations and affiliations. But, in contrast, the expectation would be that in regard to cultural consumption it is status that will carry the greater weight. This is because differences in status are typically expressed in lifestyles, and cultural consumption is one important aspect of lifestyle through which status 'markers' can be readily laid down. In order to test whether this expectation holds good, it is of course essential that one should be able to distinguish class and status, conceptually and operationally.

In addition to treating class and status separately, we also draw on information available from the Arts in England Survey on respondents' incomes and on their educational qualifications. The latter are coded to the six National Vocational Qualification levels that range from 'no qualifications' to 'degree-level

qualification or higher'. In sociological analyses of cultural consumption, income and education are often taken as substitutes or proxies for more direct measures of class or status of the kind to be used in the present study. However, income and education are here considered *along with* such measures of class and status, so that their independent effects, if any, can be established.

Finally, also included in the present analyses is socio-demographic information collected in the Arts in England survey, in particular regarding respondents' sex, age, marital status, family composition and region of residence. Given that the primary concern is with the social stratification of theatre, dance and cinema attendance, these socio-demographic variables are intended to serve primarily as 'controls': that is to say, they are brought into analyses chiefly in order to remove the possibility of any hidden confounding of their effects with those of class, status, income and education on which present interest centres.

Analytical Strategy

In Table 1 we show the percentage of respondents who in the last year had attended a theatre, for performances of the kinds previously indicated, or a cinema. As can be seen, there is some wide variation in the probabilities of attendance, although much on lines that might be expected. Cinema attendance is by far the most frequently reported, while, at the other extreme, going to the theatre for a ballet performance is at a very low level.

As earlier remarked, we wish to treat the data in question primarily as basis for obtaining an understanding of individuals' patterns of theatre and cinema attendance and of the different types of consumer in this cultural domain. To this end, a statistical technique known as latent class analysis is employed, which can be intuitively understood as follows.

There are six questions that serve as indicators of theatre or cinema attendance. Since each question has a two-option (yes/no) an-

Table 1. Percentage of Respondents Who Have Visited a Cinema or a Theatre for Various Kinds of Performance in the Past Twelve Months

| | |
|-------------|------|
| Ballet | 1.9 |
| Other dance | 12.7 |
| Pantomime | 14.6 |
| Musical | 25.4 |
| Play/drama | 29.0 |
| Cinema | 62.7 |

swer, there are in fact 2^6 or 64 different possible response sets. The overall pattern of individuals' responses will therefore be complex. But the answers given by respondents to the six questions can be expected to show some degree of association. Thus, for example, those who say that they have been to a play are also likely to report having been to a cinema. Conversely, those who say that they have not been to a ballet are also likely to report that they have not been to other dance events, and so on. What the technique of latent class analysis aims to do is to simplify matters by exploiting this association among the six indicators. It seeks to identify a limited number of discrete classes, or categories, of respondents such that, *conditional on their belonging to one or other of these classes*, individuals' responses on the indicator items become independent of each other—i.e. there is no longer any association between them. Insofar as this can be done, it can be said that it is individuals' membership of the latent classes that is the *source* of the association initially found among their responses, and each latent class can be taken as representing a quite distinctive pattern of response.

Results

Latent Class Analysis of Theatre and Cinema Attendance

It turns out in fact that, as is shown in Table 2, a very simple latent class solution can be obtained for our data on theatre and cinema attendance. With the minor technical

Table 2. Latent Class Models Fitted to Data on Cultural Participation in the Domain of Theatre, Dance and Cinema

| <i>Model</i> | <i>Number of classes</i> | <i>G²</i> | <i>df</i> | <i>p</i> |
|--------------|--------------------------|----------------------|-----------|----------|
| 1 | 1 | 1583.64 | 57 | 0.00 |
| 2 | 2 | 268.16 | 50 | 0.00 |
| 3 | 2 ^a | 53.22 | 49 | 0.31 |

^aA local dependence term is included in this model to allow for an association between attendance at ballet and other dance events.

Table 3. Estimated Size of the Latent Classes and Conditional Probabilities (Percent) of Attendance Under Our Preferred Model

| | <i>Latent class</i> | |
|------------------------------------|---------------------|------|
| | 1 | 2 |
| Relative size (%), initial | 62.5 | 37.5 |
| Relative size (%), post-assignment | 64.2 | 35.8 |
| Probabilities of attendance (%) | | |
| Ballet | 0.1 | 5.0 |
| Other dance | 5.6 | 24.6 |
| Pantomime | 6.7 | 27.9 |
| Musical | 6.9 | 56.2 |
| Play/drama | 6.1 | 67.1 |
| Cinema | 48.0 | 87.1 |

modification that is noted in Table 2, a model proposing just two latent classes fits the data satisfactorily: that is, just two latent classes prove sufficient to capture virtually all of the association that exists among responses on the six indicator items. Or, one could say, it emerges that underlying the results previously reported in Table 1 on these six different kinds of attendance, a clear, essentially dichotomous, patterning prevails.

On the basis of this solution we can then go on to assign each individual in the sample to one or other of the two latent classes that are identified—that is, to whichever he or she has the highest probability of belonging to, given his or her own set of responses on the six indicator items; in this way the respondents are divided into two types of consumer of theatre, dance and cinema. In Table 3 it is shown, first of all, that this process of assignment does not result in any major change in the relative size of the latent classes from that initially estimated under the

model; or, in other words, no great degree of uncertainty appears to arise about the latent class with which particular respondents should be affiliated. Second, the probabilities of individuals reporting each of the six different kinds of attendance considered are shown, given their latent class membership.

What, then, can be discovered about the two types of consumer that are derived from the latent class analysis? Our findings are in fact rather clear cut. As can be seen, latent class 1, which accounts for almost two-thirds of the sample, comprises individuals who have a very low probability—less than 10 per cent—of having attended a theatre in the year before the interview for any of the kinds of performance distinguished, and whose consumption is effectively limited to a fairly modest—48 per cent—probability of having visited a cinema. Latent class 2, in contrast, which accounts for somewhat over one-third of the sample, comprises individuals who have a relatively high probability (i.e. as com-

pared to the overall rates shown in Table 1) of having attended a theatre for each of the kinds of performance covered and of having been to the cinema as well.

These findings would then, so far as they go, appear highly consistent with the omnivore–univore argument initially referred to. The latent Class 2 represents the theatre and cinema omnivores, and latent Class 1 the univores, whose consumption is in fact more or less restricted to the cinema. Certainly, we find no evidence of the kind that might be expected from the homology argument of a cultural elite who, in pursuit of ‘distinction’, attend the theatre for, say, drama and ballet performances but who at the same time display ‘aesthetic distancing’ in shunning musicals and pantomime. Members of latent Class 2 have the highest probability of attendance at not only drama and ballet but also all other kinds of theatre performance covered and the cinema. Furthermore, the very fact that the sample divides so readily into just two types of consumer is in itself sufficient to throw serious doubt on the individualization argument. There is no evidence here of the kind of individual diversity in cultural consumption that would, were it present, effectively defy latent class analysis or at all events require that an unmanageably large number of latent classes be distinguished, and ones to which individuals could be assigned only with great uncertainty.

Theatre, Dance and Cinema Attendance and Social Stratification

As already noted, in our work on musical consumption (Chan & Goldthorpe, 2005), we also find, with some qualification, support for the omnivore–univore argument. Our latent class analyses in this case point in fact to three types of musical consumer: univores, whose consumption is largely restricted to pop and rock, and then two kinds of omnivore—‘true’ omnivores and omnivore listeners. The former have a high probability both of attending musical events and

of listening to music across all the genres distinguished, while the latter are omnivorous only in their listening to broadcast or recorded music. Further analysis then reveals that the chances of being an omnivore, and especially a true omnivore, rather than a univore, increase with status, although—following the expectations earlier mentioned—the effects of class are negligible once status is included in the analysis. In addition, it is shown that even when the effects of status (and class) are controlled, the chances of being a musical omnivore rather than a univore still increase fairly steadily with level of educational qualifications, but that, in contrast, these chances do not appear to be affected by income when other stratification variables are controlled. How far, then, are similar results obtained in regard to cultural consumption in the form of theatre and cinema attendance?

To begin with, we may examine the simple two-way relationships that exist between the chances of being in this regard an omnivore or a cinema-only univore (according to the previous analyses) and class and status respectively. In Table 4 we show the distribution of univores (latent Class 1) and omnivores (latent Class 2) within the seven classes of NS-SEC. It is evident that omnivores are most common in the professional and managerial classes, 1 and 2, where they are in fact in a slight majority, while univores dominate in Classes 5, 6 and 7, those of lower supervisory and technical, semi-routine and routine workers.

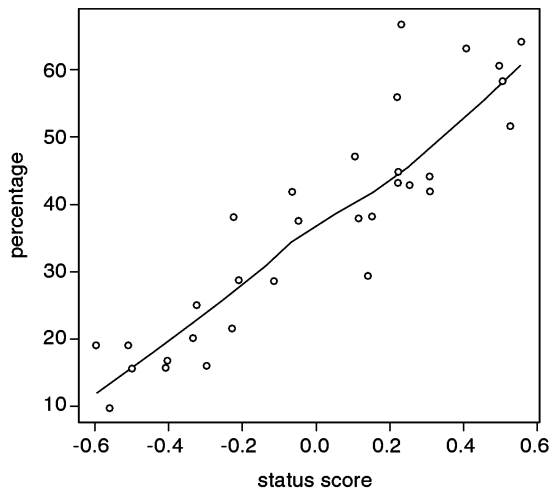
Figure 1 is then analogous to Table 4, but with the 31 categories of our status scale replacing the seven NS-SEC classes. An obvious ‘status gradient’ exists in the chances of being an omnivore rather than a univore.

These results are then consistent with the general idea that it is members of higher social strata who are more likely to be culturally omnivorous, and members of lower strata who are more likely to be univorous. But to test this idea more rigorously against

Table 4. Distribution of Univores (U) and Omnivores (O) Within NS Social Classes

| <i>NS social class</i> | <i>U (%)</i> | <i>O (%)</i> | <i>N</i> |
|--|--------------|--------------|----------|
| 1 Higher managerial and professional occupations | 43.9 | 56.2 | 488 |
| 2 Lower managerial and professional occupations | 49.4 | 50.6 | 1023 |
| 3 Intermediate occupations | 63.2 | 36.8 | 574 |
| 4 Small employers and own-account workers | 72.7 | 27.3 | 275 |
| 5 Lower supervisory and technical occupations | 77.7 | 22.3 | 359 |
| 6 Semi-routine occupations | 77.1 | 22.9 | 620 |
| 7 Routine occupations | 85.8 | 14.2 | 480 |
| | 64.2 | 35.8 | 3819 |

(a) Theatre and Cinema Omnivores



(b) Theatre and Cinema Univores

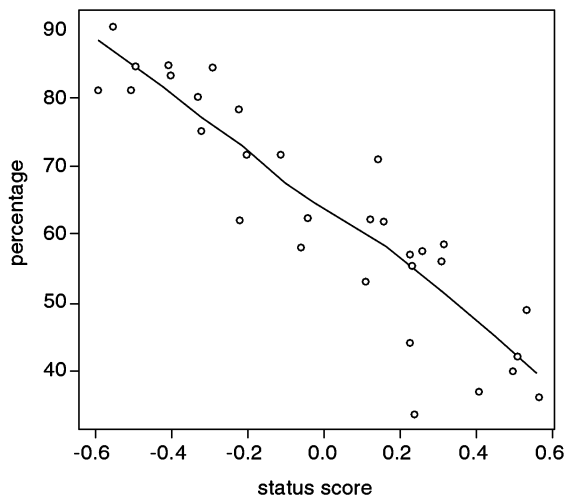


Figure 1. The Proportion of Respondents Being Univores and Omnivores Within Status Categories by Status Score

Table 5. Coefficients from Binary Logistic Regression Analysis for Effects of Covariates on the Probability of Being an Omnivore Rather than a Univore

| | $\hat{\beta}$ | SE |
|--------------------------|---------------|-------|
| Female ^a | 0.615** | 0.092 |
| Married ^b | 0.148 | 0.112 |
| Separated | 0.188 | 0.139 |
| Age ^c | 0.005 | 0.004 |
| Child (0–4) ^d | –0.562** | 0.113 |
| Child (5–10) | 0.070 | 0.100 |
| Child (11–15) | 0.088 | 0.105 |
| The North ^e | –0.231 | 0.124 |
| Midlands | –0.207 | 0.123 |
| South East | 0.083 | 0.135 |
| South West | –0.189 | 0.153 |
| Income | 0.026** | 0.005 |
| CSE/others ^f | 0.169 | 0.152 |
| O levels | 0.668** | 0.128 |
| A levels | 1.130** | 0.145 |
| Sub-degree | 1.027** | 0.160 |
| Degree | 1.223** | 0.151 |
| Class 2 ^g | 0.078 | 0.126 |
| Class 3 | –0.161 | 0.160 |
| Class 4 | –0.205 | 0.203 |
| Class 5 | –0.134 | 0.218 |
| Class 6 | –0.199 | 0.195 |
| Class 7 | –0.507* | 0.230 |
| Status | 0.631** | 0.179 |
| Constant | –2.118** | 0.292 |

Notes: ^aMale is reference category. ^bSingle is the reference category. ^cCentred at age 20. ^dNot having children is the reference category. ^eLondon is the reference category. ^fNo qualifications is the reference category. ^gClass 1 is the reference category. * $p < 0.05$, ** $p < 0.01$.

the data on theatre and cinema attendance, it is necessary to move on from merely two-way, or bivariate, analysis to analysis of a multivariate kind. That is to say, it is necessary to relate the chances of an individual being an omnivore rather than a univore to the full range of stratification variables referred to earlier and also to the socio-demographic variables of sex, age, marital status, family composition and region of residence that are introduced as controls. Only if the effects of all these variables are considered simultaneously can we hope to gain some reliable idea of their relative importance.

In Table 5 we report results from a binary logistic regression analysis. Such an analysis is

appropriate where the ‘dependent’ variable—that on which our explanatory interest centres—has just two possible values, ‘ x ’ or ‘not x ’: or, in our case, being an omnivore or not being an omnivore, and thus a univore. The β coefficients shown in the second column of the table represent the estimated effects of the variables listed in the first column. A positive coefficient implies that the higher the value of the explanatory variable, the higher the probability of being an omnivore rather than a univore, and a negative coefficient implies the opposite. Coefficients that are starred are statistically significant.

In assessing the results obtained from our regression analysis, we may begin with those relating to the socio-demographic control variables at the top of the table, since these turn out to be methodologically instructive. First, we may note a result that serves to confirm what several other investigators (e.g. O’Hagan, 1999; Quine, 1999) have previously reported: namely, that a highly significant gender effect exists in that women are more likely to be theatre goers than men; or, in our terms, women are more likely than men to be theatre and cinema omnivores rather than simply cinema-only univores. Second, though, the further finding of other investigators that the probability of theatre attendance increases with age cannot be confirmed. What we do find is that having a family that includes children below age 5, as compared with having no children, has a significant negative effect on the chances of being an omnivore. This would then suggest that where positive age effects on theatre attendance do show up in two-way analyses, or indeed in multivariate analyses in which family composition variables are not included, they should be very cautiously interpreted, and with the possibility being kept in mind that they may well reflect life-cycle stage rather than generation.

A third result likewise brings out the advantage of multivariate analysis. It can be seen that the coefficients for living in regions outside London, although often negative in sign, in no case achieve significance (though that for the North comes close). Again, then, the implication is that where in two-way analyses living outside London (or the South East) appears to have a negative effect on theatre attendance, this finding could easily mislead. If stratification variables are not simultaneously considered, it could be that what region variables largely pick up is not any specifically geographical effects, such as the location of venues, travelling times, etc., but rather the (concealed) effects of stratification variables, on account of the populations of regions differing in their class and status composition and in their average levels of income and educational attainment.

Now turning to the effects of stratification variables in our own analysis, which is our main focus of interest, the following points stand out. First, it can be seen that when class and status are included in the analysis together, the effect of status is highly significant and positive—i.e. the chances of being an omnivore rather than a univore increase with status—while class effects are for the most part insignificant. Only membership of Class 7, that of routine, largely manual wage workers, has a significant—negative—effect on the chances of being an omnivore. In other words, our results in the domain of theatre, dance and cinema do in this regard largely replicate those obtained in the domain of music, even if in the present case the preponderance of status over class effects is somewhat less marked, and thus our general theoretical expectation that status will be more closely associated with cultural consumption than will class, because such consumption represents an aspect of lifestyle through which status is readily expressed and displayed, is further borne out.

Second, Table 5 reveals that level of educational qualification tends also to have a significant and positive effect, over and above that of status (and class), on the chances of being an omnivore rather than a univore. Thus, in this regard, too, our findings in the case of music are broadly confirmed, although it should be noted that with theatre and cinema attendance, in contrast with musical consumption, the effects of education are not entirely ‘monotonic’. That is to say, the effects of having some educational qualifications rather than none on the chance of being an omnivore do not consistently increase with level of qualification. As can be seen, having CSE-level qualifications rather than none has no significant effect, and having tertiary but sub-degree qualifications has a weaker effect than having only A levels.

Third, our regression analysis also shows that even when all other stratification variables are taken into account, a highly significant and positive effect of income on theatre and cinema attendance still remains: the higher an individual’s income the more likely he or she is to be an omnivore rather than a univore. This result is then that which is most at variance with what we find in the analysis of musical consumption in which, as earlier noted, income proved to have no significant effect on the probability of being an omnivore rather than a univore, or in fact on the probability of being a true omnivore rather than an omnivore listener.

Conclusions

Three rival arguments concerning the relationship between social stratification and cultural consumption have been advanced and widely discussed in the sociological literature: what we have labelled as the homology, individualization and omnivore—univore arguments. In previous work on musical consumption in England, we have

presented findings that broadly favour the omnivore–univore argument, while lending little support to its competitors. In the present paper, we turn to the domain of theatre, dance and cinema, and ask how far the results of the analyses of musical consumption can be replicated.

In one respect, it could be said that the case of theatre, dance and cinema provides a yet more straightforward confirmation of the omnivore–univore argument than that of music. Our latent class analyses indicate just two main patterns of attendance, and, in turn, two main types of consumer: those, around one-third of the sample, who do appear omnivorous in having a relatively high probability of attending theatre performances of *all* the kinds considered *and* of going to the cinema; and those, around two-thirds of the sample, who are univorous in being cinema-goers only, if indeed they are consumers in the domain of theatre and cinema at all. That it is these two types of consumer that are empirically identifiable, and only these types, does then in itself serve to call both the homology and the individualization arguments into question. Our latent class analyses fail to reveal a cultural elite who systematically discriminate among different kinds of theatre performance; but at the same time, theatre and cinema attendance clearly cannot be regarded as simply forms of individual expression, devoid of all social patterning. Moreover, when stratification variables are introduced into the analyses the results obtained are generally those that would be expected under the omnivore–univore argument. Higher status, higher educational qualifications and a higher income all increase individuals' chances of being an omnivore rather than a univore. In sum, theatre and cinema attendance, like musical consumption, is quite evidently socially stratified, but on omnivore–univore rather than elite–mass lines.

At the same time, though, some differences from our findings in the domain of music have also to be recognized. To begin with, while status effects do generally dominate class effects on theatre and cinema attendance, as would be expected to be the case, class is not so completely overshadowed as it is in relation to musical consumption. The chances of being a theatre and cinema omnivore significantly decrease for those holding the least advantaged class positions (Class 7), even when status is controlled for. Further, while education is clearly an important influence on theatre and cinema attendance, its effects would appear to be somewhat less consistent and also relatively less strong than they are on musical consumption. And finally, and most notably, income, which appears statistically non-significant in regard to musical omnivorousness, exerts a highly significant and fairly substantial effect on the probability of being a theatre and cinema omnivore, even when other stratification variables are included in the analysis.

REFERENCES

- Bauman, Z. (1988). *Freedom*. Milton Keynes: Open University Press.
- Bourdieu, P. (1984). *Distinction: A social critique of the judgement of taste*. London: Routledge & Kegan Paul.
- Chan, T. W., & Goldthorpe, J. H. (2004). Is there a status order in contemporary British society? Evidence from the occupational structure of friendship. *European Sociological Review*, 20(5), 383–401.
- Chan, T. W., & Goldthorpe, J. H. (2005). *Social stratification and cultural consumption: music in England*. Paper presented at a one-day workshop on 'The social bases of cultural consumption', Oxford, March 18.
- Coulangeon, P. (2003). La stratification sociale des goûts musicaux. *Revue française de sociologie*, 44, 3–33.
- Featherstone, M. (1987). Lifestyle and consumer culture. *Theory, Culture and Society*, 4(1), 55–70.
- McCutcheon, A. L. (1987). *Latent class analysis*. Newbury Park, CA: Sage.

- O'Hagan, J. (1999). Policy analysis of theatre in the UK. *Cultural Trends*, 9(34), 25–29.
- Peterson, R. A., & Kern, R. M. (1996). Changing highbrow taste: from snob to omnivore. *American Sociological Review*, 61(5), 900–907.
- Peterson, R. A., & Simkus, A. (1992). How musical tastes mark occupational status groups. In M. Lamont & M. Fournier (Eds.), *Cultivating differences: Symbolic boundaries and the making of inequality* (pp. 152–186). Chicago, IL: University of Chicago Press.
- Quine, M. (1999). Audiences for live theatre in Britain: the present situation and some implications. *Cultural Trends*, 9(34), 1–24.
- Rose, D., & Pevalin, D. J. (Eds.). (2003). *A researcher's guide to the National Statistics Socio-economic Classification*. London: Sage.
- Skelton, A., Bridgwood, A., Duckworth, K., Hutton, L., Fenn, C., Creaser, C., & Babbidge, A. (2002). *Arts in England: Attendance, participation and attitudes in 2001* (Res. report 27). London: Arts Council England.
- Van Eijck, K. (2001). Social differentiation in musical taste patterns. *Social Forces*, 79(3), 1163–1184.
- Weber, M. (1922/1968). *Economy and society*. Berkeley, CA: University of California Press.

97 ■ Annette Lareau

Unequal Childhoods

Class, Race, and Family Life

There are many studies that tell us of the detrimental effects of poverty on children's lives, but it is less clear what the mechanisms are for the transmission of class advantage across generations.

I suggest that social class has an important impact on the cultural logic of childrearing (see Lareau 2003 for details). Middle-class parents, both white *and* black, appear to follow a cultural logic of childrearing that I call “concerted cultivation.” They enroll their children in numerous age-specific organized activities that come to dominate family life and create enormous labor, particularly for mothers. Parents see these activities as transmitting important life skills to children. Middle-class parents also stress language use and the development of reasoning. Talking plays a crucial role in the disciplinary strategies of middle-class parents. This “cultivation” approach results in a frenetic pace for parents, creates a

cult of individualism within the family, and emphasizes children's performance.

Among white and black working-class and poor families, childrearing strategies emphasize the “accomplishment of natural growth.” These parents believe that as long as they provide love, food, and safety, their children will grow and thrive. They do not focus on developing the special talents of their individual children. Working-class and poor children have more free time and deeper and richer ties within their extended families than the middle-class children. Some participate in organized activities, but they do so for different reasons than their middle-class counterparts. Working-class and poor parents issue many more directives to their children and, in some households, place more emphasis on physical discipline than do middle-class parents.

The pattern of concerted cultivation, with its stress on individual repertoires of activi-

Table 1. Argument of Unequal Childhoods: Class Differences in Childrearing

| | Childrearing Approach | |
|-------------------------------|--|---|
| | <i>Concerted Cultivation</i> | <i>Accomplishment of Natural Growth</i> |
| Key Elements | Parent actively fosters and assesses child's talents, opinions, and skills | Parent cares for child and allows child to grow |
| Organization of Daily Life | *multiple child leisure activities orchestrated by adults | *child "hangs out" particularly with kin |
| Language Use | *reasoning/directives *child contestation of adult statements *extended negotiations between parents and child | *directives *rare for child to question or challenge adults *general acceptance by child of directives |
| Interventions in Institutions | *criticisms and interventions on behalf of child *training of child to take on this role | *dependence on institutions *sense of powerlessness and frustrations *conflict between childrearing practices at home and at school |
| Consequences | Emerging sense of entitlement on the part of the child | Emerging sense of constraint on the part of the child |

ties, reasoning, and questioning, encourages an *emerging sense of entitlement* in children. Of course, not all parents and children are equally assertive, but the pattern of questioning and intervening among the white and black middle-class parents in the study contrasts sharply with the definitions of how to be helpful and effective observed among the white and black working-class and poor families. The pattern of the accomplishment of natural growth, with its emphasis on child-initiated play, autonomy from adults, and directives, encourages an *emerging sense of constraint*. Members of these families, adults as well as children, tend to be deferential and outwardly accepting (with sporadic moments of resistance) in their interactions with professionals such as doctors and educators. At the same time, however, compared to their middle-class counterparts, the white and black working-class and poor families are more distrustful of professionals in institutions. These are differences with long-term consequences. In a historical moment where the dominant society privileges active, informed, assertive clients of health and educa-

tional services, the various strategies employed by children and parents are not equally valuable. In sum, differences in family life lie not only in the advantages parents are able to obtain for their children, but also in the skills being transmitted to children for negotiating their own life paths.

Methodology

Study Participants

The study is based on interviews and observations of children eight to ten years of age and their families. A team of graduate research assistants and I collected the data. The first phase involved observations in third-grade public school classrooms, mainly in a metropolitan area in the Northeast. The schools serve neighborhoods in a white suburban area and two urban locales—one a white working-class neighborhood and the other a nearby poor black neighborhood. About one-half of the children are white and about one-half are black. One child is interracial. The research assistants and I carried out individual interviews (averaging two

hours each) with all of the mothers and most of the fathers (or guardians) of 88 children, for a total of 137 interviews. We also observed children as they took part in organized activities in the communities surrounding the schools. The most intensive part of the research, however, involved home observations of 12 children and their families. Nine of the 12 families came from the classrooms I observed, but the boy and girl from the two black middle-class families and the boy from the poor white family came from other sites. Most observations and interviews took place between 1993 and 1995, but interviews were done as early as 1990 and as late as 1997. This chapter focuses primarily on the findings from the observations of these 12 families since the key themes discussed here surfaced during this part of the fieldwork. I do include some information from the larger study to provide a context for understanding the family observations. All names are pseudonyms.

Intensive Family Observations

The research assistants and I took turns visiting the participating families daily, for a total of about 20 visits in each home, often in the space of one month. The observations were not limited to the home. Fieldworkers followed children and parents as they took part in school activities, church services and events, organized play, kin visits, and medical appointments. Most field observations lasted about three hours; sometimes, depending on the event (e.g., an out-of-town funeral, a special extended family event, or a long shopping trip), they lasted much longer. In most cases, there was one overnight visit. We often carried tape recorders with us and used the audiotapes for reference in writing up field notes. Families were paid \$350, usually at the end of the visits, for their participation.

A Note on Class

My purpose in undertaking the field observations was to develop an *intensive*, realistic

portrait of family life. Although I deliberately focused on only 12 families, I wanted to compare children across gender and race lines. Adopting the fine-grained differentiation of categories characteristic of current neo-Marxist and neo-Weberian empirical studies was not tenable. My choice of class categories was further limited by the school populations at the sites I had selected. Very few of the students were children of employers or of self-employed workers. I decided to concentrate exclusively on those whose parents were employees. Various criteria have been proposed to differentiate within this heterogeneous group, but authority in the workplace and “credential barriers” are the two most commonly used. I assigned the families in the study to a working-class or middle-class category based on discussions with each of the employed adults. They provided extensive information about the work they did, the nature of the organization that employed them, and their educational credentials. I added a third category: families not involved in the labor market (a population traditionally excluded from social class groupings) because in the first school I studied, a substantial number of children were from households supported by public assistance. To ignore them would have restricted the scope of the study arbitrarily. The final subsample contained 4 middle-class, 4 working-class, and 4 poor families.

Children’s Time Use

In our interviews and observations of white and black middle-class children, it was striking how busy they were with organized activities. Indeed, one of the hallmarks of middle-class children’s daily lives is a set of adult-run organized activities. Many children have three and four activities per week. In some families, every few days activities conflict, particularly when one season is ending and one is beginning. For example in the white middle-class family of the Tallingers, Garrett is on multi-

ple soccer teams—the “A” traveling team of the private Forest soccer club and the Inter-county soccer team—he also has swim lessons, saxophone lessons at school, private piano lessons at home, and baseball and basketball. These organized activities provided a framework for children’s lives; other activities were sandwiched between them.

These activities create labor for parents. Indeed, the impact of children’s activities takes its toll on parents’ patience as well as their time. For example, on a June afternoon at the beginning of summer vacation, in a white-middle-class family, Mr. Tallinger comes home from work to take Garrett to his soccer game. Garrett is not ready to go, and his lackadaisical approach to getting ready irks his father:

Don says, “Get your soccer stuff—you’re going to a soccer game!” Garrett comes into the den with white short leggings on underneath a long green soccer shirt; he’s number 16. He sits on an armchair catty-corner from the television and languidly watches the World Cup game. He slowly, abstractedly, pulls on shin guards, then long socks. His eyes are riveted to the TV screen. Don comes in: “Go get your other stuff.” Garrett says he can’t find his shorts. Don: “Did you look in your drawer?” Garrett nods. . . . He gets up to look for his shorts, comes back into the den a few minutes later. I ask, “Any luck yet?” Garrett shakes his head. Don is rustling around elsewhere in the house. Don comes in, says to Garrett, “Well, Garrett, aren’t you wearing shoes?” (Don leaves and returns a short time later): “Garrett, we HAVE to go! Move! We’re late!” He says this shortly, abruptly. He comes back in a minute and drops Garrett’s shiny green shorts on his lap without a word.

This pressured search for a pair of shiny green soccer shorts is a typical event in the Tallinger household. Also typical is the solution—a parent ultimately finds the missing object, while continuing to prod the child to hurry. The fact

that today’s frenzied schedule will be matched or exceeded by the next day’s is also par:

Don: (describing their day on Saturday) Tomorrow is really nuts. We have a soccer game, then a baseball game, then another soccer game.

This steady schedule of activity—that none of the middle-class parents reported having when they were a similar age—was not universal. Indeed, while we searched for a middle-class child who did not have a single organized activity we could not find one, but in working-class and poor homes, organized activities were much less common and there were many children who did not have any. Many children “hung out.” Television and video games are a major source of entertainment but outdoor play can trump either of these. No advanced planning, no telephone calls, no consultations between mothers, no drop-offs or pickups—no particular effort at all—is required to launch an activity. For instance, one afternoon, in a black working-class family, Shannon (in 7th grade) and Tyrec (in 4th grade) walk out their front door to the curb of the small, narrow street their house faces. Shannon begins playing a game with a ball; she soon has company:

(Two boys from the neighborhood walk up.) Shannon is throwing the small ball against the side of the row house. Tyrec joins in the game with her. As they throw the ball against the wall, they say things they must do with the ball. It went something like this: Johnny Crow wanted to know. . . . (bounces ball against the wall), touch your knee (bounce), touch your toe (bounce), touch the ground (bounce), under the knee (bounce), turn around (bounce). Shannon and Tyrec played about four rounds.

Unexpected events produce hilarity:

At one point Shannon accidentally threw the ball and it bounced off of Tyrec’s head. All the

kids laughed; then Tyrec, who had the ball, went chasing after Shannon. It was a close, fun moment—lots of laughter, eye contact, giggling, chasing.

Soon a different game evolves. Tyrec is on restriction. He is supposed to remain inside the house all day. So, when he thinks he has caught a glimpse of his mom returning home from work, he dashes inside. He reappears as soon as he realizes that it was a false alarm. The neighborhood children begin an informal game of baiting him:

The kids keep teasing Tyrec that his mom's coming—which sends him scurrying just inside the door, peering out of the screen door. This game is enacted about six times. Tyrec also chases Shannon around the street, trying to get the ball from her. A few times Shannon tells Tyrec that he'd better "get inside"; he ignores her. Then, at 6:50 [P.M.] Ken (a friend of Tyrec's) says, "There's your mom!" Tyrec scoots inside, then says, "Oh, man. You were serious this time."

Informal, impromptu outdoor play is common in Tyrec's neighborhood. A group of boys approximately his age, regularly numbering four or five but sometimes reaching as many as ten, play ball games together on the street, walk to the store to get treats, watch television at each other's homes, and generally hang out together.

Language Use

In addition to differences by social class in time use, we also observed differences in language use in the home. As others have noted (Bernstein, 1971; Heath, 1983) middle-class parents used more reasoning in their speech with children while working-class and poor parents used more directives. For example, in observations of the African American home of Alex Williams, whose father was a trial lawyer and mother was a high level corporate

executive, we found that the Williamses and other middle-class parents use language frequently, pleasurably, and instrumentally. Their children do likewise. For example, one January evening, Alexander is stumped by a homework assignment to write five riddles. He sits at the dinner table in the kitchen with his mother and a fieldworker. Mr. Williams is at the sink, washing the dinner dishes. He has his back to the group at the dinner table. Without turning around, he says to Alex, "Why don't you go upstairs to the third floor and get one of those books and see if there is a riddle in there?"

Alex [says] smiling, "Yeah. That's a good idea! I'll go upstairs and copy one from out of the book." Terry turns around with a dish in hand, "That was a joke—not a valid suggestion. That is not an option." He smiled as he turned back around to the sink. Christina says, looking at Alex: "There is a word for that you know, plagiarism." Terry says (not turning around), "Someone can sue you for plagiarizing. Did you know that?" Alex: "That's only if it is copyrighted." They all begin talking at once.

Here we see Alex cheerfully (though gently) goading his father by pretending to misunderstand the verbal instruction to consult a book for help. Mr. Williams dutifully rises to the bait. Ms. Williams reshapes this moment of lightheartedness by introducing a new word into Alexander's vocabulary. Mr. Williams goes one step further by connecting the new word to a legal consequence. Alex upstages them both. He demonstrates that he is already familiar with the general idea of plagiarism and that he understands the concept of copyright, as well.

In marked contrast to working-class and poor parents, however, even when the Williamses issue directives, they often include explanations for their orders. Here, Ms. Williams is reminding her son to pay attention to his teacher:

I want you to pay close attention to Mrs. Scott when you are developing your film. Those chemicals are very dangerous. Don't play around in the classroom. You could get that stuff in someone's eye. And if you swallow it, you could die.

Alex chooses to ignore the directive in favor of instructing his misinformed mother:

Alex corrects her, "Mrs. Scott told us that we wouldn't die if we swallowed it. But we would get very sick and would have to get our stomach pumped." Christina does not follow the argument any further. She simply reiterates that he should be careful.

Possibly because the issue is safety, Ms. Williams does not encourage Alex to elaborate here, as she would be likely to do if the topic were less charged. Instead, she restates her directive and thus underscores her expectation that Alex will do as she asks.

Although Mr. and Ms. Williams disagreed on elements of how training in race relations should be implemented, they both recognized that their racial and ethnic identity profoundly shaped their and their son's everyday experiences. They were well aware of the potential for Alexander to be exposed to racial injustice, and they went to great lengths to try to protect their son from racial insults and other forms of discrimination. Nevertheless, race did not appear to shape the dominant cultural logic of childrearing in Alexander's family or in other families in the study. All of the middle-class families engaged in extensive reasoning with their children, asking questions, probing assertions, and listening to answers. Similar patterns appeared in interviews and observations with other African American middle-class families.

A different pattern appeared in working-class and poor homes where there was simply less verbal speech than we observed in middle-class homes. There was also less speech between parents and children, a finding noted

by other observational studies (Hart and Risley, 1995). Moreover, interspersed with intermittent talk are adult-issued directives. Children are told to do certain things (e.g., shower, take out the garbage) and not to do others (e.g., curse, talk back). In an African American home of a family living on public assistance in public housing, Ms. McAllister uses one-word directives to coordinate the use of the single bathroom. There are almost always at least four children in the apartment and often seven, plus Ms. McAllister and other adults. Ms. McAllister sends the children to wash up by pointing to a child, saying, "Bathroom," and handing him or her a washcloth. Wordlessly, the designated child gets up and goes to the bathroom to take a shower.

Children usually do what adults ask of them. We did not observe whining or protests, even when adults assign time-consuming tasks, such as the hour-long process of hair-braiding. Lori McAllister is told to do for the four-year-old daughter of Aunt Dara's friend Charmaine:

Someone tells Lori, "Go do [Tyneshia's] hair for camp." Without saying anything, Lori gets up and goes inside and takes the little girl with her. They head for the couch near the television; Lori sits on the couch and the girl sits on the floor. [Tyneshia] sits quietly for about an hour, with her head tilted, while Lori carefully does a multitude of braids.

Lori's silent obedience is typical. Generally, children perform requests without comment. For example, at dinner one night, after Harold McAllister complains he doesn't like spinach, his mother directs him to finish it anyway:

Mom yells (loudly) at him to eat: "EAT! FINISH THE SPINACH!" (No response. Harold is at the table, dawdling.) Guion and Runako and Alexis finish eating and leave. I finish with Harold; he eats his spinach. He leaves all his yams.

The verbal world of Harold McAllister and other poor and working-class children offers some important advantages as well as costs. Compared to middle-class children we observed, Harold is more respectful towards adults in his family. In this setting, there are clear boundaries between adults and children. Adults feel comfortable issuing directives to children, which children comply with immediately. Some of the directives that adults issue center on obligations of children to others in the family (“don’t beat on Guion” or “go do [her] hair for camp”). One consequence of this is that Harold, despite occasional tiffs, is much nicer to his sister (and his cousins) than the siblings we observed in middle-class homes. The use of directives and the pattern of silent compliance are not universal in Harold’s life. In his interactions with peers, for example on the basketball “court,” Harold’s verbal displays are distinctively different than inside the household, with elaborated and embellished discourse. Nevertheless, there is a striking difference in linguistic interaction between adults and children in poor and working-class families when compared to that observed in the home of Alexander Williams. Ms. McAllister has the benefit of being able to issue directives without having to justify their decisions at every moment. This can make childrearing somewhat less tiring.

Another advantage is that Harold has more autonomy than middle-class children in making important decisions in daily life. As a child, he controls his leisure schedule. His basketball games are impromptu and allow him to develop important skills and talents. He is resourceful. He appears less exhausted than ten-year old Alexander. In addition, he has important social competencies, including his deftness in negotiating the “code of the street.”¹ His mother has stressed these skills in her upbringing, as she impresses upon her children the importance of “not paying no mind” to others, including drunks and drug dealers who hang out in the neighborhoods which Harold and Alexis negotiate.

Still, in the world of schools, health care facilities, and other institutional settings, these valuable skills do not translate into the same advantages as the reasoning skills emphasized in the home of Alexander Williams and other middle-class children. Compared to Alexander Williams, Harold does not gain the development of a large vocabulary, an increase of his knowledge of science and politics, a set of tools to customize situations outside the home to maximize his advantage, and instruction in how to defend his argument with evidence. His knowledge of words, which might appear, for example, on future SAT tests, is not continually stressed at home.

In these areas, the lack of advantage is *not* connected to the intrinsic value of the McAllister family life or the use of directives at home. Indeed, one can argue raising children who are polite and respectful children and do not whine, needle, or badger their parents is a highly laudable childrearing goal. Deep and abiding ties with kinship groups are also, one might further argue, important. Rather, it is the specific ways that institutions function that end up conveying advantages to middle-class children. In their standards, these institutions also permit, and even demand, active parent involvement. In this way as well, middle-class children often gain an advantage.

Intervention in Institutions

Children do not live their lives inside of the home. Instead, they are legally required to go to school, they go to the doctor, and many are involved in church and other adult-organized activities. In children’s institutional lives, we found differences by social class in how mothers monitored children’s institutional experiences. While in working-class and poor families children are granted autonomy to make their own way in organizations, in the middle-class homes, most aspects of the children’s lives are subject to their mother’s *ongoing* scrutiny.

For example in an African American middle-class home, where both parents are college graduates and Ms. Marshall is a computer worker and her husband a civil servant, their two daughters have a hectic schedule of organized activities including gymnastics for Stacey and basketball for Fern. When Ms. Marshall becomes aware of a problem, she moves quickly, drawing on her work and professional skills and experiences. She displays tremendous assertiveness, doggedness, and, in some cases, effectiveness in pressing institutions to recognize her daughters' individualized needs. Stacey's mother's proactive stance reflects her belief that she has a duty to intervene in situations where she perceives that her daughter's needs are not being met. This perceived responsibility applies across all areas of her children's lives. She is no more (or less) diligent with regard to Stacey and Fern's leisure activities than she is with regard to their experiences in school or church or the doctor's office. This is clear in the way she handles Stacey's transition from her township gymnastics classes to the private classes at an elite private gymnastic program at Wright's.

Ms. Marshall describes Stacey's first session at the club as rocky:

The girls were not warm. And these were little . . . eight and nine year old kids. You know, they weren't welcoming her the first night. It was kinda like eyeing each other, to see, you know, "Can you do this? Can you do that?"

More importantly, Ms. Marshall reported that the instructor is brusque, critical and not friendly toward Stacey. Ms. Marshall cannot hear what was being said, but she could see the interactions through a window. A key problem is that because her previous instructor had not used the professional jargon for gymnastic moves, Stacey does not know these terms. When the class ends and she walks out, she is visibly upset. Her mother's reaction is a common one among

middle-class parents: She does not remind her daughter that in life one has to adjust, that she will need to work even harder, or that there is nothing to be done. Instead, Ms. Marshall focuses on Tina, the instructor, as the source of the problem:

We sat in the car for a minute and I said, "Look, Stac," I said. She said, "I-I," and she started crying. I said, "You wait here." The instructor had come to the door, Tina. So I went to her and I said, "Look." I said, "Is there a problem?" She said, "Aww . . . she'll be fine. She just needs to work on certain things." Blah-blah-blah. And I said, "She's really upset. She said you-you-you [were] pretty much correcting just about everything." And [Tina] said, "Well, she's got—she's gotta learn the terminology."

Ms. Marshall acknowledges that Stacey isn't familiar with specialized and technical gymnastics terms. Nonetheless, she continues to defend her daughter:

I do remember, I said to her, I said, "Look, maybe it's not all the student." You know, I just left it like that. That, you know, sometimes teaching, learning and teaching, is a two-way proposition as far as I'm concerned. And sometimes teachers have to learn how to, you know, meet the needs of the kid. Her style, her immediate style was not accommodating to—to Stacey.

Here Ms. Marshall is asserting the legitimacy of an individualized approach to instruction. She frames her opening remark as a question ("Is there a problem?"). Her purpose, however, is to alert the instructor to the negative impact she has had on Stacey ("She's really upset."). Although her criticism is indirect ("Maybe it's not all the student . . ."), Ms. Marshall makes it clear that she expects her daughter to be treated differently in the future. In this case, Stacey does not hear what her mother says, but she

knows that her wishes and feelings are being transmitted to the instructor in a way that she could not do herself.

Although parents were equally concerned about their children's happiness, in working-class and poor homes we observed different patterns of oversight for children's institutional activities. For example in the white working-class home of Wendy Driver, Wendy's mother does not nurture her daughter's language development like Alexander Williams' mother does her son's. She does not attempt to draw Wendy out or follow up on new information when Wendy introduces the term mortal sin while the family is sitting around watching television. But, just like Ms. Williams, Ms. Driver cares very much about her child and just like middle-class parents she wants to help her daughter succeed. Ms. Driver keeps a close and careful eye on her Wendy's schooling. She knows that Wendy is having problems in school. Ms. Driver immediately signs and returns each form Wendy brings home from school and reminds her to turn the papers in to her teacher.

Wendy is "being tested" as part of an ongoing effort to determine why she has difficulties with spelling, reading, and related language-based activities. Her mother welcomes these official efforts but she did not request them. Unlike the middle-class mothers we observed, who asked teachers for detailed information about every aspect of their children's classroom performance and relentlessly pursued information and assessments outside of school as well, Ms. Driver seems content with only a vague notion of her daughter's learning disabilities. This attitude contrasts starkly with that of Stacey Marshall's mother, for example. In discussing Stacey's classroom experiences with fieldworkers, Ms. Marshall routinely described her daughter's academic strengths and weaknesses in detail. Ms. Driver never mentions that Wendy is doing grade-level work in math but is reading at a level a full

three years below her grade. Her description is vague:

She's having problems. . . . They had a special teacher come in and see if they could find out what the problem is. She has a reading problem, but they haven't put their finger on it yet, so she's been through all kinds of special teachers and testing and everything. She goes to Special Ed, I think it's two classes a day . . . I'm not one hundred percent sure—for her reading. It's very difficult for her to read what's on paper. But then—she can remember things. But not everything. It's like she has a puzzle up there. And we've tried, well, they've tried a lot of things. They just haven't put their finger on it yet.

Wendy's teachers uniformly praise her mother as "supportive" and describe her as "very loving," but they are disappointed in Ms. Driver's failure to take a more active, interventionist role in Wendy's education, especially given the formidable nature of her daughter's learning problems. From Ms. Driver's perspective, however, being actively supportive means doing whatever the teachers tell her to do.

Whatever they would suggest, I would do. They suggested she go to the eye doctor, so I did that. And they checked her and said there was nothing wrong there.

Similarly, she monitors Wendy's homework and supports her efforts to read:

We listen to her read. We help her with her homework. So she has more attention here in a smaller household than it was when I lived with my parents. So, we're trying to help her out more, which I think is helping. And with the two [special education] classes a day at the school, instead of one like last year, she's learning a lot from that. So, we're just hoping it takes time and that she'll just snap out of it.

But Ms. Driver clearly does not have an *independent* understanding of the nature or degree of Wendy's limitations, perhaps because she is unfamiliar with the kind of terms the educators use to describe her daughter's needs (e.g., a limited "sight vocabulary," underdeveloped "language arts skills"). Perhaps, too, her confidence in the school staff makes it easier for her to leave "the details" to them: "Ms. Morton, she's great. She's worked with us for different testing and stuff." Ms. Driver depends on the school staff's expertise to assess the situation and then share the information with her:

I think they just want to keep it in the school till now. And when they get to a point where they can't figure out what it is, and then I guess they'll send me somewhere else. . . .

Her mother is not alarmed, because "the school" has told her not to worry about Wendy's grades:

Her report card—as long as it's not spelling and reading—spelling and reading are like F's. And they keep telling me not to worry, because she's in the Special Ed class. But besides that, she does good. I have no behavior problems with her at all.

Ms. Driver wants the best possible outcome for her daughter and she does not know how to achieve that goal without relying heavily on Wendy's teachers:

I wouldn't even know where to start going. On the radio there was something for children having problems reading and this and that, call. And I suggested it to a couple different people, and they were like, wait a second, it's only to get you there and you'll end up paying an arm and a leg. So I said to my mom, "No, I'm going to wait until the first report card and go up and talk to them up there."

Thus, in looking for the source of Ms. Driver's deference toward educators, the answers don't seem to lie in her having either a shy personality or underdeveloped mothering skills. To understand why Wendy's mother is accepting where Stacey Marshall's mother would be aggressive, it is more useful to focus on social class position, both in terms of how class shapes worldviews and how class affects economic and educational resources. Ms. Driver understands her role in her daughter's education as involving a different set of responsibilities from those perceived by middle-class mothers. She responds to contacts from the school—such as invitations to the two annual parent-teacher conferences—but she does not initiate them. She views Wendy's school life as a separate realm, and one in which she, as a parent, is only an infrequent visitor. Ms. Driver expects that the teachers will teach and her daughter will learn and that, under normal circumstances, neither requires any additional help from her as a parent. If problems arise, she presumes that Wendy will tell her; or, if the issue is serious, the school will contact her. But what Ms. Driver fails to understand, is that the educators expect her to take on a pattern of "concerted cultivation" where she actively monitors and intervenes in her child's schooling. The teachers asked for a complicated mixture of deference and engagement from parents; they were disappointed when they did not get it.

Conclusions

I have stressed how social class dynamics are woven into the texture and rhythm of children and parents' daily lives. Class position influences critical aspects of family life: time use, language use, and kin ties. Working-class and middle-class mothers may express beliefs that reflect a similar notion of "intensive mothering," but their behavior is quite different. For that reason, I have described sets of paired beliefs and actions as a "cultural logic"

of childrearing. When children and parents move outside the home into the world of social institutions, they find that these cultural practices are not given equal value. There are signs that middle-class children benefit, in ways that are invisible to them and to their parents, from the degree of similarity between the cultural repertoires in the home and those standards adopted by institutions.

NOTE

1. Elijah Anderson, *Code of the Street*, New York: W. W. Norton (1999).

REFERENCES

- Anderson, Elijah. 1999. *Code of the Street*. New York, NY: W. W. Norton.
- Bernstein, Basil. 1971. *Class, Codes, and Control: Theoretical Studies Towards a Sociology of Language*. New York, NY: Schocken.
- Hart, Betty and Todd R. Risley. 1995. *Meaningful Differences in the Everyday Experiences of Young American Children*. New Haven: Yale University Press.
- Heath, Shirley Brice. 1983. *Ways with Words: Language, Life, and Work in Communities and Classrooms*. Cambridge: Cambridge University Press.
- Lareau, Annette. 2003. *Unequal Childhoods: Class, Race, and Family Life*. Berkeley, CA: University of California Press.

98 ■ Eszter Hargittai

The Digital Reproduction of Inequality

By the beginning of the twenty-first century, information and communication technologies (ICT) had become a staple of many people's everyday lives. The level of instantaneous connectivity—to others and to an abundance of information—afforded by advances in ICT is unprecedented. With economies increasingly dependent on knowledge-intensive activities, the unequal distribution of knowledge and information access across the population may be linked increasingly to stratification. No sooner did the Internet start diffusing to the general population in the mid-1990s than did debates spring up about its implications for social inequality. From the perspective of social mobility, digital media could offer people, organizations, and societies the opportunity to improve their positions regardless of existing constraints. From the point of view of social reproduction, however, ICT could exacerbate existing inequalities by increasing the oppor-

tunities available to the already privileged while leading to the growing marginalization of the disadvantaged.

Most initial attention concerning ICT's implications for social stratification focused on what segments of the population have access to the Internet or are Internet users (e.g., Bimber 2000; Hoffman and Novak 1998). Access is usually defined as having a network-connected machine in one's home or workplace. Use more specifically refers to people's actual use of the medium beyond merely having access to it. The term "digital divide" became a popular expression to sum up concerns about the unequal diffusion of the medium. The concept is most often understood in binary terms: someone either has access to the medium or does not, someone either uses the Internet or does not.

However, as an increasing portion of the population has gone online, a dichotomous

Original chapter commissioned for this book. The author is grateful for valuable comments from David Grusky and for helpful discussions with Paul DiMaggio and Russ Neuman (who also deserves credit for the title). She thanks the Center for Advanced Study in the Behavioral Sciences and the John D. and Catherine T. MacArthur Foundation for their support.

approach is no longer sufficient to address the different dimensions of inequality associated with digital media uses. The term *digital inequality* better captures the spectrum of differences associated with ICT uses (DiMaggio et al. 2004). A more refined approach considers different aspects of the divide, focusing on such details as quality of equipment, autonomy of use, the presence of social support networks, experience and user skills, in addition to differences in types of uses (Barzilai-Nahon 2006; Dewan and Riggins 2006; DiMaggio et al. 2004; Mossberger, Tolbert, and Stansbury 2003; Norris 2001; van Dijk 2005; Warschauer 2003).

Variation in basic usage rates continues to exist, so considering the core digital divide of access versus no access remains an important undertaking. However, to understand in a nuanced manner the implications of ICT for social inequality, it is important to analyze differences among users as well. This chapter will do both, starting with a historical look at connectivity patterns by population segments. This discussion is then followed by an explanation of why it is important to distinguish among users of digital media. A conceptual framework lays out the processes through which users' social position influences their ICT uses and how this in turn may contribute to social inequality even among the connected. Although the primary focus here is on Internet use in the United States, the main arguments made can be extended to the use of other digital devices in other national contexts as well.

The Haves and Have Nots

In 1995, the National Telecommunications Information Administration of the U.S. Department of Commerce published a report entitled "Falling Through the Net: A Survey of the 'Have Nots' in Rural and Urban America" in which policy makers analyzed data from the Current Population Survey about computer and modem use among Americans. Findings suggested that different segments of

the population were using digital technologies at varying rates. In subsequent years, these reports began to focus increasingly on Internet access as opposed to computer use only, documenting continued differences among various population groups (NTIA 1998, 1999, 2000). The reports' titles highlighted concerns about inequality as they all began with the phrase "Falling Through the Net."

Breaking with tradition, the fifth report of the NTIA published in 2002, based on data collected in 2001, was called "A Nation Online: How Americans Are Expanding Their Use of the Internet" (NTIA 2002). The title of this last report no longer focused on differences. Rather, it highlighted the fact that more and more Americans were going online. While significant differences remained among various population segments regarding their rates of connectivity, the report focused on the growing number of people accessing the Internet through high-speed connections. This change in focus may imply that Internet use had reached universal levels, but that was not the case.

Overall findings from the reports suggested that while the Internet may have been spreading to an increasing portion of the American population, certain segments were much more likely to be online than others. In particular, men, younger people, whites, non-Hispanics, urban residents, the more highly educated, and those with higher income were more connected than their counterparts. Gender differences leveled off after a few years with respect to basic access (Ono and Zavodny 2003) although not regarding amount of use and skill (Hargittai and Shafer 2006). In contrast, all other differences in access among different population segments remained throughout the years.

Looking at adoption figures over time, we find that while all segments increased their participation significantly, disparities continued to persist. Figures 1 and 2 illustrate this point for income and education, respectively. As Figure 1 shows, people in all income

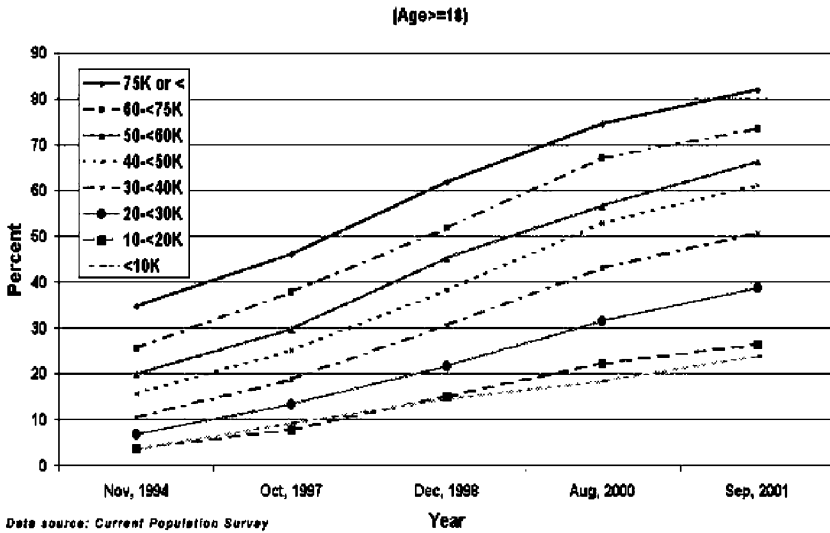


Figure 1. Internet Adoption by Level of Income in the United States, 1994–2001

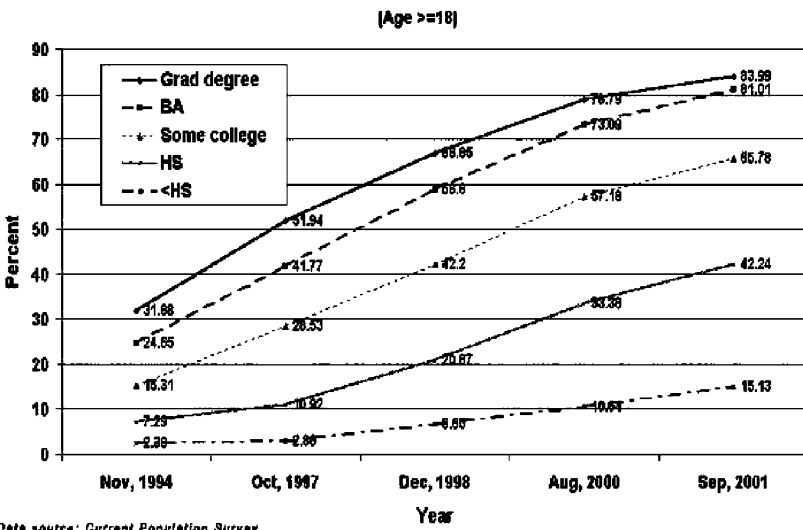


Figure 2. Internet Adoption by Level of Education in the United States, 1994–2001

brackets increased their participation over time, but the slopes in the higher income brackets are somewhat steeper, leading to an increased gap among groups over time. The data points in Figure 2 tell a similar story. Although the gap between those who have a college degree and graduate education narrowed over the years, all other gaps widened over time. In particular, the least educated—those with less than a high school degree—increased their connectivity minimally over the eight-

year period. Overall, these trend data suggest that while all population segments may have become increasingly connected, serious divides persist with the most disadvantaged trailing behind the more privileged in significant ways.

We have less data on the diffusion of cell phones, but the little evidence that has surfaced suggests similar patterns of unequal distribution among the population. Looking at the earlier years of diffusion using data from 1994–1998, researchers found that mobile

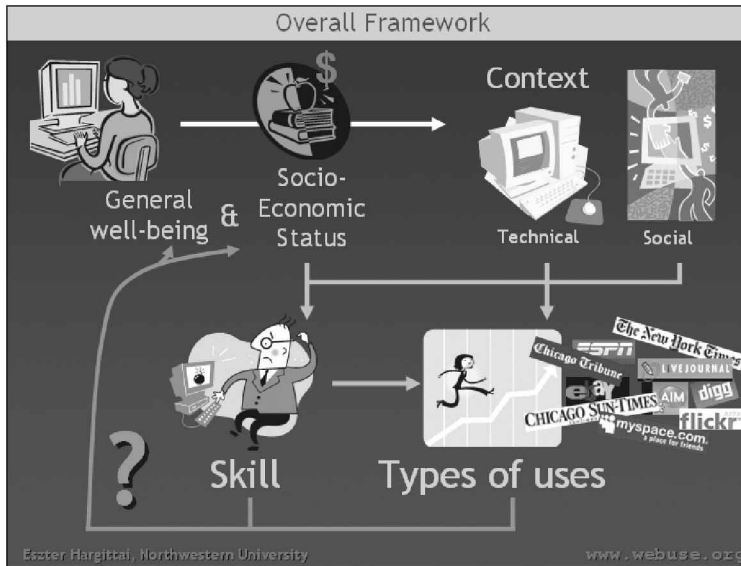


Figure 3. Framework for Studying the Implications of ICT Use for Social Inequality

technology adoption was positively related to both income and education (Wareham, Levy, and Shi 2004). Based on data from 2006, Horrigan showed that people with lower levels of income were less likely to be users (Horrigan 2006). Analyses (by Hargittai for this chapter) of these same data collected by the Pew Internet and American Life Project also found that those with higher levels of education were more likely to own cell phones, and these findings are robust (also for income) when controlling for other factors. Moreover, those with higher income tend to own cell phones with more functionality (e.g., the ability to send and receive text messages, take photos, and go online). While this literature is not as elaborate as the one on different rates of Internet connectivity, these findings clearly suggest that the digital divide expands beyond Internet use into the domain of mobile technology adoption as well.

Differences Among the Connected

The uses of ICT can differ considerably with divergent outcomes for one's life chances. Therefore, it is imperative to examine varia-

tions in use among those who have crossed the digital divide fault line to the land of the connected. Baseline Internet use statistics do not distinguish among those who go online for no more than checking sports scores or TV schedules and those who use the medium for learning new skills, finding deals and job opportunities, participating in political discussions, interacting with government institutions, and informing themselves about health matters. Yet such differentiated uses can have significant implications for how ICT uses may relate to life outcomes. This section describes how various user characteristics and one's social surroundings influence digital media uses.

People's Internet uses do not happen in isolation of their societal position and the social institutions they inhabit. A refined approach to digital inequality recognizes that people's socioeconomic status influences the ways in which they have access to and use ICT. In addition to factors such as age, gender, race, ethnicity, disability status, education, and wealth, one's social surroundings are also relevant to one's ICT experiences. Figure 3 presents a graphical representation of this framework.

The basic premise is that the societal position that users inhabit influences aspects of their digital media uses, such as the technical equipment to which they have access and the level of autonomy they possess when using the medium. Autonomy of use is understood as the freedom to use digital media when and where one wants to. Twenty-four-hour access at home can yield a much more autonomous user experience than having to drive half an hour to a public library where one competes with others for usage time and where filtering software limits the types of materials within reach. Similarly, a workplace that allows Web use without constraints results in a very different experience from a job environment where one's online actions are constantly monitored. Quality of equipment (available hardware, software, and connection speed) and autonomy of use can both be a function of one's socioeconomic status.

The use of and learning about digital media both happen in social contexts. In addition to autonomy of use, which itself is a certain social context, the availability of other users in one's social circles can have important implications for one's online experiences. The relevant mechanisms through which social networks matter can be grouped into two main categories: informal and more directed information exchange. The former refers to knowledge one amasses through everyday discussions with peers about digital media uses and includes suggestions passed along by others through email or at the water cooler. The latter highlights the importance of support networks when users encounter a specific problem during their experiences with ICT. When faced with a difficulty, it makes a difference to have access to knowledgeable networks that help in finding a solution.

All of these factors then influence users' online abilities and what they are able to accomplish using digital media. While many online actions may seem trivial to the experienced user, most activities require some level of know-how. From recognizing what mater-

ial is available online to being able to find it efficiently, from knowing how to contribute to online content production to knowing where to find relevant networks, from having the ability to evaluate content credibility to being vigilant about privacy and security concerns, informed uses of digital media rely on many important skills (Hargittai 2007b).

One's social position, the context of one's use, and one's online abilities then all have the potential to influence the types of uses to which one puts the medium. Some uses are more likely to yield beneficial outcomes than others. The next section will enumerate the ways in which ICT uses may improve, or in some cases impede, one's life chances.

How ICT Use Matters

The most pressing question for students of social inequality is whether the usage dimensions described above then loop back and translate into differences in users' socioeconomic position. What are the processes through which more informed and frequent uses of ICT may privilege some users over others? There are several ways in which differential ICT uses may influence access to the types of assets, resources, and valued goods that underlie stratification systems (Grusky 2008). The overarching idea is that certain types of ICT uses can result in increased human capital, financial capital, social capital, and cultural capital while other types of uses may outright disadvantage the uninformed.

With more and more jobs requiring the synthesis and analysis of varying types of information, employees with advanced Internet user skills can perform their jobs more effectively and efficiently. The Internet makes vast amounts of information available *so long as* one knows how to locate desired material. While theoretically all public Web pages are equally available to all users, not everybody possesses the necessary skills to (1) recognize in all situations the types of content relevant to a task that can be found online; (2) find

the information; (3) do so efficiently; and (4) carefully evaluate the results to avoid misinformation or, worse, fraudulent outcomes.

Even if people do not know how to perform certain tasks, advanced skills will allow them to find assistance online. Since skill encompasses the ability to find others who may have the desired information and efficiently contact them for guidance, even when lacking know-how most relevant to the task at hand, the skilled user can benefit through informed use of the medium. This all leads to more tasks getting done quicker and more efficiently with possibly higher-quality results than would be possible if relying on fewer resources. In addition to helping with the performance of on-the-job tasks, ICT also allow people to develop additional skills that may advantage them in the labor market. Free tutorials exist online for training in a myriad of domains from foreign languages to software applications, from design skills to networking to productivity enhancement tips.

Enterprising ways of using the Internet can save a person significant amounts of money. Several services exist that make comparison shopping very easy, allowing the user to find the best deals on an item without spending money on gas and time on driving from one store to the next. The use of auction sites expands options even further. Moreover, the especially knowledgeable can take advantage of other people's mistakes by searching for items with spelling mistakes thereby avoiding bidding wars given that misspelled items are seen by few (Hargittai 2006). In addition to savings through informed purchasing, people can also make money by selling products on the Web. While putting one's items on the market used to require significant upfront investment, ICT have lowered the barriers to entry for putting things up for sale accessible to a large buyer base, assuming one knows what services help with reaching the largest or most relevant purchasers.

The potential of ICT for expanding one's social networks is enormous, although efficient

and relevant ways of doing so are not trivial and require some amount of know-how. In some cases the Internet simply complements more traditional methods of networking while in others the medium is the main facilitator of a new relationship. The former refers to use of the tool to contact people who are already in one's extended network. The latter occurs thanks to various online interactions that can range from the exchange of information on a mailing list to the exchange of goods and services extending well beyond the Web. People find rides, coordinate meetings, and get emotional support from others online. But as with all other aspects of Internet use, skill matters. Finding the relevant communities and being sufficiently vigilant not to place oneself in harm's way are all important aspects of building social capital on the Web.

Familiarity with the latest trends can serve as status markers. Being able to discuss special topics in depth can help create bonds between people. Thanks to the Internet, certain subjects formerly much less accessible to the general public are more widely available. It is no longer necessary to go see a museum's special exhibition to have the facility to discuss what is on display since many galleries now put their pieces online. It is possible to develop a reasonably informed opinion about a restaurant halfway across the world simply by reading the many reviews available and constantly updated online. Knowing how to locate information about travel destinations—from driving directions to entertainment options—can yield more influence to the informed. Being able to draw on a myriad of topics while conversing with higher-ups can leave a good impression. While resources about diverse topics have long been available to the public, the ease and speed with which the Internet delivers information is unprecedented.

Informed users can be more engaged in the political process than those who rely exclusively on broadcast media for their political information seeking. Whether finding like-minded people or informing oneself in depth

about the other side's perspective, the Internet allows for the exchange of political opinions much more than any other medium. Creating petitions and mobilizing others around a cause can be facilitated significantly with various online tools. Again, however, knowledge of what is available in this domain and how one can implement the services to benefit one's specific needs and interests is an essential part of any such undertaking.

The above are examples of how informed uses can have beneficial outcomes. There is another side to online actions, however. Uninformed uses may have outright problematic if not detrimental consequences. Do users stop to think about the context of, for example, an email message that requests confidential information from them? If everyone was aware of these issues, then phishing emails—messages that pretend to be from a reputable source to extract confidential information from users—would not lead to people giving up their passwords to Web sites that contain private information such as their bank accounts. Yet, we know that even among young adults—the generation that is growing up with these media—many lack the necessary knowledge to approach possibly malicious email with care (Hargittai *In press*). While fraud has always existed, the scope of malicious activities and their consequences have skyrocketed.

Related to online social interactions discussed above, but sufficiently distinct to merit its own discussion, is one's reputation developed from one's online pursuits. Sending emails from the privacy of one's home or office leads many to behave less carefully than they would in a public social setting. Few interactions on the Web are truly anonymous yet many people do not realize this when sending critical messages or posting comments on Web sites. An unwelcome remark can have negative consequences if targeting the wrong person. Alternatively, critical comments by others can tarnish the reputation of the person under attack. In contrast, a well-thought-out online presence can result in sig-

nificant benefits. Those who participate regularly in online discussions and maintain Web sites frequented by many can amass fame that can later translate into fortune as well.

Generally speaking, many of the skills needed to reap the benefits listed here—or sidestep negative implications—can be learned from one's immediate networks. Growing up in a household that has the latest gadgets and digital media resources will make a difference when a student then encounters these tools in the classroom. Having siblings who can navigate the technologies will help in the transfer of relevant know-how. Living in neighborhoods where many in one's proximity are also discovering the latest ICT options will allow for more opportunities to develop savvy in the domain of digital media than a situation in which one is isolated without access to relevant technologies and knowledgeable networks. Bourdieu's cultural capital (1973) applied to the twenty-first century must incorporate the differential exposure to, understanding, and use of ICT. Work looking at young adults' digital literacy has found a statistically significant positive relationship between Internet savvy and the parental education of respondents (Hargittai 2007b).

Overall, it is important to recognize that ICT do not nullify the effects of other variables on one's life chances. People's ICT uses happen in the context of their lives, influenced by their socioeconomic status and social surroundings. The question is whether ICT uses have an independent effect on life outcomes. Given the relative newness of the Internet and other digital media uses at a mass societal level, this field is in its infancy and lacks the longitudinal data necessary to answer many of the questions raised here. Nonetheless, preliminary findings seem to suggest that ICT reinforce inequalities more than alleviating differences. Although not without its critics (DiNardo and Pischke 1997), the general consensus seems to be that skill-biased technological change, and especially computerization, is an important

source of the rise in earnings inequality in recent years (Krueger 1993). A more recent study found Internet use to have an independent effect on wage differences, suggesting tangible outcomes of being among the connected (DiMaggio and Bonikowski 2006).

Luxury Good or Essential Tool?

In 2001, then chair of the Federal Communication Commission Michael Powell likened the digital divide to a luxury car divide, stating: “I think there is a Mercedes divide, I would like to have one, but I can’t afford one” (quoted in Clewley 2001).

Is Internet use simply a luxury item with people’s connectivity—or lack thereof—merely a reflection of their preferences for the medium? As ICT become ever more central to our social infrastructure one can no longer participate meaningfully in our society without deep and ongoing usage of digital media. Once an entire society is built around these tools, they can no longer be considered simply as luxury goods. While the car and the telephone may have, at one time, been regarded as extravagant expenditures of the wealthy, once contemporary society was thoroughly built around these innovations they became necessities for operating in society and those who lacked them were socially excluded.

While it may be that some people opt out of ICT use based on an informed understanding of all that the Internet has to offer, much more likely is that people do not realize the many necessities and benefits of digital media. As an increasing number of activities between institutions and individuals move online—concerning both the public and the private sector—being a nonuser will have growing implications for people’s access to various services. If government institutions assume a familiarity with and access to the medium, then lacking access to and understanding of such resources some will be unable to interact with and navigate the system optimally.

Take, for example, the case of Medicare Part D in 2006. The government introduced a new system and required the elderly to make important choices about their health insurance. In response to concerns about the difficulty of navigating the system, the administration created a Web site and directed people to it for assistance with the program (Freese, Rivas, and Hargittai 2006). However, the resource was very complicated to navigate for many and the assumption that the elderly could access and understand the site was unfounded, as many were uninformed about or confused by the system. Similarly, more and more commercial services make material available on the Web and charge extra fees to those who interact with the company offline. When important government services are primarily accessible online and when there is an extra financial cost to handling matters with businesses offline, then having access to the Internet and knowing how to use it can no longer be considered an optional luxury item.

Conclusion

Disparities in people’s Web-use abilities and uses have the potential to contribute to social inequalities rather than alleviate them. Those who know how to navigate the Web’s vast landscape and how to use digital media to address their needs can reap significant benefits from it. In contrast, those who lack abilities in these domains may have a harder time dealing with certain logistics of everyday life, may miss out on opportunities, and may also obtain incorrect information from unreliable sources or come to rely on unsubstantiated rumors. Differentiated uses of digital media have the potential to lead to increasing inequalities benefiting those who are already in advantageous positions and denying access to better resources to the underprivileged. Merton’s (1973) observation “Unto every one who hath shall be given” applies to this domain. Preliminary findings from this emerging field suggest that initial advantages translate into

increasing returns over time for the digitally connected and digitally skilled.

REFERENCES

- Barzilai-Nahon, Karine. 2006. "Gaps and Bits: Conceptualizing Measurements for Digital Divide/s." *Information Society* 22: 269–278.
- Bimber, B. 2000. "The Gender Gap on the Internet." *Social Science Quarterly* 81: 868–876.
- Bourdieu, Pierre. 1973. "Cultural Reproduction and Social Reproduction." Pp. 71–112 in *Knowledge, Education, and Cultural Change*, edited by Richard Brown. London: Tavistock.
- Clewley, Robin. 2001. "I Have a (Digital) Dream." In *Wired*. <http://www.wired.com/politics/law/news/2001/04/43349>.
- Dewan, S., and F. Riggins. 2006. "The Digital Divide: Current and Future Research Directions." *Journal of the Association of Information Systems* 6: 298–337.
- DiMaggio, Paul, and Bart Bonkowski. Forthcoming. "Make Money Surfing the Web? The Impact of Internet Use on the Earnings of U.S. Workers." *American Sociological Review*.
- DiMaggio, Paul, Eszter Hargittai, Coral Celeste, and Steven Shafer. 2004. "Digital Inequality: From Unequal Access to Differentiated Use." Pp. 355–400 in *Social Inequality*, edited by Kathryn Neckerman. New York: Russell Sage Foundation.
- DiNardo, John E., and Jorn-Steffen Pischke. 1997. "The Returns to Computer Use Revisited: Have Pencils Changed the Wage Structure Too?" *Quarterly Journal of Economics* 112: 291–303.
- Freese, Jeremy, Salvador Rivas, and Eszter Hargittai. 2006. "Cognitive Ability and Internet Use Among Older Adults." *Poetics* 34: 236–249.
- Grusky, David B. 2008. *Social Stratification: Class, Race, and Gender in Sociological Perspective*. Boulder, CO: Westview Press.
- Hargittai, Eszter. 2006. "Hurdles to Information Seeking: Explaining Spelling and Typographical Mistakes in Users' Online Search Behavior." *Journal of the Association of Information Systems* 7. <http://jais.aisnet.org/articles/default.asp?vol=7&art=1>.
- _____. 2007a. "Beyond Gigs of Log Data: The Social Aspects of Internet Use." In *Tech Talks*. Mountain View, CA: Google, Inc. January 30. <http://video.google.com/videoplay?docid=6884513936578531954>.
- _____. 2007b. "A Framework for Studying Differences in People's Digital Media Uses." In *Cyberworld Unlimited?*, edited by Nadia Kutscher and Hans-Uwe Otto. Wiesbaden: VS Verlag für Sozialwissenschaften/GWV Fachverlage GmbH. Pp. 121–137.
- _____. In press. "The Role of Expertise in Navigating Links of Influence." In *The Hyperlinked Society*, edited by Joseph Turow and Lokman Tsui. Ann Arbor: University of Michigan Press.
- Hargittai, Eszter, and Steven Shafer. 2006. "Differences in Actual and Perceived Online Skills: The Role of Gender." *Social Science Quarterly*. 87(2):432–448. June.
- Hoffman, D. L., and T. P. Novak. 1998. "Bridging the Racial Divide on the Internet." Pp. 390–391 in *Science* 280(5362).
- Horrigan, John. 2006. "Tech Users: What They Have, How It Matters." In KMB Video Journal Conference. St. Petersburg Beach, Florida.
- Krueger, Alan. 1993. "How Computers Have Changed the Wage Structure: Evidence from Microdata, 1984–1989." *Quarterly Journal of Economics* 108: 33–60.
- Merton, R. K. 1973. *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago: University of Chicago Press.
- Mossberger, Karen, Caroline J. Tolbert, and Mary Stansbury. 2003. *Virtual Inequality: Beyond the Digital Divide*. Washington, DC: Georgetown University Press.
- National Telecommunications and Information Administration. 1998. "Falling Through the Net II: New Data on the Digital Divide." Washington, DC: NTIA.
- _____. 1999. "Falling Through the Net: Defining the Digital Divide." Washington, DC: NTIA.
- _____. 2000. "Falling Through the Net: Toward Digital Inclusion." Washington, DC: NTIA.
- _____. 2002. "A Nation Online: How Americans Are Expanding Their Use of the Internet." Washington, DC: NTIA.
- Norris, P. 2001. *Digital Divide: Civic Engagement, Information Poverty and the Internet in Democratic Societies*. New York: Cambridge University Press.
- Ono, Hiroshi, and Madeline Zavodny. 2003. "Gender and the Internet." *Social Science Quarterly* 84: 111–121.
- van Dijk, Jan A. G. M. 2005. *The Deepening Divide*. London: Sage Publications.
- Wareham, Jonathan, Armando Levy, and Wei Shi. 2004. "Wireless Diffusion and Mobile Computing: Implications for the Digital Divide." *Telecommunications Policy* 28: 439–457.
- Warschauer, M. 2003. *Technology and Social Inclusion*. Cambridge, MA: MIT Press.

Politics

99 ■ Michael Hout and Benjamin Moodie

The Realignment of U.S. Presidential Voting, 1948–2004

Introduction

U.S. presidential elections since the 1960s have offered ample material for political scientists and political sociologists who have contended that the stable class politics of the industrial era—roughly 1932–1964 (or maybe as late as 1976)—gave way to new, “postmaterial” politics (e.g., Abramson and Inglehart 1995; Clark and Lipset 1991; Inglehart 1977, 1990; Lipset 1981). Scholars who subscribe to this view point to newer cleavages based on gender, identity, concern for the environment, and family values, which have, they argue, displaced class from its central place in U.S. politics. When Democrats appeal to middle-class voters and the British Labour party touts a “third way” in twenty-first century politics, some of these ideas ring true.

Other scholars have countered the postmaterialists. First, they note that although elections are zero-sum—an individual’s vote for one candidate precludes her vote for another in the same election—cleavages are not. Voters can choose candidates based on a single dimension of their social identity, whether it

be class, gender, religion, or something else, or they can balance a variety of considerations. Nothing in the political arithmetic of social cleavages mandates that a rise in the weight given to one factor lowers the weight given to others. In fact, research has shown that, to a first approximation, all the factors that were important for voting in U.S. presidential elections in the 1960s were more important in the 1990s (Greeley and Hout 2006, ch. 3). Second, the empirical evidence contradicts the postmaterialists. In Britain, class voting fluctuates without trend (Goldthorpe 1999; Heath, Jowell, and Cur-tice 1985). In the United States, classes re-aligned (Hout, Brooks, and Manza 1995; Hout, Manza, and Brooks 1999), and the effect of income on the vote increased (McCarty, Poole, and Rosenthal 2005).

Class Voting and Class Politics

Interest in class voting goes back to the dawn of modern understandings of class; that is, back to the grand theorists of the nineteenth century. The roots of contemporary debates

about class voting, however, are planted in data, not in class theory. The national election surveys that have accumulated in many countries since the 1960s and the World Values Surveys that began in the 1970s yielded a harvest of class-voting time series. Findings through the early 1970s showed a strong, if variable, relationship between class (as indicated by occupation) and voter's choices of candidates and parties (Alford 1963; Lipset 1960, 1981; Rose 1974). Lipset and Rokkan's (1967) influential synthesis of the empirical record pointed to two nineteenth-century revolutions—the National Revolution and the Industrial Revolution. Lipset and Rokkan argued that these two revolutions initiated everywhere processes of social differentiation and conflict. As democracy spread, the four axes of social differentiation—(1) church versus secular state, (2) dominant versus subject subcultures, (3) rural areas versus cities, and (4) employers versus workers—came to influence voting behavior. The details varied from country to country, depending on history, the presence and influence of churches and minority subcultures, and timing. But everywhere the advance of industrialization pushed the first three axes to the side and the fourth—class—to the fore.

In many Western European nations (and Australia) the cleavage structure was “frozen” in the dominance of class-based political parties (see Bartolini and Mair 1990; Lipset and Rokkan 1967; Mair 1999; and Rose and Urwin 1970). Where this happened, it makes sense to speak of “class politics” (Mair 1999). Elsewhere the association between class and voting behavior is better thought of as “class voting.” Without a party structure that freezes both classes and voters in place, there is the prospect of shifting alliances, new coalitions, and realignment. In the United States, the lack of explicit links between classes and parties frees voters to weigh each party's appeals and frees each party to revise its strategies from time to time.

This process has been explicit in U.S. presidential elections since 1992. First Bill Clinton and then John Kerry dropped the Democrats' familiar appeals to “working people” in favor of direct calls to the middle class. Al Gore made more traditional class appeals in his campaign, and the Republicans accused him of “inciting class warfare.” At the same time, then-Governor Bush addressed the attendees at a high-priced fund-raiser as the “haves and the have-mores,” wisecracking that “some people call you the elites; I call you my base” (featured in Michael Moore's *Fahrenheit 911*). It is well known in the United States that many blue-collar workers identify with the middle class (Halle 1984). And, in any event, the working class is too small in the post-industrial United States to be an effective political base (Brooks and Manza 1999).

Even without a labor party, unions used to bind blue-collar Americans to the Democratic Party. From 1936 to 1968 (and again in 1984) Democrats gave unions an important voice in the selection of candidates. In exchange, the unions delivered their share of the working-class vote to the Democrats. That link was snapped by electoral reforms in the 1970s. The popular party primary system of selecting candidates moved politics from back rooms to TV screens. And unions lost out. They were, at the same time, losing membership and the trust of members and nonmembers alike (Hout, Manza, and Brooks 1999; Lipset and Schneider 1983). In the age of Jimmy Hoffa, identification with unions became more of a political liability than an asset, and unions never regained the clout they once had at the ballot box. Union members still lean strongly toward Democratic candidates (Hout, Manza, and Brooks 1999), but the heyday of identification between the Democrats and the labor movement is over.

The erosion of confidence in unions along with Republican appeals for “the right to work” and less industrial regulation seem

to have had the greatest impact on the very people who had the most to lose in the new economy that emerged during the 1970s. As energy shocks and stagflation (inflation plus high unemployment) buffeted the economy and the Democrats seemed unable to deliver the economic prosperity and security that had been their hallmark midcentury, Republican candidates' promises of relief from regulations and taxes began to seem more appealing. In particular, lower taxes and less regulation had the greatest influence on small businesspeople and manual workers who were not in unions. To these people, Reagan's firing of striking air-traffic controllers in 1981 may have seemed like a dramatic rebuke of special pleaders who wanted to dodge the apparently inexorable economic changes that everyone else was facing. If we are reading these signs correctly, then the persistence of union membership as a factor in U.S. politics is as likely a reflection of a rightward tilt among people without union protection as it is a reflection of a leftward stance among those with it.

Following Hout, Brooks, and Manza (1995) and Hout, Manza, and Brooks (1999), we distinguish between "traditional" and "total" class voting and use statistical models appropriate to each. Traditional class voting hinges on the correspondence between the working class and parties of the left and the middle classes and parties of the center or right. This correspondence can be embedded in statistical models of class voting that identify a "natural" party for each class (Weakliem 1995). A slightly weaker version of the traditional approach arrays classes and parties as ordered points on latent continua and examines the strength of association between the two latent variables (Weakliem and Heath 1999). However the researcher approaches the data, the traditional model is really only appropriate to understanding the historically significant but specific identification of classes and parties.

Total class voting, on the other hand, is not tied to any specific correspondence of classes and parties. That trend is useful and informative when traditional class voting alliances may not be the strongest or most evident. For example, the recent emergence of affinities between Democrats and teachers may be as strong and significant as the former links between Democrats and blue-collar workers. Throw in trial lawyers, college professors, nurses, and half of the doctors, and the Democrats' courting of professionals starts to yield more votes than traditional appeals to blue-collar workers could. A traditional approach would miss this kind of class voting; a total approach can capture it. As long as traditional class voting is but one of several patterns of association between classes and parties, the total approach is preferable for its flexibility. In previous work, the total approach allowed us to see the class *realignment* in U.S. presidential politics, whereas the traditional approach miscast the same trends as *dealignment* (see Hout, Manza, and Brooks 1999).

Another advantage of the total class voting approach is that it incorporates the class differences in turnout as well as class differences in votes. Middle-class people are more likely to vote than working- or lower-class Americans (Verba, Brady, and Schlossman 1996). The class skew in participation is likely to affect the party system and public policy (Piven and Cloward 1986; Rosenthal 2004). Setting it aside, as the traditional class voting approach does, misses an important element of class politics.

Our distinction between traditional and total class voting is related to Mair's (1999) distinction between "class politics" and "class voting." According to Mair, class politics signifies the kind of institutionalized class alliances that we call traditional. Class voting, for Mair, signifies a tendency for classes to ally themselves with parties in a more ad hoc endorsement for a specific election. The deal, far from institutionalized in party structure,

must be renegotiated each election cycle. Total class voting, as we have defined it, only requires class voting, not class politics.

The concept of traditional class voting is entrenched in the literature. It is unavoidable when class is conceived of or reduced to a dichotomy. Admitting to more than one class distinction (ergo more than two classes) opens the prospect of recombination. Contemporary theory elaborates several class distinctions, and history is full of examples of shifting class alliances. Thus the total class approach has the advantages of being both truer to theory and truer to history than the traditional class-voting approach is. Previous research (Hout, Brooks, and Manza 1995 and Hout, Manza, and Brooks 1999) amply demonstrated the superiority of the total class voting model in accounting for trends in voting patterns from 1948 to 1992.

Data and Methods

We use data from the American National Election Surveys (ANES), a time series that stretches back to the 1952 presidential election (with a question in 1952 about the 1948 election). The ANES is a stratified random sample of the English-speaking adult population with sample sizes varying from approximately 1,200 to 2,500 respondents per election.

The dependent variable for our analysis is self-reported vote for president recoded in four categories, (1) the Democratic candidate, (2) the Republican candidate, (3) some other candidate, and (4) did not vote. Self-reports exaggerate turnout but reproduce the partisan split very well (e.g., Abramson, Aldrich, and Rhode 1994).

We include election, class, gender, age, region, education, and racial ancestry as independent variables. Age and education are coded in years; the other variables are coded in standard, obvious ways using dummy variables. Our class scheme involves six substantive classes (as in Hout, Brooks, and Manza 1995) plus separate categories for housewives and retired per-

sons. The six substantive classes are (1) professionals, (2) managers, (3) routine white-collar workers (mostly clerical, sales, and white-collar service workers), (4) self-employed (other than professional), (5) skilled blue-collar workers, and (6) less-skilled blue-collar workers (including blue-collar service workers).

We test for changes across elections in the association between vote and the independent variables by generating interaction effects between the election and each of the independent variables. For statistical efficiency, we exclude interaction effects that fail to reach the conventional significance level (.05) from further calculations. For interaction effects involving class, Hout, Brooks, and Manza (1995) also investigated two intermediate forms, again to improve the statistical efficiency of tests. The first idea is the uniform differences model (see Hout, Brooks, and Manza 1995 for details); the second is a simple linear change. Both of these dramatically reduce the degrees of freedom used to detect change, and the second proved to be the preferred way of modeling class realignment. We fit the following, unrestricted model:

for $i = 1, \dots, N$ indexes persons; $j = 1, \dots, 4$ indexes outcomes (1 = Republican, 2 = Democrat, 3 = other, 4 = did not vote); T_j ,

$$\begin{aligned}
 y_{ij} &= \ln \left(\frac{p_{ij}}{p_{i1}} \right) \\
 &= \beta_0 + \sum_{k=1}^{14} \beta_k T_{ki} + \sum_{k=15}^{20} \beta_k X_{(k-14)i} \\
 &\quad + \sum_{k=21}^{P+20} \beta_k Z_{(k-20)i} + \sum_{k=1}^6 \sum_{k'=1}^{14} \gamma_{kk'} X_{ki} T_{k'i} \\
 &\quad + \sum_{k=1}^P \sum_{k'=1}^{14} \delta_{k,k'} Z_{ki} T_{k'i}
 \end{aligned}$$

$\dots T_{14}$ are dummy variables for elections, $X_1 \dots X_6$ are dummy variables for classes, $Z_1 \dots Z_P$ are dummy variables for other independent variables, $\beta_0 \dots \beta_{P+21}$, $\gamma_1 \dots \gamma_6$, and $\delta_{1,1} \dots \delta_{14,P}$ are parameters to be estimated. For identification, we constrain

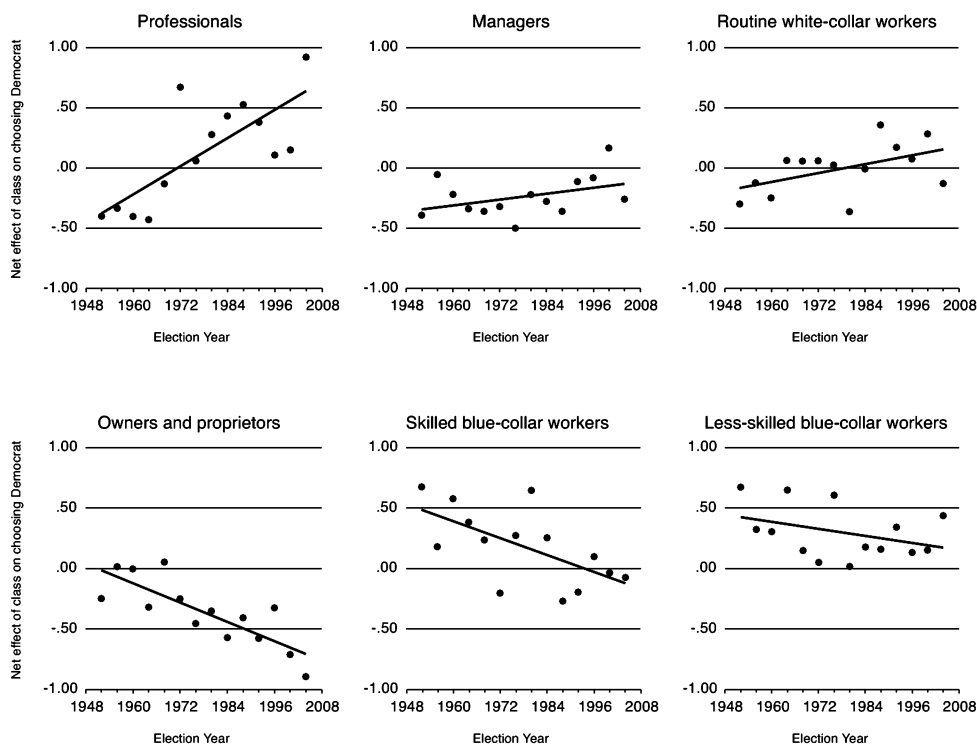


Figure 1. Net Effect of Class on Partisan Choice by Election, Class, and Model: Experienced Civilian Labor Force, 1948–2004

Note: Points and lines in these graphs represent multinomial logistic regression coefficients, normed to sum to zero within an election year. The models control for the effects of gender, race, region, age, education, and statistically significant changes in those effects.

$\sum_{k=1}^{14} \beta_k = \sum_{k=15}^{20} \beta_k = \sum_{k=1}^6 \sum_{k'=1}^{14} \gamma_{kk'} = 0$. For our graph, we obtained the points by summing the $\beta_k + \gamma_{kk'}$ for all possible combinations of $k = 1, \dots, 6$ and $k' = 1, \dots, 14$; to that we added a linear trend line (fit by unweighted OLS).

The Realignment of Class Voting

The traditional pattern of class voting held in U.S. presidential elections from 1948 to 1960. With the exception of the self-employed, middle-class voters supported Republican candidates while working-class voters supported the Democrats. The self-employed split their votes between the two major parties. After 1964 professionals shifted rapidly toward the Democrats; routine white-collar workers followed at a slower pace. At about

the same time, the self-employed and skilled blue-collar workers shifted toward the Republicans. For the self-employed this trend amounted to a shift from ambivalence to strong Republican support. For the skilled blue-collar workers, it was a shift from being part of the Democrats' base to ambivalence about the two parties by 1988. Less-skilled blue-collar workers moved about as much in the Republican direction as routine white-collar workers moved in the Democratic direction.

We summarize these trends in Figure 1. The dots show the year-by-year votes in detail; the lines show the expected percentages under the linear change model in the equation above. Statistical tests indicate that the lines fit the data well at conventional significance levels.

The most significant feature of these trends for political sociology is the realignment of the professions since the 1960s. This trend was first identified by Hout, Brooks, and Manza (1995) using data from 1948 to 1992. Brooks and Manza (1997) and Brooks (2000) show convincing evidence that professionals' rising commitments to civil rights for blacks (and maybe for other groups) prompted their shift from right to left. Professionals, when picking candidates, began to give greater weight to issues of equal opportunity than they had in the past, even bypassing their personal economic interests to do so. In extending the analysis onward by three more elections, we find that the trend that was already significant by 1992 continued into the 1996, 2000, and 2004 elections. Comparing the dots and the lines in Figure 1, we see that the Democrats got less support from professionals in 1996 and 2000 but recouped it in 2004.

In one sense the triumph of civil rights movements (along with those for civil liberties and the environment) over personal material interests is a point in favor of postmaterialist accounts of electoral change. But the postmaterialists assert that class no longer correlates with vote. The correlation is evident here; it is just nontraditional (although it is by now forty years old). Here we see one class moved from one party to another in response to these rising concerns and did so as a class. Furthermore, as we move through the rest of the class structure, we will see much more evidence of "traditional," that is, self-interested, class voting.

Managers voted in recent elections more or less as they did in the 1960s. They changed their attitudes toward civil rights through the 1970s, but their changing attitudes failed to influence their votes. Their personal interests (and perhaps the interests of the firms they managed) in cutting taxes and deregulating the economy proved to be more crucial to how they voted. Over time, the voting pattern of the self-employed became more and more conservative, eventually surpassing even

managers' preference for Republicans. The self-employed were, in the 1950s, evenly divided between Democrats and Republicans. They moved strongly toward the Republicans and were their staunchest supporters by 1976 or perhaps 1980. We suspect that some of the change reflects issues and some reflects a shift from farming to small business within the ranks of the self-employed.

Skilled blue-collar workers moved from strong Democratic support to an even split between Republicans and Democrats (or perhaps a slight preference for Republicans, although their partisan split does not pass the statistical test for being significantly different from even). The evidence through 1992 was unclear on this point. Hout, Brooks, and Manza (1995) wrote of "volatility." The continued drift of the upper blue-collar voters in the last three elections makes it clear that they are moving toward the Republicans. Disaffection with the welfare state, a preference for a tougher approach to criminal justice, and support for the military are possible explanations for this shift.

Change among less-skilled blue-collar workers is more subtle but also, overall, favors Republican candidates. It is offset by a modest shift toward Democrats among routine white-collar workers.

All of these changes might suggest that total class voting has changed significantly. That is not the case, though. The summary measure of total class voting, *kappa*, introduced by Hout, Brooks, and Manza (1995), shows no net change from 1948 to 2004. Figure 2 shows that there are some up and down cycles in its value over a span of a few elections, but the experience of the 1996–2004 elections added here does not alter the conclusion Hout, Brooks, and Manza (1995) reached regarding 1952–1992—the long-term trend line is flat.

Conclusions

U.S. class coalitions realigned after the civil rights movement in the 1960s. Voters changed

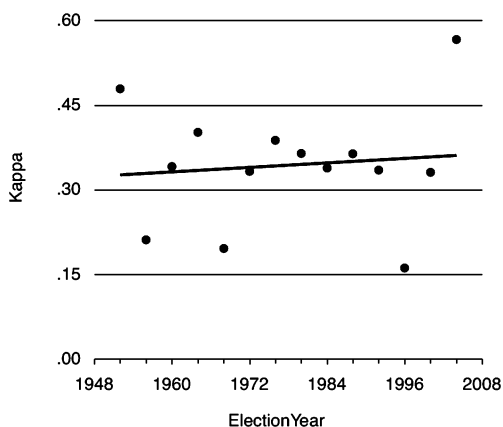


Figure 2. Total Effect of Class on Partisan Choice by Election: Experienced Civilian Labor Force, 1948–2004

what they expect from presidential politics and parties changed what they promise. The Republican platform of lower taxes and deregulation deepened the loyalty of managers and won over the self-employed workers who used to back the Democrats. Blue-collar workers also prefer low taxes. Republicans courted and won their votes in most elections since 1972. But Democrats countered with strong plays for professionals' interests in many kinds of social spending and their commitment to civil rights. "New Democrats"—most notably Clinton—appealed directly to middle-class voters and shored up support among professionals by balancing the budget in the 1990s. Overall, the Republicans gained and the Democrats lost among employed people. But the Democrats' defense of Social Security and Medicare allowed them to gain support among retirees and pensioners—people out of the labor force and, thus, out of this analysis. These changes helped balance the political arithmetic.

We do not have enough evidence to know if the class-voting realignment in U.S. politics is a general trend that extends to other Western democracies or another instance of "American exceptionalism." We note that English and

German parties have started to adopt some American-style appeals. In Britain, "New Labour's" platform is known as "the third way"; Schröder moved the German Social Democrats in that direction, too. The statistical evidence makes us cautious. In Blair's first win, he drew proportionally higher support in each class, leading to no change in the association between class and vote in Britain (Goldthorpe 1999).

In this chapter we have shown that American social classes realigned: professionals replaced skilled blue-collar workers as the class base of the Democratic Party, whereas shopkeepers and proprietors joined managers as the class base of the Republican Party. This historical realignment coupled with no net change in total class voting shows how class can be important for elections without lining up in the traditional way.

REFERENCES

- Abramson, P. R., and R. Inglehart. 1995. *Value Change in Global Perspective*. Ann Arbor: University of Michigan Press.
- Abramson, P. R., J. H. Aldrich, and D. W. Rhode. 1994. *Change and Continuity in the 1992 Elections*. Washington, DC: Congressional Quarterly Press.
- Alford, R. 1963. *Party and Society: The Anglo-American Democracies*. Westport, CT: Greenwood Press.
- Bartolini, S., and P. Mair. 1990. *Identity, Competition, and Electoral Availability: The Stabilization of European Electorates, 1885–1985*. Cambridge: Cambridge University Press.
- Brooks, C. 2000. "Civil Rights Liberalism and the Suppression of a Republican Political Realignment in the United States, 1972–1996." *American Sociological Review* 65: 483–505.
- Brooks, C., and J. Manza. 1997. "Class Politics and Political Change in the United States, 1952–1992." *Social Forces* 76: 397–408.
- Clark, T. N., and S. M. Lipset. 1991. "Are Social Classes Dying?" *International Sociology* 6: 397–410.
- Goldthorpe, J. H. 1999. "Modelling the Pattern of Class Voting in British Elections, 1964–1992." In Geoffrey Evans, ed., *The End of Class Politics?*, pp. 59–83. Oxford: Oxford University Press.
- Greeley, Andrew M., and Michael Hout. 2006. *The Trust About Conservative Christians: What They*

- Believe and What They Think*. Chicago: University of Chicago Press.
- Halle, D. 1984. *America's Working Man*. Chicago: University of Chicago Press.
- Heath, A. F., R. Jowell, and J. Curtice. 1985. *How Britain Votes*. Oxford: Pergamon Press.
- Hout, M., C. Brooks, and J. Manza. 1995. "The Democratic Class Struggle in U.S. Presidential Elections." *American Sociological Review* 60: 805–828.
- Hout, M., J. Manza, and C. Brooks. 1999. "Classes, Unions, and the Realignment of U.S. Presidential Voting, 1952–1992." In Geoffrey Evans, ed., *The End of Class Politics?* pp. 83–95. Oxford: Oxford University Press.
- Inglehart, R. 1977. *The Silent Revolution: Changing Values and Political Styles Among Western Publics*. Princeton: Princeton University Press.
- _____. 1990. *Culture Shift*. Princeton: Princeton University Press.
- Lipset, S. M. 1960. *Political Man*. New York: Doubleday.
- _____. 1981. *Political Man*. 2nd, expanded ed. Baltimore: Johns Hopkins University Press.
- Lipset, S. M., and S. Rokkan. 1967. "Cleavage Structures, Party Systems, and Voter Alignments." In S. M. Lipset and S. Rokkan, eds., *Party Systems and Voter Alignments*, pp. 1–64. New York: Free Press.
- Lipset, S. M., and W. S. Schneider. 1983. *The Confidence Gap: Business, Labor, and Government in the Public Mind*. New York: Free Press.
- Mair, P. 1999. "Critical Commentary." In Geoffrey Evans, ed., *The End of Class Politics?* pp. 308–312. Oxford: Oxford University Press.
- McCarty, N., K. T. Poole, and H. Rosenthal. 2005. "Political Polarization and Income Inequality." Department of Politics, Princeton University Web site, available at <http://www.princeton.edu/~nmccarty/ineqpold.pdf>.
- Piven, F. F., and R. A. Cloward. 1986. *Why Americans Don't Vote*. New York: Pantheon.
- Rose, R., ed. 1974. *Electoral Behavior: A Comparative Handbook*. New York: Free Press.
- Rose, R., and D. Urwin. 1970. "Persistence and Change in Western Party Systems since 1945." *Political Studies* 18: 287–319.
- Rosenthal, H. 2004. "Politics, Public Policy, and Inequality: A Look Back at the 20th Century." In Kathryn Neckerman, ed., *Social Inequality*, 861–892. New York: Russell Sage Foundation.
- Verba, S., H. Brady, and K. Schlossman. 1996. *Participation and Equality*. Cambridge, MA: Harvard University Press.
- Weakliem, D. 1995. "Two Models of Class Voting." *British Journal of Political Science* 25: 254–270.
- Weakliem, D., and A. F. Heath. 1999. "The Secret Life of Class Voting: Britain, France, and the United States since the 1930s." In Geoffrey Evans, ed., *The End of Class Politics?* pp. 97–136. Oxford: Oxford University Press.

PART IX

The Future of Inequality

Theories of Industrialism and Modernity

100 ■ Clark Kerr, John T. Dunlop, Frederick H. Harbison,
and Charles A. Myers

Industrialism and Industrial Man

The Logic of Industrialization

Although industrialization follows widely differing patterns in different countries, some characteristics of the industrialization process are common to all. These “universals” arise from the imperatives intrinsic to the process. They are the prerequisites and the concomitants of industrial evolution. Once under way, the logic of industrialization sets in motion many trends which do more or less violence to the traditional pre-industrial society.

In the actual course of history, the inherent tendencies of the industrial process are never likely to be fully realized. The pre-existing societies and conditions shape and constrain these inherent features. The leaders of economic development influence the directions and the rate of industrial growth; and the existing resources and the contemporaneous developments in other countries are also likely to affect actual events. These influences will be important in every case of industrialization. They do not, however, deny the validity of searching for some of the fundamental directions in which industrialization will haul and pull. Indeed, an understanding

of these tendencies is requisite to a full appreciation of the influence of historical, cultural, and economic factors on the actual course of industrialization.

What, then, are some of the imperatives of the industrialization process? Given the character of science and technology and the requirements inherent in modern methods of production and distribution, what may be deduced as to the necessary or the likely characteristics of workers and managers and their interrelations? What are the inherent implications of industrialization for the work place and the larger community? What, in sum do the actual histories of those societies with either brief or more extensive industrializing experience suggest about the principal forces implicit in an “industrial revolution”?

The Industrial Work Force

[The workman] becomes an appendage of the machine, and it is only the most simple, most monotonous, and most easily acquired knack, that is required of him. (*Manifesto*, p. 65). . . . Hence, in the place of the hierarchy of specialized workmen that characterizes manufacture, there steps, in the automatic factory, a tendency

to equalise and reduce to one and the same level every kind of work that has to be done by the minders of the machines . . . (*Capital*, p. 420). The more modern industry becomes developed, the more is the labour of men superseded by that of women (*Manifesto*, p. 66) . . . The various interests and conditions of life within the ranks of the proletariat are more and more equalised, in proportion as machinery obliterates all distinctions of labour, and nearly everywhere reduces wages to the same low level (*Manifesto*, p. 69). . . . The essential division is, into workmen who are actually employed on the machines (among whom are included a few who look after the engine), and into mere attendants (almost exclusively children) of these workmen. . . . In addition to these two principal classes, there is a numerically unimportant class of persons, whose occupation is to look after the whole of the machinery and repair it from time to time such as engineers, mechanics, joiners, etc. (*Capital*, p. 420).¹

These quotations from the *Manifesto* and *Capital* envisage as the consequence of capitalist production the destruction of the hierarchy of specialized workmen in pre-industrial society and the subsequent leveling of skill, a minor number of skilled labor, engineers, and managers, and the use of women and children for a growing number of unskilled tending and feeding jobs. The historical evidence of the past century, however, suggests a quite different pattern of evolution for the industrial work force.

Industrialization in fact develops and depends upon a concentrated, disciplined industrial work force—a work force with new skills and a wide variety of skills, with high skill levels and constantly changing skill requirements.

The industrialization process utilizes a level of technology far in advance of that of earlier societies. Moreover, the associated scientific revolution generates continual and rapid changes in technology which have decisive consequences for workers, managers, the state, and their interrelations.

The industrial system requires a wide range of skills and professional competency broadly distributed throughout the work force. These specialized human resources are indispensable to the science and technology of industrialism, and their development is one of the major problems of a society engaged in industrialization. The absence of a highly qualified labor force is as serious an impediment as a shortage of capital goods. The professional, technical, and managerial component of the labor force is particularly strategic since it largely carries the responsibility of developing and ordering the manual and clerical components.

Mobility and the Open Society. The dynamic science and technology of the industrial society creates frequent changes in the skills, responsibilities, and occupations of the work force. Some are made redundant and new ones are created. The work force is confronted with repeated object lessons of the general futility of fighting these changes, and comes to be reconciled, by and large, to repeated changes in ways of earning a living. But there may be continuing conflict over the timing of change and the division of the gains. The industrial society requires continual training and retraining of the work force; the content of an occupation or job classification is seldom set for life, as in the traditional society. Occupational mobility is associated with a high degree of geographical movement in a work force and with social mobility in the larger community both upwards and downwards.

One indication of the extent of occupational shifts which occur in the course of industrialization is reflected in Table 1 for the United States in the period 1900–1960.

Industrialization tends to produce an open society, inconsistent with the assignment of managers or workers to occupations or to jobs by traditional caste, by racial groups, by sex, or by family status. There is no place for the extended family; it is on balance an

TABLE I
Occupational Distribution of Employment, United States, 1900-1960 (thousands of people)
15 years of age and over

| | 1900 | 1960 | Change 1900-1960 | % Total Employment |
|---|--------|---------|---------------------|-----------------------|
| Total | 56,075 | 290,690 | 177,129 | 100 |
| Total Farm Employment | 11,007 | 10,498 | -509 | 1.9 |
| Total Non-Farm Employment | 45,068 | 180,192 | 135,124 | 98.1 |
| White-Collar Workers | | | 112,197 | 38.6 |
| Professional, Technical, and Managerial Workers | | | 46,187 | 15.9 |
| Managers, Officials, and Proprietors | | | 41,705 | 14.4 |
| Clerical and Kindred Workers | | | 24,105 | 8.3 |
| Sales Workers | | | 19,999 | 6.9 |
| Blue-Collar Workers | | | 113,879 | 38.7 |
| Craftsmen, Foremen, and Kindred Workers | 8,908 | 1,997 | -6,911 | 2.4 |
| Operatives and Kindred Workers | 12,989 | 11,111 | -1,878 | 4.2 |
| Laborers | 1,689 | 1,179 | -510 | 0.4 |
| Service Workers | | | 49,009 | 16.7 |
| Service Workers, Excepting Household Workers | 1,779 | 1,947 | 168 | 0.3 |
| Household Workers | 2,207 | 1,379 | -828 | 0.3 |

**P/u from
Social Strat. 2e,
p. 795**

Source: U.S. Department of Labor, Bureau of Labor Statistics, *Industrial Structure of Commerce, 1900 and 1960*; U.S. Department of Commerce, *Occupation in the United States*, Bureau of Economic Warfare Paper No. 7 (Washington, 1948), pp. 6-7.

impediment to requisite mobility. The primary family provides a larger and more mobile labor force. The function of the family under industrialism is constricted: it engages in very little production; it provides little, if any, formal education and occupational training; the family business is substantially displaced by professional management. “. . . economic growth and a transference of women’s work from the household to the market go closely hand in hand.”² In the industrial society the primary family is largely a source of labor supply, a unit of decision-making for household expenditures and savings, and a unit of cultural activity.

This society is always in flux and in motion. As a result of its science and technology, it is continuously rearranging what people do for a living, where they work and where they live, and on what they spend their incomes. Their children come to expect to live different lives from their parents. But mobility in the industrial society is not random; it comes to be organized and governed by a complex of rules of the work community.

Education—The Handmaiden of Industrialism. Industrialization requires an educational system functionally related to the skills and professions imperative to its technology. Such an educational system is not primarily concerned with conserving traditional values or perpetuating the classics; it does not adopt a static view of society, and it does not place great emphasis on training in the traditional law. The higher educational system of the industrial society stresses the natural sciences, engineering, medicine, managerial training—whether private or public—and administrative law. It must steadily adapt to new disciplines and fields of specialization. There is a relatively smaller place for the humanities and arts, while the social sciences are strongly related to the training of managerial groups and technicians for the enterprise and the government. The increased leisure time of industrialism, however, can afford a broader public appreciation of the humanities and the arts.

As in all societies, there is debate over what the youth is to be taught. The largest part of

the higher educational system tends to be specialized and designed to produce the very large volume of professionals, technicians, and managers required in the industrial society. There is a case for some degree of generality in the educational system because of the rapidity of change and growth of knowledge during the course of a career. A technically trained work force needs to be able to follow and to adapt to changes in its specialties and to learn to shift to new fields. Generality is also requisite for those coordinating and leading the specialists.

The industrial society tends to create an increasing level of general education for all citizens, not only because it facilitates training and flexibility in the work force, but also because, as incomes rise, natural curiosity increases the demand for formal education, and education becomes one of the principal means of vertical social mobility. It will be observed that the industrial society tends to transform drastically the educational system of the pre-industrial society. Further, the high level of technical and general education requisite to the industrial society cannot but have significant consequences for political life. The means of mass communication play a significant role both in raising standards of general education and in conditioning political activity and shaping political control.

The Structure of the Labor Force. The labor force of the industrial society is highly differentiated by occupations and job classifications, by rates of compensation, and by a variety of relative rights and duties in the work place community. It has form and structure vastly different from the more homogeneous labor force of the traditional society.

The variety of skills, responsibilities, and working conditions at the work place requires an ordering or a hierarchy. There are successive levels of authority of managers and the managed, as well as considerable specialization of function at each level of the hierarchy. Job evaluation and salary plans symbol-

ize the ordering of the industrial work force by function and compensation.³

The work force in the industrial society is also structured in the sense that movement within the work community is subjected to a set of rules; hiring, temporary layoffs, permanent redundance, promotions, shift changes, transfers, and retirement are applied to individual workers and managers according to their position, station, seniority, technical competency, or some other measure of status in a group rather than in random fashion. Not all jobs are open at all times to all bidders. The ports of entry into an enterprise are limited, and priorities in selection are established. Movement tends to be relatively easier within than among job families in an enterprise. Delineation of job families may also vary and will often depend upon whether movement involves promotion, layoff, or transfer.

The industrial system changes the hours of work that prevail in predominantly agricultural societies. The silent night of pre-industrial society yields to the sometimes insistent requirements of continuous operations. The work force is geared to shift operations and the community to a changed attitude toward working at night. Even the holidays and religious days of the traditional society do not escape transformation.

Scale of Society

The technology and specialization of the industrial society are necessarily and distinctively associated with large-scale organizations. Great metropolitan areas arise in the course of industrialization. The national government machinery expands significantly. Economic activity is carried on by large-scale enterprises which require extensive coordination of managers and the managed. A wide variety of rules and norms are essential to secure this coordination.

Urban Dominance. The industrial society is an urban society, concentrated in metropolitan areas with their suburbs and satellite

communities. While substantial cities have arisen in pre-industrial societies as commercial and religious centers,⁴ urban ways come to permeate the whole of industrial society. Rapid means of transportation and mass communication reduce the variance of subcultures, particularly those based on geography and the contrast between farm and city.

In the industrial society agriculture is simply another industry; it is not a “way of life” to be preserved for its own value or because it constituted a traditional and antecedent form of society. Agricultural units of production (farms) tend to be specialized according to products, and the general farm, substantially self-sufficient, has little place. Indeed, the proportion of the work force engaged in agriculture is a rough index of the degree of industrialization of a society.⁵ “To the economic eye a community which needs to have the majority of its people working on the land is merely demonstrating its inefficiency.”⁶

Table 2 shows the relative role of agriculture in a number of countries; it reflects the low proportion of agriculture in the economically advanced countries and the high percentage of agricultural population in the less economically advanced societies.

Industrialization tends to promote the values, folkways, and heroes of the city and to weaken those of the farm. Even the art and music of the highly industrialized society can be expected to be substantially different from that of the pre-industrial society.

Large Role for Government. The industrial society is necessarily characterized by a substantial range and scale of activities by the government. In a society of advanced technology there are a larger number of activities for government; for instance, roads and highways, airports, the regulation of traffic, radio and television. Urban development has the same consequences. Technology also creates a more complex military establishment, extending in

TABLE 2
Agriculture and Population

| | Percentage of Total Population |
|---------------------|--------------------------------|
| Canada | 1.0 |
| United Kingdom 1951 | 1.0 |
| United States 1950 | 1.0 |
| Sweden 1950 | 3.0 |
| Switzerland 1950 | 7.0 |
| France 1951 | 8.0 |
| Italy 1951 | 12.0 |
| Japan 1951 | 14.0 |
| USSR 1948 | 14.0 |
| Greece 1950 | 18.0 |
| Spain 1941 | 21.0 |
| India | 21.0 |
| Iran 1950 | 25.0 |
| Algeria 1951 | 30.0 |
| Norway 1948 | 30.0 |
| Nepal 1950 | 40.0 |

**P/u from
Social Strat. 2e,
p. 797**

Source: American Council on Education, *International Population Review*, 1950, p. 98. Population figures are in millions. The University of Chicago Press, 1950, pp. 11-12.

many directions the activities of government. The more integrated character of the world increases governmental activities in the area of international relations. The scale of some scientific applications and the capital needs of new technologies, such as atomic energy development or space exploration, increase the scope of public agencies. As income rises, the demand of consumers may be for services largely provided by governments, such as education, parks, roads, and health services.

The role of government in countries entering upon industrialization, regardless of political form, may therefore be expected to be greater than before. There is wisdom in the observation: “. . . it is extremely unlikely that the highly modernized systems of the world today could have developed indigenously on the basis of any system other than ones that relied very heavily indeed on private individual operations, and it is extremely unlikely that latecomers can carry out such development without relying very heavily on public operations.”⁷

The industrial society and individual freedom, however, are not necessarily to be re-

garded as antagonists. A high degree of discipline in the work place and a large range of governmental activities is fully consistent with a larger freedom for the individual in greater leisure, a greater range of choice in occupations and place of residence, a greater range of alternatives in goods and services on which to use income, and a very wide range of subgroups or associations in which to choose participation. It is a mistake to regard the industrial society as antithetical to individual freedom by citing ways in which the scope of public and private governments has increased without also noting ways in which industrialization expands individual freedom.

The Web of Rules. The production of goods and services in the industrial society is largely in the hands of large-scale organizations. They consist of hierarchies composed of relatively few managers and staff advisers and a great many to be managed. The managers and the managed are necessarily connected by an elaborate web of rules that is made the more intricate and complex by technology, specialization, and the large scale of operations.

At any one time, the rights and duties of all those in the hierarchy must be established and understood by all.

The web of rules of the work place concerns compensation, discipline, layoffs, transfers and promotions, grievances, and a vast array of matters, some common to all work places and others specialized for the type of activity—factory, airline, railroad, mine, or office—and to the specific establishment. The rules also establish norms of output, pace, and performance. Moreover, the web of rules is never static, and procedures arise for the orderly change of these rules. The industrial system creates an elaborate “government” at the work place and in the work community. It is often observed that primitive societies have extensive rules, customs, and taboos, but a study of the industrial society reflects an even greater complex and a quite different set of detailed rules.

The web of rules depends partially on those technological features and market or budgetary constraints of the work place which are generally common to all types of industrializing countries, and partially on the particular resources and the political and economic forms of the country. The relative strength of these factors, and their mode of interaction, is important to an understanding of any particular industrial society. Cultural and national differences are less significant to the substantive web of rules, the more a country has industrialized. The impact of cultural and national heritage is more clearly discerned in the differences to be found in the process for formulating and promulgating the rules affecting men at work than in the content of the rules themselves.

The tug of industrialization—whatever these initial differences—is toward a greater role for the state in an eventual pluralistic rule-making system. The state does not evolve simply as a class apparatus and instrument for the oppression of another class, as Marx asserted.⁸ Nor does it “wither away” in the ultimate “good society.”⁹ Governments have a significant role in determining the substantive rules of the work community or in establishing the procedures and responsibilities of those with this power. In the highly industrialized society, enterprise managers, workers, and the government tend to share in the establishment and administration of the rules. The industrial relations system of the industrial society is genuinely tripartite.

Consensus in Society

The industrial society, like any established society, develops a distinctive consensus which relates individuals and groups to each other and provides an integrated body of ideas, beliefs, and value judgments. Various forms of the industrial society may create some distinctive ideological features, but all industrialized societies have some common values.

In the industrial society science and technical knowledge have high values, and scientists

and technologists enjoy high prestige and rewards.

Taboos against technical change are eliminated, and high values are placed on being “modern” and “up-to-date,” and in “progress” for their own sake.

Education also has a high value because of the fundamental importance of science and the utility of education as a means of social mobility.

The industrial society is an open community encouraging occupational and geographic mobility and social mobility. Industrialization calls for flexibility and competition; it is against tradition and status based upon family, class, religion, race, or caste.

It is pluralistic, with a great variety of associations and groups and of large-scale operations; the individual is attached to a variety of such groups and organizations.

Goods and services have a high value in the industrial society, and the “demonstration effect” is very strong on the part of individuals and groups seeking to imitate the standards of those with higher income levels.

The work force is dedicated to hard work, a high pace of work, and a keen sense of individual responsibility for performance of assigned norms and tasks. Industrial countries may differ with respect to the ideals and drives which underlie devotion to duty and responsibility for performance, but industrialization requires an ideology and an ethic which motivate individual workers. Strict supervision imposed on a lethargic work force will not suffice; personal responsibility for performance must be implanted within workers, front-line supervisors, and top managers.¹⁰

It is not by accident that the leaders of industrializing countries today exhort their peoples to hard work. “This generation is sentenced to hard labor” (Nehru). “We shall march forward as one people who have vowed to work and to proceed on a holy march of industrializing . . .” (Nasser). “The chief preoccupation of every Communist regime between the Elbe and the China Sea

is how to make people work; how to induce them to sow, harvest, mine, build, manufacture and so forth. It is the most vital problem which confronts them day in, day out, and it shapes their domestic policies and to a considerable extent their attitude toward the outside world.”¹¹ There are many counterparts for the Protestant ethic.

The Western tradition has been to harness the drive of individual self-interest; the communist method combines in varying proportions at varying times money incentives, devotion to a revolutionary creed, and the compulsion of terror. Regardless of means, industrialization entails a pace of work and an exercise of personal responsibility seldom known in economic activity in traditional societies.

The function of making explicit a consensus and of combining discrete beliefs and convictions into a reasonably consistent body of ideas is the task of intellectuals in every society. There are probably more intellectuals in the industrial society because of the higher levels of general education, income, and leisure. There are also new patrons to the intellectuals—the university, enterprise, labor organization, voluntary association and government—in place of the old aristocratic patrons. The function of formulating and restating the major values, premises, and consensus of a society from time to time, of reconciling the new industrial processes with the old order, plays a significant role in industrialization. The intellectuals accordingly are an influential group in the creation and molding of industrial society. . . .

The Road Ahead: Pluralistic Industrialism

Men attempt to peer ahead, to understand the structure of history, to alter the process of history, if possible, in accord with their preferences. The history of industrialization to date has not been a smoothly unilinear one; it has been uneven and multilinear. It

is likely that in the future it will continue to be both somewhat uneven and multilinear; and there will continue to be some latitude for choice and for chance. Chance may elude man, but choice need not; and the choice of men, within fairly broad limits, can shape history. To predict the future with any accuracy, men must choose their future. The future they appear to be choosing and pressing for is what might be called "pluralistic industrialism."

This term is used to refer to an industrial society which is governed neither by one all-powerful elite (the monistic model) nor by the impersonal interaction of innumerable small groups with relatively equal and fractionalized power (the atomistic model in economic theory). The complexity of the fully developed industrial society requires, in the name of efficiency and initiative, a degree of decentralization of control, particularly in the consumer goods and service trades industries; but it also requires a large measure of central control by the state and conduct of many operations by large-scale organizations.

As the skill level rises and jobs become more responsible, any regime must be more interested in consent, in drawing forth relatively full cooperation. For the sake of real efficiency, this must be freely given. The discipline of the labor gang no longer suffices. With skill and responsibility goes the need for consent, and with consent goes influence and even authority. Occupational and professional groups, of necessity, achieve some prestige and authority as against both the central organs of society and the individual members of the occupation or profession.

Education brings in its wake a new economic equality and a new community of political outlook. This in turn, along with many other developments, helps bring consensus to society. The harsh use of power by the state is no longer so necessary to hold society together at the seams. Education also opens the mind to curiosity and to inquiry, and the individual seeks more freedom to

think and to act. It brings a demand for liberty, and can help create conditions in which liberty can safely be assumed. It leads to comparisons among nations with respect to progress and participation.

Industrialism is so complex and subject to such contrary internal pressures that it never can assume a single uniform unchanging structure; but it *can* vary around a general central theme, and that theme is pluralism. While it will take generations before this theme will become universal in societies around the world, the direction of the movement already seems sufficiently clear:

The State that Does Not Wither Away. The state will be powerful. It will, at the minimum, have the responsibility for the economic growth rate; the over-all distribution of income among uses and among individuals; the basic security of individuals (the family formerly was the basic security unit); the stability of the system; providing the essential public services of education, transportation, recreational areas, cultural facilities, and the like; and the responsibility of providing a favorable physical environment for urban man.

In addition, any pluralistic society is subject to three great potential internal problems, and the state is responsible for handling each. One is the conflict among the various power elements in a pluralistic society. The state must set the rules of the game within which such conflict will occur, enforce these rules, and act as mediator; conflicts between managers and the managed are the most noticeable, but by no means the only ones. Another is the control of collusion by producers against consumers, by any profession against its clients, and by labor and management against the public. Undue aggrandizement of sectional interests is always endemic if not epidemic in a pluralistic society; in fact, one of the arguments for monism and atomism alike is the avoidance of sectionalism. Additionally, the state will come generally, under pluralistic industrialism, to

set the rules relating members to their organizations—who may get in, who may stay in, what rights and obligations the members have, what the boundaries are for the activities of the organization, and so on. It will, almost of necessity, be against too much conflict among, or collusion between, or domination of the members by the subsidiary organizations in society.

All these responsibilities mean the state will never “wither away”; that Marx was more utopian than the despised utopians. The state will be the dominant organization in any industrial society. But it may itself be less than fully unitary. It may itself be subject to checks and balances, including the check of public acceptance of its current leadership and its policies.

The Crucial Role of the Enterprise—The Middle Class and the Middle Bureaucracy. The productive enterprise, whether private or public, will be a dominant position under pluralistic industrialism. It will often be large and it must always have substantial authority in order to produce efficiently. This authority will not be complete, for it will be checked by the state, by the occupational association, by the individual employee; but it will be substantial.

The distinction between the private and the public manager will decrease just as the distinction between the private and the public enterprise will diminish; and the distinction among managers will be more according to the size, the product, and the nature of their enterprises. The controlled market and the controlled budget will bring more nearly the same pressures on them. The private enterprise, however, will usually have more freedom of action than the public enterprise; but the middle class and the middle bureaucracy will look much alike.

Associated Man. The occupational or professional association will range alongside the state and the enterprise as a locus of power in

pluralistic industrialism; and there will be more occupations and particularly more professions seeking association. Group organizations around skill and position in the productive mechanism will be well-nigh universal. These organizations will affect output norms, comparative incomes, access to employment, and codes of ethics in nearly every occupational walk of life. Their containment within reasonable limits will be a continuing problem; and some of the groups will always seek to invade and infiltrate the government mechanisms which are intended to supervise them.

The Web of Rules. Uniting the state, the enterprise, and the association will be a great web of rules set by the efforts of all the elements, but particularly by the state. This web of rules will also relate the individual to each of these elements. In the contest over who should make the web of rules, the end solution will be that they will be made or influenced by more than one element; they will not be set by the state alone or by the enterprise alone or by the association alone. The web of rules will not equally cover all aspects of life.

From Class War to Bureaucratic Gamesmanship. Conflict will take place in a system of pluralistic industrialism, but it will take less the form of the open strife or the revolt and more the form of the bureaucratic contest. Groups will jockey for position over the placement of individuals, the setting of jurisdictions, the location of authority to make decisions, the forming of alliances, the establishment of formulas, the half-evident withdrawal of support and of effort, the use of precedents and arguments and statistics. Persuasion, pressure, and manipulation will take the place of the face-to-face combat of an earlier age. The battles will be in the corridors instead of the streets, and memos will flow instead of blood. The conflict also will be, by and large, over narrower issues than in

earlier times when there was real disagreement over the nature of and the arrangements within industrial society. It will be less between the broad programs of capital and labor, and of agriculture and industry; and more over budgets, rates of compensation, work norms, job assignments. The great battles over conflicting manifestos will be replaced by a myriad of minor contests over comparative details.

From Class Movement to Special Interest Group. Labor-management relations will conform to this new context. Labor organizations will not be component parts of class movements urging programs of total reform, for the consensus of a pluralistic society will have settled over the scene. Nor may they be very heavily identified by industry, particularly with the increasing multiplication and fractionalization of industries. Rather, they may tend to take more the craft, or perhaps better, the occupational form. With skills more diverse, at a generally higher level, and obtained more through formal education, and with geographical mobility greatly increased, professional-type interests should mean more to workers than industry or class ties.

The purpose of these occupational and professional associations will be relatively narrow, mostly the improvement of the status of the occupation in terms of income, prestige, and specification of the rights and duties that accompany it. Generally these organizations will be a conservative force in society, opposed to new ways of doing things, resistant to increased efforts by members of the occupation. The enterprise managers will be the more progressive elements in the society, although they too may become heavily weighted down by checks and balances and rules.

The techniques of the professional associations for achieving their ends will be those of the bureaucratic organization everywhere; a far cry from the individual withdrawal, or the guerilla warfare, or the strike or the polit-

ical reform movement of earlier times. They will constitute the quarrels between the semi-managed and the semi-managers.

Individuals will identify themselves more closely with their occupation, particularly if it involves a formal training period for entry, and mobility will follow more the lines of the occupation than the lines of the industry or the job possibilities of the immediate geographical area. In terms of identification, the orientation will be more nearly that of the member of a guild than of a class or of a plant community. Mayo will turn out to be as wrong as Marx. Just as the class will lose its meaning, so also will the plant community fail to become the modern counterpart of the primitive tribe. The occupational interest group will represent the employee in his occupational concerns and the occupation will draw his allegiance. Status in the tribe will not give way to status in the plant; nor will status have given way to the individual contract through the march of civilization; rather interest identification will take the place of both status and individual contract in ordering the productive arrangements of men.

Education, occupation, occupational organization will all be drawn together to structure the life-line and the economic interests of many if not most employees.

The New Bohemianism. The individual will be in a mixed situation far removed either from that of the independent farmer organizing most aspects of his own life or from that of the Chinese peasant in the commune under total surveillance. In his working life he will be subject to great conformity imposed not only by the enterprise manager but also by the state and by his own occupational association. For most people, any complete scope for the independent spirit on the job will be missing. However, the skilled worker, while under rules, does get some control over his job, some chance to organize it as he sees fit, some possession of it. Within

the narrow limits of this kind of “job control,” the worker will have some freedom. But the productive process tends to regiment. People must perform as expected or it breaks down. This is now and will be increasingly accepted as an immutable fact. The state, the manager, the occupational association are all disciplinary agents. But discipline is often achieved by a measure of persuasion and incentive. The worker will be semi-independent with some choice among jobs, some control of the job, and some scope for the effects of morale; but he will also be confined by labor organizations, pensions, and seniority rules, and all sorts of rules governing the conduct of the job.

Outside his working life the individual may have more freedom under pluralistic industrialism than in most earlier forms of society. Politically he can have some influence. Society has achieved consensus and the state need not exercise rigid political control. Nor in this “Brave New World” need genetic and chemical means be employed to avoid revolt. There will not be any rebellion, anyway, except little bureaucratic revolts that can be settled piecemeal. An educated population will want political choice and can effectively use it. There will also be a reasonable amount of choice in the controlled labor market, subject to the confining limits of one’s occupation, and in the controlled product market.

The great new freedom may come in the leisure-time life of individuals. Higher standards of living, more free time, and more education make this not only possible but almost inevitable. Leisure will be the happy hunting ground for the independent spirit. Along with the bureaucratic conservatism of economic and political life may well go a New Bohemianism in the other aspects of life—partly as a reaction to the confining nature of the productive side of society. There may well come a new search for individuality and a new meaning to liberty. The economic system may be highly ordered and the political system barren ideologically;

but the social and recreational and cultural aspects of life should be quite diverse and quite changing.

The world will be for the first time a totally literate world. It will be an organization society, but it need not be peopled by “organization men” whose total lives are ruled by their occupational roles.

The areas closest to technology will be the most conformist; those farthest from the requirements of its service, the most free. The rule of technology need not, as Marx thought it would, reach into every corner of society. In fact, there may come a new emphasis on diversity, on the preservation of national and group traits that runs quite counter to the predictions of uniform mass consumption. The new slavery to technology may bring a new dedication to diversity and individuality. This is the two-sided face of pluralistic industrialism that makes it forever a split personality looking in two directions at the same time. The new slavery and the new freedom go hand in hand.

Utopia never arrives, but men may well settle for the benefits of a greater scope for freedom in their personal lives at the cost of considerable conformity in their working lives. If pluralistic industrialism can be said to have a split personality, then the individual in this society will lead a split life too; he will be a pluralistic individual with more than one pattern of behavior and one dominant allegiance.

Social systems will be reasonably uniform around the world as compared with today’s situation; but there may be substantial diversity within geographical and cultural areas as men and groups seek to establish and maintain their identity. The differences will be between and among individuals and groups and subcultures rather than primarily between and among the major geographical areas of the world. Society at large may become more like the great metropolitan complexes of Paris or London or New York or Tokyo, urbanized and committed to the in-

dustrial way of life, but marked by infinite variety in its details.

Pluralistic industrialism will never reach a final equilibrium. The contest between the forces for uniformity and for diversity will give it life and movement and change. This is a contest which will never reach an ultimate solution. Manager and managed also will struggle all up and down the line of hierarchies all around the world; quiet but often desperate little battles will be fought all over the social landscape.

The uniformity that draws on technology, and the diversity that draws on individuality; the authority that stems from the managers, and the rebellion, however muted, that stems from the managed—these are destined to be the everlasting threads of the future. They will continue in force when class war, and the contest over private versus public initiative, and the battle between the monistic and atomistic ideologies all have been left far behind in the sedimentary layers of history.

NOTES

1. The page citations in the *Manifesto* are to K. Marx and F. Engels, *Manifesto of the Communist Party* (Moscow: Foreign Languages Publishing House, 1955); page citations in *Capital* are to Karl Marx, *Capital*, First edition (Moscow: Foreign Languages Publishing House, 1954), English edition. Chapter XV in this first volume of *Capital* is entitled, "Machinery and Modern Industry," pp. 371–507.

2. W. Arthur Lewis, *The Theory of Economic Growth* (London: George Allen and Unwin Ltd., 1955), p. 116.

3. The 32 labor grades in the basic steel industry and the many thousands of jobs described and rated in the manual in use in the United States are eloquent testimony to the way in which an industrial work force is structured. While the details of the ordering vary among countries, the steel industry of all countries reflects a highly differentiated and ordered work force. See Jack Stieber, *The Steel Industry Wage Structure* (Cambridge, Massachusetts: Harvard University Press, 1959). Compare American Iron and Steel Institute, *Steel in the Soviet Union* (New York: 1959), pp. 287–376.

4. Bert F. Hoselitz, "The City, The Factory, and Economic Growth," *American Economic Review* (May 1955), pp. 166–184.

5. If industrializing countries are arrayed in groups according to product per capita, the proportion of the labor force in agriculture and related industries varies from 61.2 per cent in the least developed group to 14.4 per cent in the group with the highest product per capita. See Simon Kuznets, *Six Lectures on Economic Growth* (Glencoe, Illinois: The Free Press, 1959), pp. 44–45.

6. W. Arthur Lewis, *The Theory of Economic Growth*, p. 92.

7. Marion J. Levy, Jr., "Some Social Obstacles to 'Capital Formation' in 'Underdeveloped Areas,'" in *Capital Formation and Economic Growth*, A Conference of the Universities—National Bureau Committee for Economic Research (New Jersey: Princeton University Press, 1955), p. 461.

8. "Political power, properly so called, is merely the organized power of one class for oppressing another." K. Marx and F. Engels, *Manifesto of the Communist Party* (Moscow: Foreign Languages Publishing House, 1955), p. 95. The highest purpose of the state is the protection of private property; it is an instrument of class domination. See also F. Engels, *Origin of the Family, Private Property and the State*, translated by Ernest Untermann (Chicago: C. H. Kerr & Co., 1902), p. 130, and Paul H. Sweezy, *The Theory of Capitalist Development, Principles of Marxian Political Economy* (New York: Oxford University Press, 1942), pp. 243–244.

9. In the good society which Marx believed to be the final and inevitable result of the dialectical process, there would no longer be a division of society into economic classes. Since he held the state to be merely an instrument of class coercion, with the disappearance of classes, there would follow a concomitant "withering away" of the state. "The society that is to reorganize production on the basis of free and equal association of the producers, will transfer the machinery of state where it will then belong—into the Museum of Antiquities by the side of the spinning wheel and the bronze age." F. Engels, *Origin of the Family, Private Property and the State*, p. 211.

10. Daniel Bell, *Work and Its Discontents* (Boston: Beacon Press, 1956). "Although religion declined, the significance of work was that it could still mobilize emotional energies into creative challenges" (p. 56).

11. Eric Hoffer, "Readiness to Work" (unpublished manuscript).

Theories of Post-Industrialism, Post-Socialism, and Post-Modernity

■ POST-INDUSTRIALISM AND THE NEW CLASS ■

101 ■ Daniel Bell

The Coming of Post-Industrial Society

The concept of the post-industrial society deals primarily with changes *in the social structure*, the way in which the economy is being transformed and the occupational system reworked, and with the new relations between theory and empiricism, particularly science and technology. These changes can be charted, as I seek to do in this [chapter]. But I do not claim that these changes in social structure *determine* corresponding changes in the polity or the culture. Rather, the changes in social structure pose *questions* for the rest of society in three ways. First, the social structure—especially the social structure—is a structure of roles, designed to coordinate the actions of individuals to achieve specific ends. Roles segment individuals by defining limited modes of behavior appropriate to a particular position, but individuals do not always willingly accept the requirements of a role. One aspect of the post-industrial society, for example, is the increasing bureaucratization of science and the increasing specialization of intellectual work into minute parts. Yet it is not clear that individuals entering science will accept this segmentation, as

did the individuals who entered the factory system a hundred and fifty years ago.

Second, changes in social structure pose “management” problems for the political system. In a society which becomes increasingly conscious of its fate, and seeks to control its own fortunes, the political order necessarily becomes paramount. Since the post-industrial society increases the importance of the technical component of knowledge, it forces the hierophants of the new society—the scientists, engineers, and technocrats—either to compete with politicians or become their allies. The relationship between the social structure and the political order thus becomes one of the chief problems of power in a post-industrial society. And, third, the new modes of life, which depend strongly on the primacy of cognitive and theoretical knowledge, inevitably challenge the tendencies of the culture, which strives for the enhancement of the self and turns increasingly anti-nomian and anti-institutional.

In this [chapter], I am concerned chiefly with the social structural and political consequences of the post-industrial society. In a

later work I shall deal with its relation to culture. But the heart of the endeavor is to trace the societal changes primarily within the social structure.

“Too large a generalization,” Alfred North Whitehead wrote, “leads to mere barrenness. It is the large generalization, limited by a happy particularity, which is the fruitful conception.”¹ It is easy—and particularly so today—to set forth an extravagant theory which, in its historical sweep, makes a striking claim to originality. But when tested eventually by reality, it turns into a caricature—viz. James Burnham’s theory of the managerial revolution thirty years ago, or C. Wright Mills’s conception of the power elite, or W. W. Rostow’s stages of economic growth. I have tried to resist that impulse. Instead, I am dealing here with *tendencies*, and have sought to explore the meaning and consequences of those tendencies if the changes in social structure that I describe were to work themselves to their logical limits. But there is no guarantee that they will. Social tensions and social conflicts may modify a society considerably; wars and recriminations can destroy it; the tendencies may provoke a set of reactions that inhibit change. Thus I am writing what Hans Vahinger called an “as if,” a fiction, a logical construction of what *could* be, against which the future social reality can be compared in order to see what intervened to change society in the direction it did take.

The concept of the post-industrial society is a large generalization. Its meaning can be more easily understood if one specifies [eleven] dimensions, or components, of the term:

1. *The centrality of theoretical knowledge.* Every society has always existed on the basis of knowledge, but only now has there been a change whereby the codification of theoretical knowledge and materials science becomes the basis of innovations in technology. One sees this primarily in the new science-based industries—computers, electronics, optics, polymers—that mark the last third of the century.

2. *The creation of a new intellectual technology.* Through new mathematical and economic techniques—based on the computer linear programming, Markov chains, stochastic processes and the like—we can utilize modeling, simulation and other tools of system analysis and decision theory in order to chart more efficient, “rational” solutions to economic and engineering, if not social, problems.

3. *The spread of a knowledge class.* The fastest growing group in society is the technical and professional class. In the United States this group, together with managers, made up 25 percent of a labor force of eight million persons in 1975. By the year 2000, the technical and professional class will be the largest single group in the society.

4. *The change from goods to services.* In the United States today more than 65 out of every 100 persons are engaged in services. By 1980, the figure will be about 70 in every 100. A large service sector exists in every society. In a pre-industrial society this is mainly a household and domestic class. (In England, it was the single largest class in the society until about 1870.) In an industrial society, the services are transportation, utilities, and finance, which are auxiliary to the production of goods, and personal service (beauticians, restaurant employees, and so forth). But in a post-industrial society, the new services are primarily human services (principally in health, education and social services) and professional and technical services (e.g., research, evaluation, computers, and systems analysis). The expansion of these services becomes a constraint on economic growth and a source of persistent inflation.

5. *A change in the character of work.* In a pre-industrial world, life is a game against nature in which men wrest their living from the soil, the waters, or the forests, working usually in small groups, subject to the vicissitudes of nature. In an industrial society, work is a game against fabricated nature, in which men become dwarfed by machines as they turn

out goods and things. But in a post-industrial world, work is primarily a “game between persons” (between bureaucrat and client, doctor and patient, teacher and student, or within research groups, office groups, service groups). Thus in the experience of work and the daily routine, nature is excluded, artifacts are excluded, and persons have to learn how to live with one another. In the history of human society, this is a completely new and unparalleled state of affairs.

6. *The role of women.* Work in the industrial sector (e.g., the factory) has largely been men’s work, from which women have been usually excluded. Work in the post-industrial sector (e.g., human services) provides expanded employment opportunities for women. For the first time, one can say that women have a secure base for economic independence. One sees this in the steadily rising curve of women’s participation in the labor force, in the number of families (now 60 percent of the total) that have more than one regular wage earner, and in the rising incidence of divorce as women increasingly feel less dependent, economically, on men.

7. *Science as the imago.* The scientific community, going back to the seventeenth century, has been a unique institution in human society. It has been charismatic, in that it has been revolutionary in its quest for truth and open in its methods and procedures; it derives its legitimacy from the credo that knowledge itself, not any specific instrumental ends, is the goal of science. Unlike other charismatic communities (principally religious groups and messianic political movements), it has not “routinized” its creeds and enforced official dogmas. Yet until recently, science did not have to deal with the bureaucratization of research, the subordination of its inquiries to state-directed goals, and the “test” of its results on the basis of some instrumental payoff. Now science has become inextricably intertwined not only with technology but with the military and with social technologies and societal needs. In all this—a central feature of the

post-industrial society—the character of the new scientific institutions—will be crucial for the future of free inquiry and knowledge.

8. *Situses as political units.* Most of sociological analysis has focused its attention on classes or strata, horizontal units of society that exist in superior-subordinate relation to each other. Yet for the post-industrial sectors, it may well be that *situses* (from the Latin *situ*, location), a set of vertical orders, will be the more important loci of political attachment. On page 974, I sketch the possible *situses* of the post-industrial order. There are four *functional* *situses*—scientific, technological (i.e., applied skills: engineering, economics, medicine), administrative and cultural—and five *institutional* *situses*—economic enterprises, government bureaus, universities and research complexes, social complexes (e.g., hospitals, social-service centers), and the military. My argument is that the major interest conflicts will be between the *situs* groups, and that the attachments to these *situses* might be sufficiently strong to prevent the organization of the new professional groups into a coherent class in society.²

9. *Meritocracy.* A post-industrial society, being primarily a technical society, awards place less on the basis of inheritance or property (though these can command wealth or cultural advantage) than on education and skill. Inevitably the question of a meritocracy becomes a crucial normative question. In this [chapter] I attempt to define the character of meritocracy and defend the idea of a “just meritocracy,” or of place based on achievement, through the respect of peers.

10. *The end of scarcity?* Most socialist and utopian theories of the nineteenth century ascribed almost all the ills of society to the scarcity of goods and the competition of men for these scarce goods. In fact, one of the most common definitions of economics characterized it as the art of efficient allocation of scarce goods among competing ends. Marx and other socialists argued that abundance was the precondition for socialism and claimed, in

TABLE 1
Stratification and Power

| | PRE-INDUSTRIAL | INDUSTRIAL | POST-INDUSTRIAL |
|------------------|--------------------------------|-----------------------------------|---|
| Resource | Land | Machinery | Knowledge |
| Social locus | Farm, Plantation | Business firm | University, Research institute |
| Dominant figures | Lords, Military | Business men | Scientists, Research men |
| Means of power | Direct control | | Labour of technical-political forces, franchises and rights |
| Class base | Property, Arbitrary | Technical skill | Technical skill, Political organization |
| Access | Inheritance, Seizure by armies | Inheritance, Patronage, Education | Education, Mobilization, Co-optation |

**P/u from
Social Strat. 2e,
p. 808**

fact, that under socialism there would be no need to adopt normative rules of just distribution, since there would be enough for everyone's needs. In that sense, the definition of communism was the abolition of economics, or the "material embodiment" of philosophy. Yet it is quite clear that scarcity will always be with us. I mean not just the question of scarce resources (for this is still a moot point) but that a post-industrial society, by its nature, brings new scarcities which nineteenth- and early-twentieth-century writers had never thought of. The socialists and liberals had talked of the scarcities of goods; but in the post-industrial society there will be scarcities of information and of time. And the problems of allocation inevitably remain, in the crueler form, even, of man becoming *homo economicus* in the disposition of his leisure time.

11. *The economics of information.* Information is by its nature a collective, not a private, good (i.e., a property). In the marketing of individual goods, it is clear that a "competitive" strategy between producers is to be preferred lest enterprise become slothful or monopolistic. Yet for the optimal social investment in knowledge, we have to follow a "cooperative" strategy in order to increase the spread and use of knowledge in society. This new problem regarding information poses the most fascinating challenges to economists

and decision makers in respect to both theory and policy in the post-industrial society. . . .

Who Holds Power?

Decisions are a matter of power, and the crucial questions in any society are: *Who* holds power? And *how* is power held? How power is held is a *system* concept; who holds power is a *group* concept. How one comes to power defines the base and route; who identifies the persons. Clearly, when there is a change in the nature of the system, new groups come to power. (In the tableau of pre-industrial, industrial, and post-industrial societies, the major differences can be shown schematically—see Table 1 on Stratification and Power.)

In the post-industrial society, technical skill becomes the base of and education the mode of access to power; those (or the elite of the group) who come to the fore in this fashion are the scientists. But this does not mean that the scientists are monolithic and act as a corporate group. In actual political situations scientists may divide ideologically (as they have in the recent ABM debate), and different groups of scientists will align themselves with different segments of other elites. In the nature of politics, few groups are monolithic ("the" military, "the" scientists, "the" business class), and any group contending for power

will seek allies from different groups. (Thus, in the Soviet Union, for example, where the interest groups are more clear-cut in functional terms—factory managers, central planners, army officers, party officials—and the power struggle more naked, any faction in the Politburo *seeking* power will make alliances *across* group lines. Yet once *in* power, the victors will have to make decisions *between* groups and affect the relative distribution of power of the functional units and shift the weights of the *system*.) In the change of the system in the post-industrial society, two propositions become evident:

1. As a *stratum*, scientists, or more widely the technical intelligentsia, now have to be taken into account in the political process, though they may not have been before.
2. Science itself is ruled by an ethos which is different from the ethos of other major social groups (e.g., business, the military), and this ethos will *predispose* scientists to act in a different fashion, politically, from other groups.

Forty-five years ago Thorstein Veblen, in his *Engineers and the Price System*, foresaw a new society based on technical organization and industrial management, a “soviet of technicians,” as he put it in the striking language he loved to use in order to scare and mystify the academic world. In making this prediction, Veblen shared the illusion of that earlier technocrat, Henri de Saint-Simon, that the complexity of the industrial system and the indispensability of the technician made military and political revolutions a thing of the past. “Revolutions in the eighteenth century,” Veblen wrote, “were military and political; and the Elder Statesmen who now believe themselves to be making history still believe that revolutions can be made and unmade by the same ways and means in the twentieth century. But any substantial or effectual overturn in the twentieth century will necessarily be an industrial overturn, and by

the same token, any twentieth-century revolution can be combatted or neutralized only by industrial ways and means.”

If a revolution were to come about in the United States—as a practiced skeptic Veblen was highly dubious of that prospect—it would not be led by a minority political party, as in Soviet Russia, which was a loose-knit and backward industrial region, nor would it come from the trade-union “votaries of the dinner pail,” who, as a vested interest themselves, simply sought to keep prices up and labor supply down. It would occur, he said, along the lines “already laid down by the material conditions of its productive industry.” And, turning this Marxist prism to his own perceptions, Veblen continued: “These main lines of revolutionary strategy are lines of technical organization and industrial management; essentially lines of industrial engineering; such as will fit the organization to take care of the highly technical industrial system that constitutes the indispensable material foundation of any modern civilized community.”

The heart of Veblen’s assessment of the revolutionary class is thus summed up in his identification of the “production engineers” as the indispensable “General Staff of the industrial system.” “Without their immediate and unremitting guidance and correction the industrial system will not work. It is a mechanically organized structure of the technical processes designed, installed, and conducted by the production engineers. Without them and their constant attention to the industrial equipment, the mechanical appliances of industry will foot up to just so much junk.”

This syndicalist idea that revolution in the twentieth century could only be an “industrial overturn” exemplifies the fallacy in so much of Veblen’s thought. For as we have learned, no matter how technical social processes may be, the crucial turning points in a society occur in a political form. It is not the technocrat who ultimately holds power, but the politician.

TABLE 2
Evolved Model

| | | | |
|-----------------|--|------------|------------|
| Mode of Power: | P/u from Social Strat. 2e, p. 810 | Membership | Skill |
| Mode of Access: | | | Education |
| Social Unit: | | Party | Individual |

The major changes that have reshaped American society over the past thirty years—the creation of a managed economy, a welfare society, and a mobilized polity—grew out of political responses: in the first instances to accommodate the demands of economically insecure and disadvantaged groups—the farmers, workers, blacks and the poor—for protection from the hazards of the market; and later because of the concentration of resources and political objectives following the mobilized postures of the cold war and the space race.

All of this opens up a broader and more theoretical perspective about the changing nature of class and social position in contemporary society. *Class, in the final sense, denotes not a specific group of persons but a system that has institutionalized the ground rules for acquiring, holding, and transferring differential power and its attendant privileges.* In Western society, the dominant system has been property, guaranteed and safeguarded by the legal order, and transmitted through a system of marriage and family. But over the past twenty-five to fifty years, the property system has been breaking up. In American society today, there are three modes of power and social mobility, and this baffles students of society who seek to tease out the contradictory sources of class positions. There is the historic mode of property as the basis of wealth and power, with inheritance as the major route of access. There is technical skill as the basis of power and position, with education as the necessary route of access to skill. And finally there is political office as a base of power, with organization of a machine as the route of access.

One can, in a simplified way, present these modes in Table 2.

The difficulty in the analysis of power in modern Western societies is that these three systems [in Table 2] co-exist, overlap, and interpenetrate. While the family loses its importance as an economic unit, particularly with the decline of family-firms and the break-up of family capitalism, family background is still advantageous in providing impetus (financial, cultural and personal connections) for the family member. Ethnic groups, often blocked in the economic access to position, have resorted to the political route to gain privilege and wealth. And, increasingly, in the post-industrial society, technical skill becomes an overriding condition of competence for place and position. A son may succeed a father as head of a firm, but without the managerial skill to run the enterprise, the firm may lose out in competition with other, professionally managed corporations. To some extent, the owner of a firm and the politician may hire technicians and experts; yet, unless the owner or politician themselves know enough about the technical issues, their judgments may falter.

The rise of the new elites based on skill derives from the simple fact that knowledge and planning—military planning, economic planning, social planning—have become the basic requisites for all organized action in a modern society. The members of this new technocratic elite, with their new techniques of decision-making (systems analysis, linear programming, and program budgeting), have now become essential to the formulation and

analysis of decisions on which political judgments have to be made, if not to the wielding of power. It is in this broad sense that the spread of education, research, and administration has created a new constituency—the technical and professional intelligentsia.

While these technologists are not bound by a sufficient common interest to make them a political class, they do have common characteristics. They are, first, the products of a new system in the recruitment for power (just as property and inheritance were the essence of the old system). The norms of the new intelligentsia—the norms of professionalism—are a departure from the hitherto prevailing norms of economic self-interest which have guided a business civilization. In the upper reaches of this new elite—that is, in the scientific community—men hold significantly different values, which could become the foundation of the new ethos for such a class.

Actually, the institution of property itself is undergoing a fundamental revision, in a significant way. In Western society for the past several hundred years, property, as the protection of private rights to wealth, has been the economic basis of individualism. Traditionally the institution of property, as Charles Reich of the Yale Law School has put it, “guards the troubled boundary between individual man and the state.” In modern life property has changed in two distinctive ways. One of these is elementary: Individual property has become corporate, and property is no longer controlled by owners but by managers. In a more subtle and diffuse way, however, a new kind of property has emerged, and with it a different kind of legal relationship. To put it more baldly, property today consists not only of visible things (land, possessions, titles) but also of claims, grants, and contracts. The property relationship is not only between persons but between the individual and the government. As Reich points out, “The valuables dispensed by government take many forms, but they all share one characteristic. They are steadily taking the place of the traditional

forms of wealth—forms which are held as private property. Social insurance substitutes for savings, a government contract replaces a businessman’s customers and goodwill. . . . Increasingly, Americans live on government largess—allocated by government on its own terms, and held by recipients subject to conditions which express ‘the public interest.’”³

While many forms of this “new property” represent direct grants (subsidies to farmers, corporations, and universities) or are contracts for services or goods (to industry and universities), the most pervasive form is claims held by individuals (social security, medical care, housing allowances) which derive from a new definition of social rights: claims on the community to ensure equality of treatment, claims legitimately due a person so that he will be able to share in the social heritage. And the most important claim of all is full access to education, within the limits of one’s talent and potential.

The result of all this is to enlarge the arena of power, and at the same time to complicate the modes of decision-making. The domestic political process initiated by the New Deal was in effect a broadening of the “brokerage” system—the system of political deals between constituencies—although there are now many participants in the game. But there is also a new dimension in the political process, which has given the technocrats a new role. Matters of foreign policy have not been a reflex of internal political forces, but a judgment about the national interest, involving strategy decisions based on the calculation of an opponent’s strength and intentions. Once the fundamental policy decision was made to oppose the communist power, many technical decisions, based on military technology and strategic assessments, took on the highest importance in the shaping of subsequent policy. Even a reworking of the economic map of the United States followed as well, with Texas and California gaining great importance because of the electronics and aerospace industries. In these instances technology and strategy laid

down the requirements, and only then could business and local political groups seek to modify, or take advantage of, these decisions so as to protect their own economic interests.

In all this, the technical intelligentsia holds a double position. To the extent that it has interests in research, and positions in the universities, it becomes a new constituency—just as the military is a distinct new constituency, since this country has never before had a permanent military establishment seeking money and support for science, for research and development. Thus the intelligentsia becomes a claimant, like other groups, for public support (though its influence is felt in the bureaucratic and administrative labyrinth, rather than in the electoral system or mass pressure). At the same time, the technicians represent an indispensable administrative staff for the political office holder with his public following. . . .

If one turns, then, to the societal structure of the post-industrial society considered along these two historical axes [of class and power], two conclusions are evident. First, the major class of the emerging new society is primarily a professional class, based on knowledge rather than property. But second, the control system of the society is lodged not in a successor-occupational class but in the political order, and the question of who manages the political order is an open one. (See, on p. 974, “Schema: The Societal Structure of the Post-Industrial Society.”)

In terms of status (esteem and recognition, and possibly income), the knowledge class may be the highest class in the new society, but in the nature of that structure there is no intrinsic reason for this class, on the basis of some coherent or corporate identity, to become a new economic interest class, or a new political class which would bid for power. The reasons for this are evident from an inspection of the Schema.

The professional class as I define it is made up of four estates: the scientific, the technological, the administrative, and the cultural.⁴ While the estates, as a whole, are bound by a

common ethos, there is no intrinsic interest that binds one to the other, except for a common defense of the idea of learning; in fact there are large disjunctions between them. The scientific estate is concerned with the pursuit of basic knowledge and seeks, legitimately, to defend the conditions of such pursuit, untrammelled by political or extraneous influence. The technologists, whether engineers, economists, or physicians, base their work on a codified body of knowledge, but in the application of that knowledge to social or economic purposes they are constrained by the policies of the different *situses* they are obedient to. The administrative estate is concerned with the management of organizations and is bound by the self-interest of the organization itself (its perpetuation and aggrandizement) as well as the implementation of social purposes, and may come into conflict with one or another of the estates. The cultural estate—artistic and religious—is involved with the expressive symbolism (plastic or ideational) of forms and meanings, but to the extent that it is more intensively concerned with meanings, it may find itself increasingly hostile to the technological and administrative estates. As I noted in the introduction, the axial principle of modern culture, in its concern with the self, is antinomian and anti-institutional, and thus hostile to the functional rationality which tends to dominate the application of knowledge by the technological and administrative estates. Thus in the post-industrial society one finds increasingly a disjunction between social structure and culture which inevitably affects the cohesiveness if not the corporate consciousness of the four estates.⁵

While the classes may be represented, horizontally, by *statuses* (headed by the four estates), the society is organized, vertically, by *situses*, which are the actual loci of occupational activities and interests. I use this unfamiliar sociological word *situses* to emphasize the fact that in day-to-day activities the actual play and conflict of interests exist between the organizations to which men belong, rather than between the

Schemata: The Societal Structure of the Post-Industrial Society (U.S. Model)

-
- I. *Statuses: Acts of Stratification—Based on Knowledge*
(Horizontal Structures)
 - A. The professional class: the four estates
 - 1. Scientific
 - 2. Technological (applied skills: engineering, economics, medicine)
 - 3. Administrative
 - 4. Cultural (artistic and religious)
 - B. Technicians and semi-professional
 - C. Clerical and sales
 - D. Craftsmen and semi-skilled (blue-collar)
 - II. *Situses: Locations of Occupational Activities*
(Vertical Structures)
 - A. Economic enterprises and business firms
 - B. Government
 - C. Universities
 - D. Social service
 - E. The military
 - III. *Control: Political and Social Structures*
 - A. The directorate
 - 1. Office of the President
 - 2. Legislative leaders
 - 3. Bureaucratic chiefs
 - 4. Military chiefs
 - B. The politics: constituencies and champions
 - 1. Parties
 - 2. Elites (scientific, academic, business, military)
 - 3. Mobilized groups
 - a) Functional groups (business, professional, labor)
 - b) Ethnic groups
 - c) Special-focus groups
 - (1) Functional majority of elites, poor, etc.
 - (2) Expressive (youth, women, homosexual, etc.)

**P/u from
Social Strat. 2e,
p. 812**

more diffuse class or status identities. In a capitalist society, the property owner or businessman, as a class, is located exclusively in the business firm or corporation, so that status and situs are joined. In the post-industrial society, however, the four estates are distributed among many different situses. Scientists can work for economic enterprises, government, universities, social complexes, or the military (though the bulk of the “pure” scientists are to be found in the university). And the same distributions hold for the technologists and the managers. Because of this “cross-cutting,” the likelihood of a pure “estate” consciousness for political purposes tends to diminish.

Finally, if the major historical turn in the last quarter-century has been the subordination of the economic function to societal goals, the political order necessarily becomes the control system of the society. But who runs it, and for whose (or what) ends? In one respect, what the change may mean is that traditional social conflicts have simply shifted from one arena to another, so that what the traditional classes fought out in the economic realm, where men sought comparative advantage in place, privilege and domination, is now transferred to the political realm, and as that arena widens, the special foci and ethnic groups (the poor and the blacks) now seek to

gain through politics the privileges and advantages they could not obtain in the economic order. This is what has been taking place in recent years, and it will continue. The second, and structurally more pervasive, shift is that in the post-industrial society the *situses* rather than the *statuses* would be the major political-interest units in the society. To some extent this is evident in the familiar phenomenon of pressure groups. But in the post-industrial society it is more likely that the *situses* will achieve greater corporate cohesiveness vis-à-vis one another and become the major claimants for public support and the major constituencies in the determination of public policy.⁶ And yet the very forces which have re-emphasized the primacy of the political order in a technical world make it imperative to define some coherent goals for the society as a whole and, in the process, to articulate a public philosophy which is more than the sum of what particular situses or social groups may want. In the efforts to forge some such coherence one may find the seeds of the cohesiveness of the professional class in the post-industrial society.

A new social system, contrary to Marx, does not always arise necessarily within the shell of an old one but sometimes outside of it. The framework of feudal society was made up of noblemen, lords, soldiers, and priests whose wealth was based on land. The bourgeois society that took hold in the thirteenth century was made up of artisans, merchants, and free professionals whose property lay in their skills or their willingness to take risks, and whose mundane values were far removed from the fading theatrics of the chivalric style of life. It arose, however, outside the feudal landed structure, in the free communes, or towns, that were no longer seignorial dependencies. And these self-ruling small communes became the cornerstones of the future European mercantile and industrial society.⁷

So, too, the process today. The roots of post-industrial society lie in the inexorable influence of science on productive methods, particularly

in the transformation of the electrical and chemical industries at the beginning of the twentieth century. But as Robert Heilbroner has observed: “Science, as we know it, began well before capitalism existed and did not experience its full growth until well after capitalism was solidly entrenched.” And science, as a quasi-autonomous force, would extend beyond capitalism. By this token, one can say that the scientific estate—its ethos and its organization—is the monad that contains within itself the imago of the future society. . . .⁸

Meritocracy and Equality

In 1958, the English sociologist Michael Young wrote a fable, *The Rise of the Meritocracy*.⁹ It purports to be a “manuscript,” written in the year 2033, which breaks off inconclusively for reasons the “narrator” failed to comprehend. The theme is the transformation of English society, by the turn of the twenty-first century, owing to the victory of the principle of achievement over that of ascription (i.e., the gaining of place by assignment or inheritance). For centuries, the elite positions in the society had been held by the children of the nobility on the hereditary principle of succession. But in the nature of modern society, “the rate of social progress depend[ed] on the degree to which power is matched with intelligence.” Britain could no longer afford a ruling class without the necessary technical skills. Through the successive school-reform acts, the principle of merit slowly became established. Each man had his place in the society on the basis of “IQ and Effort.” By 1990 or thereabouts, all adults with IQs over 125 belonged to the meritocracy.

But with that transformation came an unexpected reaction. Previously, talent had been distributed throughout the society, and each class or social group had its own natural leaders. Now all men of talent were raised into a common elite, and those below had no excuses for their failures; they bore the stigma of rejection, they were known inferiors.

By the year 2034 the Populists had revolted. Though the majority of the rebels were members of the lower classes, the leaders were high-status women, often the wives of leading scientists. Relegated during the early married years to the household because of the need to nurture high-IQ children, the activist women had demanded equality between the sexes, a movement that was then generalized into the demand for equality for all, and for a classless society. Life was not to be ruled by “a mathematical measure” but each person would develop his own diverse capacities for leading his own life.¹⁰ The Populists won. After little more than half a century, the Meritocracy had come to an end.

Is this, too, the fate of the post-industrial society? The post-industrial society, in its initial logic, is a meritocracy. Differential status and differential income are based on technical skills and higher education. Without those achievements one cannot fulfill the requirements of the new social division of labor which is a feature of that society. And there are few high places open without those skills. To that extent, the post-industrial society differs from society at the turn of the twentieth century. The initial change, of course, came in the professions. Seventy years or so ago, one could still “read” law in a lawyer’s office and take the bar examination without a college degree. Today, in medicine, law, accounting, and a dozen other professions, one needs a college degree and accrediting, through examination, by legally sanctioned committees of the profession, before one can practice one’s art. For many years, until after World War II, business was the chief route open to an ambitious and aggressive person who wanted to strike out for himself. And the rags-to-riches ascent (or, more accurately, clerk-to-capitalist, if one follows the career of a Rockefeller, Harriman, or Carnegie) required drive and ruthlessness rather than education and skills. One can still start various kinds of small businesses (usually, now, by franchise from a larger corporation), but the

expansion of such enterprises takes vastly different skills than in the past. Within the corporation, as managerial positions have become professionalized, individuals are rarely promoted from shop jobs below but are chosen from the outside, with a college degree as the passport of recognition. Only in politics, where position may be achieved through the ability to recruit a following, or through patronage, is there a relatively open ladder without formal credentials.

Technical skill, in the post-industrial society, is what the economists call “human capital.” An “investment” in four years of college, according to initial estimates of Gary Becker, yields, over the average working life of the male graduate, an annual return of about 13 percent.¹¹ Graduation from an elite college (or elite law school or business school) gives one a further differential advantage over graduates from “mass” or state schools. Thus, the university, which once reflected the status system of the society, has now become the arbiter of class position. As the gate-keeper, it has gained a quasi-monopoly in determining the future stratification of the society.¹²

Any institution which gains a quasi-monopoly power over the fate of individuals is likely, in a free society, to be subject to quick attack. Thus, it is striking that the populist revolt, which Michael Young foresaw several decades hence, has already begun, at the very onset of the post-industrial society. One sees this in the derogation of the IQ and the denunciation of theories espousing a genetic basis of intelligence; the demand for “open admission” to universities on the part of minority groups in the large urban centers; the pressure for increased numbers of blacks, women, and specific minority groups such as Puerto Ricans and Chicanos in the faculties of universities, by quotas if necessary; and the attack on “credentials” and even schooling itself as the determinant of a man’s position in the society. A post-industrial society reshapes the class structure of society by creating new technical elites. The populist reaction, which

has begun in the 1970s, raises the demand for greater “equality” as a defense against being excluded from that society. Thus the issue of meritocracy versus equality. . . .

The claim for group rights stands in formal contradiction to the principle of individualism, with its emphasis on achievement and universalism. But in reality it is no more than the extension, to hitherto excluded social units, of the group principle which has undergirded American politics from the start. The group process—which was the vaunted discovery of the “realists” of American political science—consisted largely of economic bargaining between functional or pressure groups operating outside the formal structure of the party system. What we now find are ethnic and ascriptive groups claiming formal representation both in the formal political structure and in all other institutions of the society as well.

These claims are legitimated, further, by the fact that America has been a pluralist society, or has come to accept a new definition of pluralism rather than the homogeneity of Americanism. Pluralism, in its classic conceptions,¹³ made a claim for the continuing cultural identity of ethnic and religious groups and for the institutional autonomy of cultural institutions (e.g., universities) from politics. Pluralism was based on the separation of realms. But what we have today is a thoroughgoing politicizing of society in which not only the market is subordinated to political decision but all institutions have to bend to the demands of a political center and politicize themselves in group representational terms. Here, too, there has been another change. In functional group politics, membership was not fixed, and one could find cross-cutting allegiances or shifting coalitions. Today the groups that claim representation—in the political parties, in the universities, in the hospitals and the community—are formed by primordial or biological ties, and one cannot erase the ascriptive nature of sex or color.

And yet, once one accepts the principle of redress and representation of the disadvantaged in the group terms that were initially formulated, it is difficult for the polity to deny those later claims. That is the logic of democracy which has always been present in the ambiguous legacy of the principle of equality.

NOTES

1. Alfred North Whitehead, *Science and the Modern World* (New York, 1960; original edition, 1925), p. 46.

2. What is striking is that in the communist world, it is quite clear that *situses* play the major role in politics. One analyzes the play of power, not in class terms, but on the basis of the rivalries among the party, the military, the planning ministries, the industrial enterprises, the collective farms, the cultural institutions—all of which are *situses*.

3. Charles Reich, “The New Property,” *The Public Interest*, no. 3 (Spring 1966), p. 57.

4. The suggestion of four estates is derived, of course, from Don K. Price’s fruitful book *The Scientific Estate* (Cambridge, Mass., 1965). Price defines four functions in government—the scientific, professional, administrative, and political—and converts each function, as an ideal type, into an estate. My differences with Price are twofold: I think the estates can be represented more accurately as social groups, rather than functions; more importantly, I do not consider the *political* function coeval logically with the others, for I see the political as the control system of the entire societal structure. Terminologically, I have substituted the word “technological” (for the applied skills) where Price uses “professional,” since I would reserve “professional” for the larger meaning of the entire class, and I have added a cultural estate, where Price has none. Nonetheless, my indebtedness to Price is great.

5. One might note that the more extreme forms of the “new consciousness” such as Theodore Rozsak’s *The Making of a Counter-Culture* and Charles Reich’s *The Greening of America* manifest a distinct hostility not only to scientism, but to science as well.

6. The limitation of this analysis is that while the post-industrial society, in its societal structure, increasingly becomes a *functional* society, the political order is not organized in functional terms. Thus the continuing existence of the traditional geographical districts and the dispersal of persons in this fashion means that the political issues at any one time are

much more diffuse than the interests of the particular statuses or situses. It would also indicate that the situses would, like the pressure groups, operate primarily through the lobbying of the legislative and executive branches, rather than work directly through the electoral process. Reality complicates immeasurably any ideal-type schemas.

7. Paradoxically, the growth of that society came about only after the self-contained economic life of the commune—its roots—was broken by the rise of larger-scale industry which, in branching out, could buy its raw materials in one town and sell in another, and which made its way, against both the older feudal society and the regulative restrictions of the commune, in alliance with the monarchical centralization of the newly emerging national state.

8. This is, indeed, Heilbroner's suggestion. See Robert Heilbroner, *The Limits of American Capitalism* (New York, 1966), p. 115.

9. Michael Young, *The Rise of the Meritocracy, 1870–2033* (London, 1958).

10. A theoretician of the Technicians party, Professor Eagle, had argued that marriage partners, in

the national interest, should consult the intelligence register, for a high-IQ man who mates with a low-IQ woman is wasting his genes. The activist women, on the other hand, took romance as their banner and beauty as their flag, arguing that marriage should be based on attraction. Their favorite slogan was "Beauty is achievable by all."

11. Gary S. Becker, *Human Capital* (New York, 1964), p. 112. Later writers have suggested this figure may be too high; the point remains that a college degree does provide an investment "yield."

12. For a comprehensive discussion of this major social change, see Jencks and Riesman, *The Academic Revolution* (New York, 1968). For a survey of the reaction, see Stephen Graubard and Geno Balloiti, eds., *The Embattled University* (New York, 1970).

13. See, for example, the work of R. M. MacIver, *The More Perfect Union: A Program for the Control of Inter-group Discrimination* (New York, 1948), and on the religious side, John Courtney Murray, *We Hold These Truths: Reflections on the American Proposition* (New York, 1960).

102 ■ Gøsta Esping-Andersen

Social Foundations of Postindustrial Economies

A social trend is basically a projection from the past. Many of the early postindustrial theorists, such as Daniel Bell (1976), predicted a future in which most of what they deemed positive in the era of the 'democratic class struggle' would come to full fruition. Bell's vision was a coming society of professionals and technicians, one where 'situs' rather than class conflict would reign. This was a radical reinterpretation in so far as postindustrial society would do away with class altogether. Also Lipset has now embraced this position (Clark and Lipset, 1991; Clark, Lipset, and Rempel, 1993).

Today's visionaries find much less cause for optimism; their projections are most likely to range from the sombre to the outright gloomy. The sombre view insists that little of substance has changed (Wright, 1989; Erikson and Goldthorpe, 1992; and Hout *et al.*, 1993); that the cleavages of the past remain pretty much intact. The gloom comes from those who see a new era of polarization. American or British observers see a world with a 'declining middle', job polarization, and a new underclass (Harrison and Bluestone, 1988; Jencks and Peterson, 1991; Levy, 1988; Burtless, 1990). Europeans, in con-

trast, see a two-thirds society with social exclusion, marginalization, and outsider classes (Van Parijs, 1987; Offe, 1985; Esping-Anderesen, 1993; Brown and Crompton, 1994).

What the new pessimists see, in brief, is the possibility of a resurgent proletarian underclass and, in its wake, a menacing set of new ‘class correlates’. The transAtlantic difference of accent is clearly related to job performance: in North America labour market exclusion is less dramatic than is growing pay inequality, declining real wages, and a swelling army of the working poor. Europe’s social safety net manages to stem the tide of inequality but its inferior job performance induces mass exclusion.

How do we understand the postindustrial employment problem? Is there indeed any problem? And, if new class configurations are emerging, what is driving them? Besides the radical optimists, such as Lipset *et al.*, who basically deny that there is any problem worthy of discussion, the reigning answers all fall back on the great ‘equality-jobs’ trade-off. According to mainstream analysis, exclusion in Europe and inequality in America are two sides of the same coin, namely the inevitable consequence of technology and the new global economy. . . .

The Dilemmas of Globalization and Technological Change

Global trade and capital mobility are often seen as the culprits of de-industrialization. The Asian tigers can produce steel, scooters, stereos, or shoes much cheaper than Europe or North America. Even if their regimes are not authoritarian, unionism is weak and labour discipline arguably high. Regardless, they hold a massive competitive edge in terms of wage costs in mass-production industries as well as in many services. Less-skilled workers in advanced economies are no longer cushioned by Keynesianism or import protection. They must now compete in a global labour market.

The special vulnerability of the low-skilled is highlighted in Wood’s (1995) study on the impact of imports from the ‘South’ on labour markets in the ‘North’. He paints a fairly dramatic picture of worker displacement and downward wage pressures, in particular for the low-skilled. His findings are, however, widely contested. For one, the lion’s share of European trade is *intra-European*. The share of total European Union trade with the outside world is, according to all estimates, less than 10 per cent of total—of which a large proportion is with North America, the Antipodes, or Eastern Europe. Hence, if there is any ‘Third World’ effect, it must be at the margin. This is more or less what the mainstream view concludes (OECD, 1994).

It is, in the first place, questionable whether globalization really is something new, and if it indeed is altering our capacity to harmonize equality and full employment. Bairoch (1996) shows that the level of internationalization of trade, finance, and capital today is hardly greater than it was 80 years ago. In fact, it was the period of protectionism and mounting tariffs following World War One and the Great Depression which was anomalous. Moreover, the spurt in globalization over the past few decades has hardly altered the trade dependency of those small, open economies (like the Benelux or Nordic countries) that were *always* open.

With these qualifiers in mind, we might ask ourselves what has *de facto* changed to make equality and social protection problematic now, when once a strong welfare state was viewed as the precondition for successful international performance? It was precisely due to their extreme degree of international vulnerability that the small European economies spearheaded strong worker protection and welfare guarantees after World War Two. Global trade competition is not novel to Belgium or Denmark; it is to the large economies, such as the American or British, or the erstwhile protected ones, such as Australia. Perhaps the real problem is that

the late-comers to globalization have chosen the ‘low road-low wage’ strategy of competition, thus forcing the hand of the vanguards?

Regardless, the facts point to technological change as the more potent source of falling demand for less-qualified workers. Employment lost due to trade competition from the ‘South’ (i.e. Asia or the Third World) is rather trivial in comparison with the volume of job loss due to ‘structural change’. In a few countries (France and Japan), jobs gained from exporting to the ‘South’ actually outnumber those lost to import penetration. Only the UK and the USA, two countries traditionally dominated by comparatively low-skilled, mass-production industry, exhibit a substantial loss of domestic employment due to competition from Asia or elsewhere.¹

The really powerful impulse, then, may come from rapid structural and technological change whether or not this was initially propelled by global trade. Whatever the root cause, the employment effect ends up being fairly similar: it raises the returns to education and reduces demand for lower-skilled and less-experienced workers. Hence the unfavourable labour market position of the unskilled in general, but also of youth and women to the extent that they lack experience and practical skills. In either case, therefore, we seem unable to dodge the evil choice between heightened pay and job inequalities, on one hand, or unemployment and exclusion, on the other hand.

Golden Age capitalism could absorb masses of low-skilled workers on simple assembly-line production, churning out mass-production goods for which there was massive demand. It is these kinds of jobs that are rapidly disappearing within the advanced economies and, as we know, virtually all *net* new job growth will have to come from services.

That simple, routine industrial jobs are vanishing must be considered a ‘Paretian’ welfare gain: the ‘South’ benefits from taking over simple mass production manufacturing, thus creating domestic jobs and

wealth and paving the way for its own process of de-ruralization and, eventually, post-industrialization; the ‘North’ gains by eliminating its most unpleasant jobs, but only in so far as these are being substituted by more pleasant and better-paid jobs. The optimistic postindustrial theorists believed this to be the case. Current opinion is sceptical, and the data seem to confirm it. We must therefore examine more closely the workings of the service economy.

Dilemmas of the New Service Economy

The decline of industrial employment began in earnest in the 1980s. Between 1979 and 1993, the OECD countries lost an (unweighted) average of 22 per cent of their manufacturing jobs. Some countries (Belgium, France, Norway, Sweden, Spain, and the UK) have been hit especially hard, with a net loss between one-third and one-half.² Such magnitudes echo post-war de-ruralization. But today’s equivalent to the assembly lines, namely low-end tertiary employment, has greater difficulty absorbing the joint impact of industrial job losses, women’s rising labour supply, and the baby-boom cohorts. Why should this be so? In this chapter, I focus on three dilemmas. The first derives from a rarely recognized, inherent characteristic of services themselves: the more we expand the tertiary labour market, the larger is the share of low-skilled services. The second comes from the well-known ‘Baumol cost-disease’ problem in services (Baumol, 1967). And the third derives from households’ economic choices and, in particular, from women’s choice to work for pay.

Today, almost all net job creation occurs in services but even with bouyant growth, their ability to absorb masses of redundant industrial workers cannot be taken for granted. Many, like business and health services, are skill-intensive, and the more routine and labour-intensive services (such as social care,

waiting, or personal services generally) often require a modicum of social or cultural skills that a redundant steel worker is unlikely to possess. Labour-intensive services, be they in the public or private sector, are very female-dominated; not merely because they are easy-entry jobs, but also because they typically represent a marketized version of conventional domestic tasks. In any case, services and female participation have grown in tandem. To better understand contemporary employment dilemmas, we need to come to grips with the logic of services. First, what precisely is the service economy? Second, what drives it?

Service Occupations

Apart from the pioneering work of Renner (1953) and, more recently, Bell (1976) and Goldthorpe (1982), little attention has been paid to occupational hierarchies in services. Since they were primarily concerned with the new knowledge class, the lower orders have been largely ignored. As in Braverman (1974) and Erikson and Goldthorpe (1992), they are simply lumped together with manual, industrial workers.

Such superimposition is not necessarily warranted. Even low-end service work is in typically less strict command hierarchies, it is less difficult to monitor, and orders will come more from customers than from bosses. Most servicing work is quite individualized. For heuristic reasons, it may therefore be useful to distinguish between the traditional skill and authority hierarchy of 'industrial society' and that of the servicing society.³ In Table 1, I propose one such classification.

The two hierarchies exclude agricultural occupations and, more importantly, also the self-employed and entrepreneurs. Omitting self-employment may actually be problematic because it is becoming one means of gaining access to the labour market when the supply of 'regular' employment contracts is scarce. This classificatory exercise must be seen as an heuristic device.⁴

TABLE 1
Industrial and Postindustrial
Occupational Hierarchies

| <i>The Industrial Hierarchy</i> | <i>The Service Hierarchy</i> |
|---------------------------------|---------------------------------|
| Managers and executives | Professionals |
| Administrators, supervisors | Semi-professionals, technicians |
| Skilled manuals | Skilled service |
| Unskilled manuals | Unskilled service |

Lousy Jobs or Outsiders?

Our present preoccupation with the equality-jobs trade-off may have blinded us to a second dilemma, i.e. professionalization with exclusion or full employment with job polarization. The highly professionalized scenario depicted by postindustrial optimists can be attained only at the price of massive exclusion, while a massive growth of services is likely to produce a mass of 'lousy' jobs. We need therefore to gain a better picture of postindustrial job trends.

The postindustrial pessimists, like Braverman (1974), err on one count: tertiarization undoubtedly implies occupational upgrading. Over the 1980s, professional-technical jobs rose 3–4 times as fast as employment overall in Belgium, Germany, Sweden, and Japan, and an astounding 7 times as fast in France. Their relative growth was less spectacular in Canada and the United States (1.7 and 1.9, respectively), two nations with unusually strong aggregate service expansion.⁵ Indications are that the professional bias is inversely related to service economy growth. This comes out in Table 2.

The statistical relationship is in fact quite strong: for each additional percentage point of service job growth, the professional ratio declines by 1.2 points.⁶

Since we must rely almost exclusively on services to furnish new jobs and thus reduce unemployment, the contours of postindustrial employment pose an unpleasant trade-off: we will come close to the professionalized

TABLE 2
Service Growth and the Professional-Technical Bias
of Job Growth, 1980–1990^a

| | <i>Annual Service Growth</i> | <i>Ratio of Professional to Total Growth</i> |
|-----------|----------------------------------|--|
| Australia | 3.3 | 1.8 |
| Canada | 2.8 | 1.7 |
| UK | 2.1 | 2.8 |
| USA | 2.7 | 1.9 |
| Denmark | 0.8 | 2.0 |
| Sweden | 1.2 | 4.3 |
| Austria | 1.9 | 2.5 |
| Belgium | 1.4 | 3.8 |
| France | 1.6 | 7.3 |
| Germany | 1.8 | 3.1 |
| Japan | 1.2 | 3.8 |

^a Years vary according to data availability.

Source: ILO, *Yearbook of Labour Statistics* (various years).

world of Daniel Bell, but only when services are relatively stagnant. And this entails unemployment and labour market exclusion. Vice versa, we can minimize exclusion but here we must pay the price of accepting a less favourable, more ‘proletarian’ mix of service occupations. This comes out quite clearly when we examine the occupational hierarchies in greater detail.

Table 3 divides the active population, *c.* mid-1980s, into our ‘industrial’ and ‘postindustrial’ hierarchy and, as somewhat of a residual, it provides an estimate of the excluded, ‘outsider’ population (here measured as a percentage of the working-age population). The three-way comparison between Germany, Sweden, and the United States is chosen because these nations are prototypical examples of three types of welfare regimes—conservative, social democratic, and liberal, respectively. But they also represent three distinct employment trajectories: Germany stands as a best-case version of European jobless growth; Sweden is the epitome of the Nordic welfare state-led model of service ex-

pansion; and the United States is, *par excellence*, the leading example of unregulated, market-driven employment.

On some counts there is convergence. Manual workers decline, and the occupations that constituted the major source of post-war ‘middle-class growth’, such as managers, clerical, and sales jobs, are now stagnant. Today’s growth is undisputably dominated by professionals and semi-professionals. This is all more or less consistent with the prognosis of Bell (1976). Accordingly, chances of upward class mobility in postindustrial societies will depend primarily on how much the skilled and professional service occupations grow. On the other hand, if the objective is full employment there will also be substantial growth of less-qualified service jobs, something that Daniel Bell did not envisage.

A postindustrial paradise would combine full employment with a slim but highly skilled industrial labour force in unison with a professionally dominated service economy. No country comes even close to this. In Germany, which often epitomizes skill-upgrading (see, for example, Kern and Schumann, 1984), skilled manual workers have actually disappeared faster than the unskilled—although from the 1980s onwards, this has reversed. None the less, Germany’s services are heavily professional and ‘non-proletarian’. High labour costs crowd out private social services; low levels of female participation make them less demanded. One reason for Germany’s lower female employment rates is the absence of public care services. And, besides costs, one reason for Germany’s stagnant personal services is that families engage in more self-servicing.

As a consequence, German postindustrialization provides no substantial employment outlet for either laid-off manual workers or less-qualified women. Instead, both these groups have been managed primarily through

TABLE 3
The Distribution of Occupations and the Size of the Outsider Population (Percentage Change, 1960–1980s, in Parentheses)

| | <i>Germany</i> 1985 | <i>Sweden</i> 1985 | <i>United States</i> 1988 |
|-------------------------------------|------------------------|-----------------------|------------------------------|
| <i>'Industrial' society</i> | | | |
| Unskilled manual | 16.5 (–0) | 12.4 (–42) | 14.4 (–33) |
| Skilled manual | 17.3 (–32) | 15.2 (–18) | 8.7 (–34) |
| Clerical and sales | 29.6 (+30) | 18.6 (+16) | 28.3 (+21) |
| Managers ^a | 4.5 (+36) | 4.0 (–15) | 9.1 (+17) |
| <i>Total</i> | 67.9 (–0) | 50.2 (–18) | 60.4 (–8) |
| <i>'Servicing' society</i> | | | |
| Unskilled service | 4.5 (–48) | 16.9 (+78) | 11.7 (–0) |
| Skilled service | 5.0 (+194) | 4.4 (–0) | 6.6 (+57) |
| Professional-technical | 17.3 (+121) | 21.9 (+89) | 18.1 (+56) |
| <i>Total</i> | 26.8 (+47) | 43.2 (+70) | 36.4 (+31) |
| <i>'Outsider' society</i> | | | |
| Not employed | 35.2 | 16.7 | 29.5 |
| Long-term unemployment ^b | 46.3 | 5.0 | 5.6 |

^aIncludes also self-employed (non-professional).

^bAs a percentage of all unemployed (1990).

Note: The table excludes primary sector occupations.

Sources: Recomputations from Esping-Andersen (1993: Tables 2.3 and 2.4); and OECD (1992; 1994).

labour reduction strategies: early retirement or unemployment insurance for industrial workers, and discouragement of female careers. Germany may come closest to the postindustrial ideal as regards job structure, but this comes at the expense of employment exclusion.

Sweden does conform to the industrial skill-upgrading thesis but suffers from strong polarization in the service occupations. The phenomenally high levels of female employment go hand-in-hand with huge numbers of (often low-skilled) public-sector jobs. And Sweden's formidable investment in retraining and active manpower programmes has, until the crash in the 1990s, helped recycle redundant industrial workers into alternative jobs. The combination of these two factors implies few labour market outsiders, be it in the form of early retirees, mass unemployment, or discouraged women.

The United States, finally, exhibits skill polarization in both industry and services. It is an economy biased towards unqualified jobs. De-industrialization has reduced manual employment equally between the skilled and unskilled; declining wages have helped maintain their relative share in industry and services. Of course, our data disguise the fact that, in absolute terms, the number of net new jobs has risen phenomenally. Hence, the absorption of immigrant masses, the huge growth in women's employment, and the relocation of redundant industrial workers were all made possible by the sheer volume of job growth.

The juxtaposition between Germany's more favourable skill mix, and the large share of unskilled servicing jobs in both Sweden and the United States highlights the basic dilemma: today, substantial employment expansion requires heavy growth in consumer

TABLE 4

The Growing Productivity Gap. Percentage Change in the Ratio of Service-to-Manufacturing Productivity, 1983–1995, for Select Countries^a

| | <i>Ratio of Restaurants to Manufacture</i> | <i>Ratio of Personal Services to Manufacture</i> | <i>Ratio of Business Services to Manufacture</i> | <i>Ratio of Social Services to Manufacture</i> |
|---------------|--|--|--|--|
| United States | –7 | –47 | –55 | –62 |
| Denmark | –35 | –30 | –11 | –2 |
| Sweden | –82 | –60 | –71 | –54 |
| France | –56 | –41 | –38 | +4 |
| Netherlands | –40 | –19 | –38 | –28 |

^aProductivity is measured as GDP (in constant prices) per person employed.

Source: OECD Data File on Services Statistics on Value Added and Employment.

and social services, both characteristic of a large unskilled quotient the more they grow—hence an apparent trade-off between either joblessness or a mass of inferior jobs.

The real issue lies in the correlates of class. It makes a notable difference whether inferior jobs provide inferior welfare. In Sweden, they are relatively well-paid and secure welfare state jobs; indeed, private sector ‘Mcjobs’ hardly exist. While relative pay in Swedish welfare service jobs has been in modest decline, job security has not. A notable aspect of Sweden’s gargantuan welfare spending cutbacks in the 1990s is that public employment is safeguarded. In the United States, the low-end service workers are mainly in the private sector, typically poorly paid and excluded from occupational welfare entitlements and basic job security. In Sweden, the concentration of women in low-end servicing jobs is extreme (women account for roughly 80 per cent of the total), while in America the bias is tendentially ethnic—Hispanics—rather than gender-based (Esping-Andersen, 1990; ch. 8).

In summary, there are undisputably forces that pull economies away from job homogeneity and wage equality, and towards polarization. It is therefore doubtful whether the great ‘equality-jobs’ trade-off overshadows everything else. It is, as we shall now see, more likely that trade-offs overlap.

The Cost-Disease and Service Expansion

Globalization, new technologies, and tertiarization all seem to produce heightened labour market polarization. The productivity lag within many services only helps reinforce this problem. As originally argued by Fourastiér, and subsequently formalized by Baumol (1967), services face a long-run ‘cost-disease’ problem. This will come about because, in the long haul, productivity grows on average much faster in manufacturing than in (most) services. As shown in Table 4, most services lag—sometimes dramatically—behind manufacturing.

There are three possible responses to the productivity gap. One would be to allow labour costs to adjust to productivity differentials—the market-clearing approach. The downside of unregulated wage adjustment is that this would push earnings in the more stagnant services towards zero. Many services, such as music concerts, psychotherapy, or aged care, are capable of almost no productivity enhancement (at least not without a quality loss), and they would therefore most likely disappear, simply because no one would be willing to perform them at earnings that correspond to relative productivity. The potential limits to the market-clearing strategy is poverty: aggregate service con-

TABLE 5
The Cost-Disease Illustrated by Laundry Servicing. The Comparative Cost of Washing and Ironing One Man's Shirt, Employment, and Self-Service Equipment, mid-1990s

| | Cost (\$US) 1996 | Working-Age Population per Laundry Worker ^a | Washing-Machines ^b |
|----------------|------------------|---|-------------------------------|
| Denmark | 5.20 | 3,500 | 74 |
| Sweden | 4.25 | 727 | 87 |
| France | 4.50 | | |
| Germany (West) | 3.70 | 667 | 88 |
| Italy | 3.25 | n.a. | 96 |
| Spain | 3.90 | 905 | 87 |
| UK | 2.20 | 750 | 87 |
| USA | 1.50 | 391 | 75 |

^aData for the UK are 1993; for the USA, Spain, and Sweden, 1990; for Germany, 1987.

^bPercentage households equipped with a washing-machine, 1991.

Sources: Employment: United States Statistical Abstract, 1995, Table 668; *Sveriges Statistiska Aarsbok*, 1993, Table 204; German Census of 1987, Part 2, Fachserie 1, Haft 10; Government Statistical Office, *Employment Gazette, Historical Supplement*, no. 4, October 1994; Ministerio de Economia y Hacienda, *Prospectiva de las Ocupaciones y la Formacion en la Espana de los Noventa*, 1991, App. Table 2. Washing-machines: United States Statistical Abstract, 1995, Table 1376. Prices: own data collection in Italy, Spain, and the United States, and special thanks to Francis Castles, Kevin Farnsworth, Joakim Palme, Jon Kvist, Karl Ulrich Mayer, John Myles, Kari Salminen, and Lucy Roberts for help on price and employment figures.

sumption may, on the one hand, be stimulated by low prices but if, on the other hand, this means that a large population mass is employed at poverty-level wages, aggregate demand will suffer. Low wages can of course be offset by income supplements, as now Canada, Australia, and the United States do with various forms of negative income tax programmes.

The second and, in most countries, typical response has been to allow service earnings to follow general wage developments in the economy. This may matter little for engineering, product design, or financial consultancy, but it risks pricing out of the market labour-intensive, low value-added activities such as personal services, entertainment, or private daycare. Both the Nordic countries and Continental Europe have in common an institutional framework which has this effect: high wage costs across the board and a more compressed earnings distribution. Hence, personal services grow very sluggishly, if at all,

and many services become almost extinct. To illustrate the problem, let us examine a prototypical labour-intensive service: laundries.

As we see in Table 5, time-saving household goods (washing-machines) are almost universal everywhere. What varies is the cost of out-servicing and, therefore, laundry jobs. Laundries literally clutter American streets because they are cheap. They are almost impossible to find in Copenhagen and Stockholm because they are priced out of the market.

In a service-led economy, this kind of cost-disease translates easily into jobless growth. But there is, of course, a third possible solution, namely to subsidize services—either directly via government production, or indirectly via subsidies to consumers. All advanced nations have to a greater or lesser degree adopted the subsidy strategy, especially for vital collective goods such as health and education, or for culturally valued goods such as opera and theatre. The European nations stress direct public provision, the United

States favours tax-subsidized private provision. The uniqueness of Scandinavia lies in government's huge role in furnishing labour-intensive, and otherwise unaffordable, care services to families. Hence, Scandinavian tertiarization is uniquely biased towards welfare state jobs and away from market services. In the United States, the lion's share of caring services is marketized while, in Continental Europe, it is basically familialized. The limits of subsidization are a question of political economy: of the balance of political power and citizens' willingness to be taxed.

The cost-disease problem and the tension between professionalized or more polarized employment growth are clearly closely related. Where the cost-disease in services is severe, and public servicing is scarce, the size of the 'outsider classes' is likely to be large; where it is countered by low-wage labour markets or by public-service provision, exclusion will be less of a problem than growing occupational polarization. It is this nexus which defines the varying employment scenarios within the three principal welfare regimes.

The Micro-Foundations of Postindustrial Employment

As we leave behind us the era of mass production, there is one thing that has not changed. The household still remains the *sine qua non* of how much, and what kind of, employment will grow. How much families save, consume, or produce affects the probability that their members will find themselves employed or unemployed.

Household market purchase is subordinate to three principal factors: levels of income, relative prices, and time constraints.⁷ As we know from Engel's Law, our propensity to buy non-essentials, like services, will rise with disposable income. The phenomenal rise in real incomes in the post-war decades allowed us initially to satisfy demand for manufactured household goods, like cars and

washing-machines. Subsequently, an increasing share went to servicing less essential needs like entertainment or foreign travel. One obstacle to tertiary sector growth today is that real incomes rise much slower. In the United States, real earnings have been stagnant—even declining for many—since the late 1970s.

But also the distribution of riches matters greatly. A very skewed income distribution may result in a 'Latin American' consumption scenario: a narrow, hyper-serviced élite being waited upon by a mass of impoverished servants. At the other extreme, a very egalitarian income distribution, as in Sweden, may very well impede mass consumption of market services because equal wages mean that service costs will be high.

Rising incomes might also push demand in the direction of 'positional' services, that is, services that are bought for their status value. Following the argument of Hirsh (1976), the 'democratization' of consumption means that traditional status-enhancing goods, like automobiles or cellular phones, eventually lose their prestige value. Hence, as mass consumption spreads so will demand for 'positional' goods (Armani suits or Porsches) and services (haircuts from 'Chez Pierre', rather than the local barbershop).

Secondly, household demand is sensitive to price relativities (the Baumol cost-disease problem). Often the lure of Third-World travel lies in the amazing amount of servicing that our 'tourist dollar' can buy. Price relativities in services are intimately linked to wage differentials and, as we have seen, high wages in low-productivity services might price them out of the market.

Of course, where a service is consumed for its 'positional' value the logic of relative prices is altered, and may even operate in reverse. It is possible that status-seeking clients choose 'Chez Pierre' precisely because he is expensive. The status-hungry may follow outrageous pricing practices as their main signal.

Aside from the more esoteric world of positional goods, the cost-disease problem should be most acute in low-skilled, labour-intensive services that compete head-on with household self-servicing. And herein lies a fundamental misunderstanding in much contemporary economics. It is too often assumed that, since they are largely protected from international competition, services can provide a safe haven of employment—service workers do not compete with, say, Malaysians (except in the case of mass-immigration). True, in most economies the lion's share of services are sheltered from global competition. However, they face an even more ferocious competitor, namely family self-servicing. And, as Gershuny (1978) has stressed, this competition stiffens as households acquire time-saving household machinery (washing-machines, dishwashers, or microwave ovens). Yet, as we saw, Americans are massive consumers of laundries, Europeans wash and iron at home.

The problem is that the marginal cost of services such as laundry, daycare, or home-help services to the elderly affects not only women's ability to pursue careers, but also society's welfare-employment nexus. As women's educational attainment increasingly matches (and surpasses) males', the social cost of housewifery is lost human capital, productive potential, *and* jobs. The high costs of out-servicing in Europe compels families to self-servicing which translates into fewer market services.

Here we come to the third, fundamental condition, namely households' time constraints. The emerging 'postindustrial' household types, be they single person, single parent, or dual-earner households, have one thing in common: a scarcity of time, great difficulties in harmonizing paid employment with family obligations, domestic duties, and leisure.

The trade-off between the burden of domestic unpaid work and paid employment can be quite severe. Since it is unrealistic to expect a major relaxation of this trade-off through male substitution (husbands do not

generally compensate significantly with more unpaid work to offset wives' declining hours), the solution must come from outside the household itself. This means either the market or the state.

The employment potential of new household forms comes out in consumer expenditure data. In a pioneering study, Stigler (1956: 88) suggested that restaurant employment should rise with the rate of women's labour force participation. Today, the dual-earner family norm is much stronger and, in some countries, almost universal. This should show up not only in family consumption behaviour, but ultimately also in service-sector job growth. . . .

Job growth in services is, accordingly, caught between rival forces. There are, on the one hand, the constraints that come from the inherent features of tertiarization or globalization and, on the other hand, the potential catalyst that comes from changing family life. If changing household forms can have positive employment effects, our focus should logically turn to which conditions would help maximize just such an effect. Women with careers are likely to reduce fertility, unless the tension between job and children is eased. If, then, access to child care is one constraint, society may find itself locked into a low-fertility equilibrium as seems to be the case in Southern Europe. And this will affect long-term job and growth prospects because, after all, aggregate demand depends on the size and growth of our population. Hence an ample social service infrastructure is one precondition for service-job growth.

The alternative scenario is that women are discouraged from economic activity. In this case, of course, a beneficial job spiral is blunted. The service economy may flourish with more inequality but also with two-earner households. In either case, we face institutional constraints. Some welfare regimes, and some modes of regulating labour markets and the distribution of welfare constrain

or ease more than others. It is to this that we turn in the next section.

Recasting Welfare Regimes for a Postindustrial Era

If there is one great question that unites all welfare state researchers, it is this: why is it that nations respond so differently to a set of social risks that, all told, is pretty similar whether you are an American, a Spaniard, or a Swede?

Why is it that Scandinavia responds with social democracy and comprehensive welfare states, Continental Europe and East Asia with familialism and corporatist social insurance, and the Anglo-Saxon world with targeted assistance and maximum markets? There is certainly no want of answers. Some point to nation-building traditions and church–state conflicts, some to etatiste initiative, and some to the configuration of class power. These are all compelling explanations but they beg the question of why, in the first place, did state building move in one or another direction? Why was working-class unity and mobilization so thorough in Sweden and not in America? And why, once formed and forged, did the institutions they built last and then reproduce themselves so successfully? The popular answer in latter-day social science is that institutions, once forged, are overpowering and rarely permit the kind of sweeping change that would be observable to the human eye. We may live in a world of path dependencies where social creations, once cemented, are incomparably tougher than the city of Troy. Very well, but who creates and sustains the hegemony of whatever path dependency that obtains?

Economists often draw large benefits from basic simplicity. Their *homo oeconomicus* is certainly not believable, but he can be a convenient tool with which to proceed and eventually make things more complicated. I shall, for the moment, strive for similar minimalism. Sociologists, of course, crave more

complexity, and one *Homo* will therefore not do. Let us allow for three ideal typical *homines*: *Homo liberalismus*, *Homo familius*, and *Homosocialdemocraticus*.⁸

Homo liberalismus resembles Mister Economics because he follows no loftier ideal than his own personal welfare calculus. The well-being of others is their affair, not his. A belief in noble self-reliance does not necessarily imply indifference to others. *Homo liberalismus* may be generous, even altruistic. But kindness towards others is a personal affair, not something dictated from above. His ethics tell him that a free lunch is amoral, that collectivism jeopardizes freedom, that individual liberty is a fragile good, easily sabotaged by sinister socialists or paternalistic authoritarians. *Homo liberalismus* prefers a welfare regime where those who can play the market do so, whereas those who cannot must merit charity.

Homo familius inhabits an altogether different planet. He abhors atomism and impersonality and, hence, markets and individualism. His worst enemy is the Hobbesian world of elbows, because self-interest is amoral; a person will find his equilibrium when he puts himself at the service of his family. Freedom, to *Homo familius*, means that he and his kin are immunized from the ceaseless threats that the greater world around him produce. He is not a go-getter with an irresistible urge to challenge the world around him. He is a satisficer, not maximizer, because what really counts is stability and security; a job for life in the postal service is heaven on earth; it will guarantee him and his kin a good life, security and, incidentally, also the means to land a postal job for his daughters and sons. Both *Homo familius* and *femina familia* see patriarchy as a good thing. The family would be unruly without authority and, in any case, respect and status is due to the father, on whose shoulders so much responsibility weighs. To them, the family is the unrivalled source of solidarity and community because it alone

knows what its members need. Alas, the family is fragile since so many individuals must depend on the *pater familias*. *Homo familium* is therefore quite happy with the idea that the state—or some higher body—eliminates whatever risk of misfortune may arise. Yet, higher bodies are there to service the family, not to command his loyalties. *Homo familium* wants a welfare regime that tames the market and exalts the virtues of close-knit solidarities.

Homo socialdemocraticus is, like a boy scout or good Christian, inclined to believe that he will do better when everybody does better. Doing good to others is not an act of charity and can, indeed, be coolly calculative. *Homo socialdemocraticus* plans his life around the one basic idea that he, and everybody else, will be better off in a world without want but also without free-riders. Society is something that we all are compelled to share, and so we had better share it well. *Homo socialdemocraticus* is, none the less, also a believer in individual, personal empowerment. Collectivism is not pursued for its own sake, but in order to bring out the utmost in each and every individual soul. That is why collective solutions are always the best. But his individualism is cautionary because no one should be granted favours, advantages, special recognition. *Homo socialdemocraticus* must constantly live with a moral clash between individualism and conformity. He loves the idea that we all be equally endowed; he hates the idea that someone may rise above others, particularly above himself. Solidarity is therefore fragile if someone desires to move beyond the common denominator. *Homo socialdemocraticus* is fully convinced that the more we invest in the public good, the better it will become. And this will trickle down to all, himself especially, in the form of a good life. Collective solutions are therefore the single best assurance of a good, if perhaps dull, individual life.

We all combine the instincts of these ideal typical *homines*; we all have moral conflicts.

And all societies combine them in one mix or another. Sweden has its *Homines liberalismi* (although many emigrate to Monte Carlo), and America its *Homines socialdemocratici* (many of which show up regularly in Sweden). But how did a sufficient mass manage to profile itself in collective expression, and thereby sway society towards its preferred welfare regime?

Genetics clearly do not create preferences and beliefs. What might account for this is society itself with all its institutions, incentive systems, and inscribed norms of proper conduct. Society's 'median *Homo*' is created, and this is obviously where labour movements or other collective actors have played a key role. If, originally, a labour movement garnered its strength (or failed to do so) from the available raw material, so it also played a constant historical role in reactivating that very same raw material. Institutional path dependency means that one society is likely to reproduce itself in the image of *Homo socialdemocraticus*, another in the image of *Homo familium*, generation after generation. . . .

If core institutional traits are so unyielding to change, it remains unlikely today—as always—that the contemporary welfare state crisis will produce an avalanche of revolutionary change, no matter how urgent such change is claimed to be. Any blueprint for reform is bound to be naïve if it calls for a radical departure from existing welfare regime practice. The IMF is naïve when it asks that European welfare states adopt, lock, stock, and barrel, the Chilean pension system; American leftists, or emergent democratic governments in ex-communist countries, are equally naïve when they call for the importation of the 'Swedish model'. One thing is to redesign or reform the ways in which the *state* delivers welfare, another to recast the entire welfare regime. Academically speaking, there may exist a blueprint for a 'win-win' strategy, for an ideal postindustrial welfare regime. But unless it is

compatible with existing welfare regime practice, it may not be practicable. As I shall argue below, optimizing welfare in a postindustrial setting will, none the less, require radical departures. . . .

A Third Way?

Up to a point, but not further, a ‘social democratic’ approach might furnish the basis for a new, positive equilibrium. A politics of collectivizing families’ needs (de-familialization) frees women from unpaid labour, and thereby nurtures the dual-earner household. And this reduces child poverty, and makes households better equipped to weather the storms of flexibilization, since they will usually have one member’s earnings to fall back on if the other is made redundant, needs temporary retraining, or suffers wage decline. Two-earner households have stronger social networks, and are less likely to run them down if one partner becomes unemployed.

A social democratic de-familialization strategy can reverse fertility decline if it helps employed mothers square the caring-work circle (mainly via daycare), and if it is willing to cover a good part of the opportunity costs of having children (which means expensive maternity and parental leave as well as generous child allowances). Working mothers may have ‘negative productivity’ in the sense that their earnings (plus public subsidies) exceed their output. But these costs are potentially recovered via their higher lifelong earnings, the smaller depreciation of their human capital, *and* the resulting fertility dividend.

A fundamental ‘postindustrial’ dilemma—generally overlooked in contemporary debates—is that families seem no longer inclined to assume the costs of bearing children. As a comparison between fertility in Southern Europe (with Japan) and Scandinavia suggests, the social price of children may be high indeed. Immigration is certainly an alternative, and herein lies a major reason behind North America’s (and Australia’s)

younger age profile. Yet, besides being a rather unrealistic scenario for Europe today, the immigration option poses hard dilemmas of its own. If one of our chief problems is how to absorb masses of low-skilled workers, this would only be compounded by Third World immigration. The advanced economies would clearly prefer highly educated immigrants, yet encouraging such a flow can hardly be considered solidaristic towards Third World nations.

Universalizing the double-earner household (with lots of children) holds yet another promise of welfare improvement. When we add its greater purchasing power to its desperate search for free time, it is a truly promising source of service consumption, from restaurants and leisure parks to child care and home-help services for their aged parents. Thereby families create jobs for waiters, park-keepers, child-minders, and home-helpers. Of course, access to child care is a pre-condition for dual-income families in the first place. This is exactly the point: services beget services; the double-earner household plays the role of employment multiplier. The employment multiplier of working mothers can be quite substantial, especially in those kinds of services that are labour intensive.

Encouraging families to consume more external services is therefore part of a potential ‘win-win’ strategy. Yet, as we know, relative costs can be a major constraint. This is why, despite the virtual universalization of two-earner households, lower-end personal services in Scandinavia grow very little. They would doubtlessly grow if costs were lower, but this implies greater wage inequality.

This is where *Homo socialdemocraticus* must yield to *Homo liberalismus*. Substantial service employment growth outside the public sector depends on flexibility and low wages, something that social democracy cannot easily accept. The social democratic strategy, as I have outlined it so far, will therefore not escape the fundamental ‘equality-jobs’

trade-off. We need therefore to resolve one last problem: the very same ‘equality-jobs’ trade-off with which we started. . . .

There is in truth only one way out of the impasse, namely to redefine what kind of equality we desire. *Homo socialdemocraticus* must be convinced that we cannot aspire for all kinds of equality at once; that some inequalities can be made compatible with some equalities.

The principle of equality that must go is exactly the same that emerged—most forcefully in Scandinavian social democracy—when welfare states sought to respond to the equality-crisis of the 1960s, namely the promise of equality for all ‘here-and-now’. In practice, this may not be so difficult a task. *Homo socialdemocraticus*, like many of his rivals, has surely held any number of lousy jobs in his youth. Like all Scandinavians, he left the parental home very early and lived for years on bread and water (this is what he tells his children). Yet, he is now a respectable citizen with a respectable career. Temporary deprivation is unimportant if it does not affect our life chances.

We can return to Schumpeter’s omnibus: always full, but always with different people. Everybody gets off at the next stop, or at least where desired. If, like *Homo socialdemocraticus*, we cling to a notion of equality for all, here and now, we shall never resolve the fundamental dilemma of our times. The kinds of inequalities that are inevitable in the world of *Homo liberalismus* can become acceptable, even welcomed, if they coincide with a welfare regime capable of guaranteeing all citizens against entrapment: no one should find him- or herself in an omnibus with locked doors.

Our search for a postindustrial welfare optimum requires, therefore, some kind of a mobility guarantee. What this will look like depends, in turn, on what are the chief causes of social exclusion and inferior life chances. A room full of academics and experts would, no doubt, draft an endlessly

long list of causes. Many standard reasons, such as physical or mental handicaps, do not concern us here because they are not inherently part of the equality–jobs trade-off. Two sources of substantial life chance problems stand out: one, the risks associated with marital instability and poverty in childhood; two, inadequate skills.

Diminishing family-induced risks calls for a standard ‘social democratic’ package of what the feminists call women-friendly policy: child-care services, incentives for mothers to work, and adequate income maintenance to take into account mothers’ reduced labour supply and the cost of children. Diminishing labour market induced risks calls for a rethinking of education, training, and marketable skills.

This is not the place to explore the broad issues of education. But there are a number of basic facts that all can agree upon: the returns to skills are rising, as we would expect in a world increasingly dominated by complex technology; the low-skilled are in rapidly declining demand. To exemplify the importance of skills, Bjorn (1995) shows that the probability of exiting from ‘marginality’ (basically unemployment or low pay) jumps by 30 points with vocational training, and 50 points with some theoretical training. Closing the skill gap is therefore an extremely effective way of catapulting people out of entrapment, of assuring good life chances. And it also pays off to society in the form of a more productive workforce.

There is clearly nothing earth-shatteringly novel about a call for more education. There is virtually no government or international organization today that does not advocate ‘active labour market policy’ or ‘life-long learning’. But there is widespread scepticism about their effectiveness. Active labour market policy with its ‘activation’ and training programmes does not always appear to pay off if by this we mean that the unemployed eventually find a stable, promising career. Too often, activation looks more

like a temporary parking-lot, or one interlude in a never-ending roundabout of unemployment, training, occasional jobs, and then unemployment again.

Effective or not, many believe that training is useless since there are no jobs for the newly trained. In a static sense this is undeniably often true. But they miss the basic point of the omnibus analogy: for any given individual, skills are the single best source of escape from underprivilege. What we are trying to resolve is a dynamic, life-course issue and not where to place everybody today. A new welfare optimum is, in fact, compatible with the possibility that many of us will experience a spell of unpleasantness.

What we really need to understand better is what kinds of skills and what kind of education to promote? We are generally aware of the rising demand for multi-skilling, flexible adaptation, and capacity for life-long learning. A worker with a good theoretical base in his vocational training is much easier to upgrade than someone with a high-school diploma unable to understand instructions on an aspirin bottle. This is why German employers use their skilled workers to run computerized production systems while in England they import engineers. We also face a reality where 'social skills' are more fundamental than muscle because of the direct face-to-face mode of service production. What these are is not easy to define, but they undoubtedly include more than the routine 'have a nice day'. A major problem in the contemporary unemployment structure is that laid-off miners and steel workers are unlikely to possess the kinds of social skills that sell a service.

We know from the OECD literacy studies (OECD, 1997) that education—even completed secondary education—may guarantee very little if it ends up producing 15 or 20 per cent Americans incapable of even rudimentary reading, writing, and arithmetic. These do not even possess the

minimal level of qualifications needed to be trainable. They are a *de facto* human capital waste, a stratum only too obviously condemned to lifelong low-wage employment or, possibly, crime.

It is this kind of result that must be eliminated from any kind of society if we seriously desire an optimal welfare regime. I therefore close this chapter inviting education experts to design a workable system of skilling entitlements, one that would befit an ideal postindustrial welfare regime. And I invite our political leaders to forge a new coalition of our assorted *homines*, one capable of breaking the deadlock of median-voter support for anachronistic modes of welfare production.

NOTES

1. The ratio of jobs lost to competition from the 'South' to jobs lost due to structural change is, for the period 1970–85: .07 in Denmark; .04 in France; .06 in Germany; .01 in Japan and the Netherlands; .05 in the UK; and .08 in the USA (calculated from OECD, 1994, i, Table 3.10).

2. In the more 'distribution-dominated' economies of Canada, the Netherlands, and the United States, manufacturing decline has been much more modest. Japan (and Denmark) actually experienced net manufacturing job growth. (Source: recalculations from OECD, *Historical Statistics*, 1995, Table 1.10).

3. The 'classes' in the industrial order are relatively unproblematic, but those in the servicing hierarchy need some clarification. Professionals are straightforward; semi-professionals refer to occupations such as social workers, nurses, technicians, and teachers. Skilled service workers embrace occupations for which a skill certification is required (hairdressers or nursing assistants, for example). Unskilled service workers are defined as those occupying jobs for which there are no particular skill requirements, tasks that, in principle, anyone could perform (cleaning, waiting, bell-hopping, car-parking, and so forth). For a detailed exposé, see Esping-Andersen (1993).

4. There is none the less evidence that this classification scheme has validity. Salido (1996, ch. 3) has tested it both in terms of construct and criterion validity against the Treiman prestige scale, and also against a battery of hierarchy criteria such as auton-

omy, decision-making, and supervision, and the results are quite comforting.

5. The job-growth calculations are based on data from ILO, *Yearbook of Labour Statistics* (various years).

6. Estimating the regression equation (using robust regression): Professional ratio = $5.377 - 1.235$ (service growth). T-statistics for the constant = 14.87, and for service growth = -6.93 . R-squared = .212 with $F = 48.07$.

7. The following argumentation lies close to the work of Gershuny (1978).

8. The ideas for this section of the chapter came from a week-end seminar with Jose Maria Maravall and Adam Przeworski (whom I thank for the idea and to whom I apologize for the result).

REFERENCES

- Bairoch, P. (1996), 'Globalization, Myths and Realities: One Century of External Trade and Foreign Investment', in R. Boyer and D. Drache (eds.), *States Against Markets*. London: Routledge, 173–93.
- Baumol, W. (1967), 'The Macroeconomics of Unbalanced Growth', *American Economic Review*, 57: 415–26.
- Bell, D. (1976), *The Coming of Postindustrial Society*. New York: Basic Books.
- Bjorn, N. H. (1995), 'Causes and Consequences of Persistent Unemployment', Ph.D. diss., Department of Economics, Copenhagen University.
- Braverman, H. (1974), *Labor and Monopoly Capital: The Degradation of Work in the Twentieth Century*. New York: Monthly Review Press.
- Brown, P. and Crompton, R. (1994), *Economic Restructuring and Social Exclusion*. London: UCL Press.
- Burtless, G. (1990) (ed.), *A Future of Lousy Jobs? The Changing Structure of U.S. Wages*. Washington: Brookings Institute.
- Clark, T. and Lipset, S. M. (1991), 'Are Social Classes Dying?', *International Sociology*, 4: 397–410.
- . Lipset, S. M., and Rempel, M. (1993), 'The Declining Political Significance of Class', *International Sociology*, 3: 293–316.
- Erikson, R. and Goldthorpe, John H. (1992), *The Constant Flux: Class Mobility in Industrial Societies*. Oxford: Clarendon Press.
- Esping-Andersen, G. (1990), *The Three Worlds of Welfare Capitalism*. Cambridge: Polity Press.
- . (1993) (ed.), *Changing Classes: Stratification and Mobility in Postindustrial Societies*. London: Sage.
- Gershuny, J. (1978), *After Industrial Society: The Emerging Self-service Economy*. London: Macmillan.
- Goldthorpe, J. (1982), 'The Service Class, its Formation and Future', in A. Giddens and G. Mackenzie (eds.), *Social Class and the Division of Labour*. Cambridge: Cambridge University Press, 162–87.
- Harrison, B. and Bluestone, B. (1988), *The Great U-Turn*, 2nd edn. New York: Basic Books.
- Hirsh, F. (1976), *Social Limits to Growth*. Cambridge, Mass.: Harvard University Press.
- Hout, M., Brooks, C., and Manza, J. (1993), 'The Persistence of Classes in Postindustrial Societies', *International Sociology*, 3: 259–78.
- ILO (1943), *Yearbook of Labour Force Statistics*. Geneva: International Labour Office.
- Jencks, C. and Peterson, P. E. (1991) (eds.), *The Urban Underclass*. Washington: Brookings Institute.
- Kern, H. and Schumann, M. (1984), *Das Ende der Arbeitsteilung?* Munich: C. H. Beck.
- Levy, Frank (1988), *Dollars and Dreams: The Changing American Income Distribution*. New York: W. W. Norton.
- OECD (1992), *Employment Outlook*. Paris: OECD.
- . (1994a), *The OECD Jobs Study: Part I and II*. Paris: OECD.
- . (1995e), *Historical Statistics*. Paris: OECD.
- . (1997c), *Literacy, Skills and the Knowledge Society*. Paris: OECD.
- Offe, C. (1985), *Disorganized Capitalism*. Cambridge, Mass.: MIT Press.
- Renner, K. (1953), *Nachgelassene Werke*, iii. *Wandlungen der Modernen Gesellschaft*. Wien: Wiener Volksbuchhandlung.
- Salido, O. (1996), 'La Movilidad Ocupacional Femenina en Espana: Una Comparacion por Sexo', Ph.D. diss., Universidad Complutense de Madrid, Facultad de Ciencias Politicas y Sociologia (June).
- Stigler, G. (1956), *Trends in Employment in the Service Industries*. Princeton: Princeton University Press.
- van Parijs, P. (1987), 'A Revolution in Class Theory', *Politics and Society*, 15/4: 453–82.
- Wood, A. (1995), *North-South Trade, Employment and Inequality*. Oxford: Clarendon Press.
- Wright, E. O. (1989), *The Debate on Classes*. London: Verso.

■ POST-SOCIALISM ■

103 ■ Victor Nee

Postsocialist Stratification

The transitions from state socialism in China, Russia and Eastern Europe provide natural experiments involving change in the stratification order on a scale reminiscent of that experienced in the West during the rise of capitalism. Comparative stratification research builds on modernization theory in assuming that industrial development, whether socialist or capitalist, leads to convergence (Lipset and Zetterberg 1959; Treiman 1970; Grusky and Hauser 1984; Hauser and Grusky 1984). Institutional theorists, however, are skeptical about theories of convergence driven by industrialism. If convergence pointed the way to state socialism's future, the sweeping measures to institute a market economy in industrially developed Eastern Europe and Russia would not have been necessary. Rather than focusing on the effects of industrial growth, institutionalists insist that research on state socialism needs also to take into account underlying differences in institutional forms. Such a focus on alternative institutional forms that provide a deep structure for economic action underlies my argument that the shift from redistribution to markets gives rise to different mechanisms of stratification.

New institutionalist theory in sociology maintains that institutions shape the structure of incentives and thereby establish the constraints within which rational actors identify

and pursue their interests (Nee and Ingram 1998; Cook and Levi 1990). It builds on Polanyi's (1944; North 1977, 1981) insight that economies are embedded in definite institutional arrangements. In Polanyi's view, the problem with neoclassical economics is that its formal models assume the existence of a market economy. Yet comparative studies of human societies, past and present, have demonstrated a variety of institutional forms giving rise to economies organized around fundamentally different operating principles (Polanyi 1957). Hence, the movement of goods and services in an economy is not simply a product of the aggregation of individual maximizing behavior, as assumed in neoclassical economics. In societies integrated by redistribution, goods and services are collected and distributed from a center in accordance with the customs, regulation, ideology, and ad hoc decisions of those social groups that hold redistributive power. Producers pass on a larger share of the economic surplus to the state precisely because they have less bargaining power over the terms of exchange than they would in a market economy. Instead, the state requisitions their products and compels them to work for prices and wages fixed by central bureaus (I. Szelényi 1978). As a consequence, the institutional logic of redistributive economies differs substantially from market economies where goods and

services are exchanged directly between buyer and seller (Polanyi 1957; Kornai 1980), and from economies based on reciprocity where trust and cooperation in local social orders allow for balanced exchange (Homans 1950; Sahlins 1972). Redistribution, market, and reciprocity, as alternative institutional forms, incorporate different structures of incentives and constraints and therefore distinct parameters of choice. Elements of each coexist in every society, at various levels, but Polanyi claimed that only one can constitute the dominant integrative mechanism of an economy. In this sense, the “economy” is not narrowly economic but is embedded in customs, social norms, laws, regulations, and the state. Fundamentally, differences in institutional framework are reflected in contrasting structures of property rights (Pryor 1973).

The transition from one dominant institutional form to another entails remaking the fundamental rules that shape economies, from formal regulations and laws to informal conventions and norms. The shift to markets—well underway but far from complete in postsocialist societies—involves changing structures of opportunity (Merton 1949). Whereas opportunities for advancement were previously centered solely on decisions made by the redistributive bureaucracy and within the economy controlled by it, markets open up alternative avenues for mobility through emergent entrepreneurship and the selling of labor power. Under conditions of expanding markets, economic actors strive to institute new rules of competition and cooperation that serve their interests, both through informal arrangements and through formal institutional channels. This entails efforts to change the structure of property rights in a manner that enables entrepreneurs and producers to capture a greater proportion of the economic surplus, previously transferred to the state by administrative fiat. Political actors contribute to instituting change in the formal rules of the game insofar as gains in productivity increase revenues to the state (North

1981; Nee 1992). Such action at the margins is cumulative and gradually results in transformation of the institutional environment.

Market transition theory (Nee 1989, 1991, 1996) maintains that as power—control over resources—shifts progressively from political disposition to market institutions, there will be a change in the distribution of rewards favoring those who hold market rather than redistributive power. Compared to nonmarket allocation, market exchange enhances the bargaining power of producers. Incentives are improved as producers retain a greater share of the economic surplus. Opportunities for gain depend less on the personal discretion of cadres in positional power. Consequently, the growth of market institutions (i.e., labor markets, subcontracting arrangements, capital markets, and business groups) causes a decline in the significance of socialist redistributive power even in the absence of fundamental change in the political order. In sectors of the socialist economy where a decisive shift to markets has occurred, officials are less likely to gain a dominant advantage from positional power in the state socialist redistributive bureaucracy. Contingent on continuing allocative control over key factor resources, their relative advantage declines as a function of the extent to which markets replace redistribution as the coordinating mechanism of the economy.

In the state socialist redistributive economy, officials act as monopolists who specify and enforce the rules of exchange by administrative fiat and exclude private entrepreneurs from taking part in legitimate economic activities. Because economic actors depend on resources allocated from above, they strive to secure favorable access to these decision makers in order to maximize their access to scarce resources (Walder 1986). By contrast, the shift to the market mechanism reduces dependence on superordinate bureaucratic agencies, as producers and consumers increasingly get needed goods and services through markets. Market-based exchanges stimulate the development of horizontal ties between buyers and sellers, in

labor and production markets, resulting in an incremental decline in the relative value of vertical connections—political capital—and an increasing importance of network ties between economic actors in society.¹ In sum, changes in the mechanisms of stratification stem from the expansion of opportunities for gain and profit centering on market institutions. Opportunities are more broadly based and diverse when markets replace and augment the opportunity structure controlled by the state. Moreover, markets provide powerful incentives for direct producers and entrepreneurs, whereas state socialist redistributive economies depress incentives because administratively set prices for labor lack sensitivity to differential performance.

The foregoing line of reasoning may be represented as four closely related hypotheses. These are as follows:

1. *Political capital hypothesis*: In market transactions, economic actors have the right to withhold their labor power or product in emergent markets until a mutually agreed upon price is set, and a greater share of the economic surplus is therefore retained by producers than in a redistributive economy. The relative returns to political capital accordingly decline.
2. *Human capital hypothesis*: There are also stronger incentives for individual effort because rewards are more closely related to individual productivity. This is likely to be reflected in higher returns to human capital, which is among the best indicators of human productivity.
3. *Sector mobility hypothesis*: Producers who shift into the marketized sectors of the transition economy are likely to experience higher returns to their labor and productivity than those who remain in the sectors of the transition economy still controlled by redistribution.
4. *Entrepreneurship hypothesis*: Finally, expanding markets open up alternative

avenues of socioeconomic mobility, giving rise to entrepreneurship as a mechanism for upward mobility into the postsocialist elite.

Critics of the theory advance the claim that departures from a centrally planned economy are likely to proceed without fundamental changes in the mechanism of stratification. According to proponents of this perspective, the initial advantages of the redistributive elite are such that this elite will come to dominate the postsocialist stratification order (Rona-Tas 1995; Walder 1996). The old redistributive elite, they argue, occupy strategic positions in political markets, which enable them to adapt flexibly to emerging markets and capture a disproportionate share of the rewards (Parish and Michelson 1996). In Russia, for example, officials and their cronies gained, by means of state-sponsored privatization, ownership rights over public assets—oilfields, mines, utility companies, state-owned enterprises, and so forth. The key issue here is the extensiveness of such power conversion relative to the frequency with which new elites emerge in the market economy. Market transition theory maintains that because an expanding market economy activates a substantially broader base of economic and social participation, even though a substantial element of the political elite succeed in converting political capital to economic capital, cadres as a social group still experience a relative decline. A recent study of the Hungarian elite confirms a very substantial decline in the representation of the nomenklatura in the composition of the postsocialist elite (S. Széleányi 1998). In urban China, where the communist party remains firmly in power, increasing relative returns to human capital are revealed not only in regressions on income (Zhou 2000), but in the lower chances of promotion for party members even in state-owned enterprises oriented to market action (Cao 1999).

Table 1 provides a meta-analysis of results reported in empirical studies testing market

TABLE 1
Summary of Empirical Studies Relevant to Market Transition Theory

| Study | Data | | | | Findings | | | |
|---------------------------|---------|------------------------------------|---------|--------------------|-------------------|---------------|-----------------|------------------|
| | Nation | Population | Year | Dependent Variable | Political Capital | Human Capital | Sector Mobility | Entrepreneurship |
| Szelényi 1988 | Hungary | Rural, whole nation | 1982–83 | Agricultural goods | Mixed | Yes | — | — |
| Nee 1989 | China | Rural, Fujian Province | 1985 | Income | Yes | Yes | — | Yes |
| Nee 1991 | China | Rural, Fujian Province | 1985 | Income | Yes | — | — | Yes |
| Peng 1992 | China | Urban & rural, selected areas | 1986 | Income | — | Yes | Yes | — |
| Rona-Tas 1994 | Hungary | Urban & rural, whole nation | 1989,91 | Employment | Inconclusive* | Yes | — | — |
| | | | | Income | Inconclusive* | Yes | — | Yes |
| Damanski and Heyns 1995 | Poland | Urban & rural, whole nation | 1987,91 | Income | Yes | Yes | Inconclusive | — |
| Parish, Zhe, and Li 1995 | China | Rural, eastern two-thirds of China | 1993 | Employment | Mixed | Yes | — | — |
| Nee 1996 | China | Rural, whole nation | 1989–90 | Income | Yes | Yes | Yes | — |
| | | | | Employment | Mixed | Yes | — | — |
| Parish and Michelson 1996 | China | Rural, whole nation | 1988 | Income | Yes | No | — | Yes |
| | | | | Employment | Inconclusive# | — | — | — |
| Xie and Hannum 1996 | China | Urban, whole nation | 1988 | Income | Inconclusive# | — | — | — |
| | | | | Income | Inconclusive# | Inconclusive# | — | — |
| Bian and Logan 1996 | China | Urban, Tianjin | 1988,93 | Income | Yes* | Yes | Yes | — |
| Gerber and Hout 1998 | Russia | Urban & rural, whole nation | 1991–95 | Income | — | No | — | Yes |
| S. Szelényi (1998) | Hungary | Whole nation | 1993 | Elite membership | Yes | No | — | — |
| Brainerd 1998 | Russia | Urban & rural, whole nation | 1991–94 | Income | — | Yes | — | — |
| Cao and Nee 1999 | China | Urban, two southern cities | 1994–95 | Income | Yes | Yes | Yes | Yes |
| Zhou 2000 | China | Urban, selected areas | 1995 | Income | Yes* | Yes* | Yes | — |

Note: “Yes” means that the finding is consistent with market transition theory, “No” means inconsistent and “—” means inapplicable.

“*” indicates that the finding reported here is based on our own interpretation of the results, which differs from the author’s.

“#” indicates that we disagree with the study’s research design and hence regard the results as inconclusive.

Source: Nee and Cao (1999, pp. 10–11).

transition theory. The findings confirm the theory's predictions of increasing relative returns to human capital, and its claim that emergent market economies open alternative pathways for upward mobility through entrepreneurship and labor markets in the private/hybrid sectors of the transition economy. However, empirical tests of the prediction of declining relative significance of political capital show mixed results, suggesting that elites in some postsocialist stratification orders have circulated more successfully than in others. Overall, the cumulative evidence from empirical studies points to path-dependent changes in the mechanisms of stratification largely in line with predictions proffered by market transition theory (Nee and Cao 1999).

The institutional changes that accompany the shift from a redistributive to a market economy comprise the causal mechanisms that cumulatively transform the stratification order. Although the transition entails hybrid stratification orders in which elements of the old elite retain and even augment their advantages, the secular trends of increasing returns to human capital and declining relative returns to political capital signal the decline of the redistributive elite. The more developed the market economy, the greater the breadth and diversity of opportunities that develop outside the boundaries of the old redistributive economy. Groups and individuals who were formerly barred from advancement in the state socialist bureaucracy and economy gain chances for social mobility through emergent labor markets and private entrepreneurship. To the extent that new bases of opportunity expand, resources become embedded in alternative networks and institutions, dependence on the established elite declines, and excluded groups gain in power relative to the established elite. Importantly, such shifts in power need not entail a direct transfer of power as in a regime change but occur as an unintended by-product of institutional change. As in the rise of capital-

ism in the West, such change in the stratification order occurs gradually over time.

NOTE

1. This argument is consistent with power dependence theories (Emerson 1962; Hechter 1987) and social resource theory (Lin 1982). As China shifts to market coordination, family firms and social networks will grow in importance in economic transactions (Hamilton 1991).

REFERENCES

- Bian, Yanjie, and John Logan. 1996. "Market Transition and the Persistence of Power: The Changing Stratification in China." *American Sociological Review* 61:739–58.
- Brainerd, Elizabeth. 1998. "Winners and Losers in Russia's Economic Transition." *American Economic Review* 88:1094–116.
- Cao, Yang. 1999. "Careers Inside Organizations: Comparative Study of Promotion Determination in Urban China." Unpublished manuscript, Department of Sociology, Cornell University.
- Cao, Yang, and Victor Nee. 1999. "Remaking Inequality: Market Dependency and Organizational Adaptation in Urban China." Working paper (Ithaca: Department of Sociology, Cornell University).
- . 2000. "Comment: Controversies and Evidence in the Market Transition Debate." *American Journal of Sociology* 105:1175–95.
- Cook, Karen Schweers, and Margaret Levi, eds. 1990. *The Limits of Rationality*. Chicago: University of Chicago Press.
- Damanski, Henrik, and Barbara Heyns. 1995. "Toward a Theory of the Role of the State in Market Transition: From Bargaining to Markets in Post-Communism." *European Journal of Sociology* 36:317–51.
- Emerson, Richard A. 1962. "Power-Dependence Relations." *American Sociological Review* 27: 31–41.
- Gerber, Theodore, and Michael Hout. 1998. "More Shock than Therapy: Market Transition, Employment, and Income in Russia, 1991–1995." *American Journal of Sociology* 104:1–50.
- Grusky, David B., and Robert M. Hauser. 1984. "Comparative Social Mobility Revisited: Models of Convergence and Divergence in 16 Countries." *American Sociological Review* 49: 19–38.
- Hamilton, Gary G., ed. 1991. *Business Networks and Economic Development in East and Southeast Asia*.

- Hong Kong: University of Hong Kong, Centre of Asian Study.
- Hechter, Michael. 1987. *Principles of Group Solidarity*. Berkeley: University of California Press.
- Homans, George C. 1950. *The Human Group*. New York: Harcourt Brace Jovanovich.
- Kornai, János. 1980. *The Economics of Shortage*. Amsterdam: North-Holland.
- Lin, Nan. 1982. "Social Resources and Instrumental Action." Pages 131–46 in *Social Structure and Network Analysis*, edited by P. Marsden and N. Lin. Beverly Hills, CA: Sage.
- Lipset, Seymour Martin, and Hans L. Zetterberg. 1959. "Social Mobility in Industrial Societies." Pp. 11–75 in *Social Mobility in Industrial Society*, edited by S. M. Lipset and R. Bendix. Berkeley: University of California Press.
- Merton, Robert K. 1949. *Social Theory and Social Structure*. New York: Free Press.
- Nee, Victor. 1989. "A Theory of Market Transition: From Redistribution to Markets in State Socialism." *American Sociological Review* 54: 663–81.
- _____. 1991. "Social Inequalities in Reforming State Socialism: Between Redistribution and Markets in China." *American Sociological Review* 56:267–82.
- _____. 1992. "Organizational Dynamics of Market Transition: Hybrid Forms, Property Rights, and Mixed Economy in China." *Administrative Science Quarterly* 37:1–27.
- _____. 1996. "The Emergence of a Market Society: Changing Mechanism of Stratification in China." *American Journal of Sociology* 100:908–949.
- Nee, Victor, and Paul Ingram. 1998. "Embeddedness and Beyond: Institutions, Exchange, and Social Structure." Pages 19–45 in *The New Institutionalism in Economic Sociology*, edited by M. Brinton and V. Nee. New York: Russell Sage Foundation.
- Nee, Victor, and Yang Cao. 1999. "Path Dependent Societal Transformation: Stratification in Hybrid Mixed Economies." *Theory and Society* 28:1–37.
- North, Douglass C. 1977. "Markets and Other Allocation Systems in History: The Challenge of Karl Polanyi." *Journal of European Economic History* 6:703–16.
- North, Douglass C. 1981. *Structure and Change in Economic History*. New York: Norton.
- Parish, William L., Xiaoye Zhe, and Fang Li. 1995. "Nonfarm Work and Marketization of the Chinese Countryside." *China Quarterly* 143:697–730.
- Parish, William L., and Ethan Michelson. 1996. "Politics and Markets: Dual Transformations." *American Journal of Sociology* 101:1042–59.
- Peng, Yusheng. 1992. "Wage Determination in Rural and Urban China: A Comparison of Public and Private Industrial Sectors." *American Sociological Review* 57:198–213.
- Polanyi, Karl. 1944. *The Great Transformation: The Political and Economic Origins of Our Time*. New York: Rinehart.
- Polanyi, Karl. 1957. "The Economy as Instituted Process." Pages 243–69 in *Trade and Market in Early Empires*, edited by K. Polanyi, C. Arensberg, and H. Pearson. Glencoe, IL: Free Press.
- Pryor, Frederick L. 1973. *Property and Industrial Organization in Communist and Capitalist Nations*. Bloomington: Indiana University Press.
- Rona-Tas, Akos. 1994. "The First Shall Be Last? Entrepreneurship and Communist Cadres in the Transition from State Socialism." *American Journal of Sociology* 100:40–69.
- Sahlins, Marshall. 1972. *Stone Age Economics*. New York: Aldine.
- Szelényi, Iván. 1978. "Social Inequalities in State Socialist Redistributive Economies." *International Journal of Comparative Sociology* 19: 63–87.
- _____. 1988. *Socialist Entrepreneurs*. Madison, WI: University of Madison Press.
- Szelényi, Szonja. 1998. *Equality by Design: The Grand Experiment in Destratification in Socialist Hungary*. Stanford, CA: Stanford University Press.
- Treiman, Donald J. 1970. "Industrialization and Social Stratification." Pages 207–34 in *Social Stratification: Research and Theory for the 1970s*, edited by E. O. Laumann. Indianapolis: Bobbs-Merrill.
- Walder, Andrew G. 1986. *Communist Neo-Traditionalism: Work and Authority in Chinese Industry*. Berkeley and Los Angeles: University of California Press.
- _____. 1996. "Markets and Inequality in Transitional Economies: Toward Testable Theories." *American Journal of Sociology* 101: 1060–73.
- Xie, Yu, and Emily Hannum. 1996. "Regional Variation in Earnings Inequality in Reform-Era Urban China." *American Journal of Sociology* 101:950–992.
- Zhou, Xueguang. 2000. "Economic Transformation and Income Inequality in Urban China." *American Journal of Sociology* 2000: 1135–74.

104 ■ Gil Eyal, Iván Széleányi, and Eleanor Townsley

Making Capitalism without Capitalists

There have been a number of schools of thought about which social groups would be best placed to take advantage of market reform after the fall of communism. Market transition theory¹ and research on socialist entrepreneurs² suggested that those who had been successful in the second economy during the late communist period would benefit most from the transition to capitalism. In contrast, theorists of political capitalism argued that former cadres were best placed to convert state-socialist privilege into economic capital.³ Against both these theories we contend that cultural capital became the dominant form of capital in post-communism. The coalition that governs post-communist societies is comprised of technocrats and managers—many of whom held senior positions in communist institutions—and former dissident intellectuals who contributed to the fall of communist regimes at the end of the 1980s.

This [chapter] offers a new theory of the transition to capitalism, by telling the story of how capitalism is being built without capitalists in post-communist Central Europe. We theorize *capitalism without capitalists* as a distinctive new strategy of transition adopted by technocratic–intellectual elites in societies where no class of private owners existed prior to the introduction of market mechanisms. Note, however, that capitalism without capitalists is not necessarily capitalism *without a bourgeoisie*. If one thinks of the bourgeoisie as plural—thus, if one conceives *bourgeoisies* as a social group composed of both possessors of

material property (the economic bourgeoisie) and possessors of culture or knowledge (the cultural bourgeoisie)—then one can claim—and we do so in this [chapter]—that post-communist capitalism is being promoted by a broadly defined intelligentsia which is committed to the cause of bourgeois society and capitalist economic institutions.

This approach to analyzing the transition to capitalism in Central Europe necessarily differs both from the classical social and economic theories of Adam Smith and Karl Marx and from twentieth-century visions of corporate, managerial, or other kinds of post-capitalist societies, such as those proposed by Ralf Dahrendorf, Berle and Means, or Daniel Bell. On one hand, classical theorists assumed that there must have been *capitalists before capitalism*. For this reason, these theorists expended much scholarly effort investigating the process of the ‘original’ or ‘primitive accumulation’ of capital. The rationale was that both logically and historically, private capital accumulation must have occurred before market institutions could operate.⁴ On the other hand, a recurrent theme in the study of existing capitalist systems, particularly since the 1930s, has been that the importance of individual private owners is waning. Observing the growth of large corporate organizations, the increasing role of financial institutions, or the growing importance of science/knowledge, varied theorists of ‘late’ capitalism have argued that the role of capital as the main source of economic growth is

Gil Eyal, Iván Széleányi, and Eleanor Townsley. *Making Capitalism without Capitalists: The New Ruling Elites in Eastern Europe*, pp. 1–2, 6–7, 32–33, 160–163, 229–231, 247–248. Copyright © 1998 by Verso Press. Reprinted by permission of Verso Press.

ending.⁵ There is little doubt within this group that the capitalist system itself remains robust; what is in contention is the shape of *capitalism after capitalists*. Our theory of the transition to capitalism in Central Europe borrows from both these theoretical traditions, but differs from and builds upon them by imagining the historical possibility of *capitalism without capitalists*.

First, like ‘capitalists *before* capitalism’, ours is primarily a theory of transition. Our central aim is to understand and explain how capitalism can emerge in an economic system with no propertied bourgeoisie. We want to know what agents are building post-communist capitalism, and on whose behalf and for what purposes they act. One possibility is that the technocratic elite of former state-socialist societies may turn itself into a new propertied bourgeoisie and thereby fulfil the classical condition for capitalist development. Indeed, some analysts claim—rather prematurely, in our view—that such a transformation has already occurred. Another possibility is that the liberal intelligentsia will act as an ‘intellectual vanguard of the economic bourgeoisie’, creating a new class of proprietors from agents other than itself. Having fulfilled its historic mission, the intelligentsia may then return to creative writing, research, or teaching, or it may keep managing capitalist enterprises owned by others.

Within this field of possibilities, however, it is not inconceivable that capitalism will be built without a class of individual proprietors. In this case, we would expect an economic system which looks like those envisioned by twentieth-century theorists of ‘capitalism *after* capitalists’. If it is true that the future of capitalist economies will be systems in which individual proprietors do not play a major role (and this proposition has been forcefully challenged by Zeitlin, Domhoff and others⁶), it is possible that Central European societies will emerge as corporate or post-capitalist societies without the historical intervention of a *grande bourgeoisie*. In other words, it is conceivable

that post-communist elites will take a historic short cut and move directly to the most ‘advanced’ stage of corporate capitalism, never sharing their managerial power (even temporarily) with a class of individual owners. . . .

Cultural Capital and Class in Post-Communism

Post-communist society can be described as a unique social structure in which cultural capital is the main source of power, prestige, and privilege. Possession of economic capital places actors only in the middle of the social hierarchy, and the conversion of former political capital into private wealth is more the exception than the rule. Indeed, the conversion of former communist privilege into a post-communist equivalent happens only when social actors possess the right kinds of capital to make the transition. Thus, those who were at the top of the social hierarchy under state socialism can stay there only if they are capable of ‘trajectory adjustment’, which at the current juncture means if they are well endowed with cultural capital. By contrast, those who relied exclusively on now devalued political capital from the communist era are not able to convert this capital into anything valuable, and are likely to be downwardly mobile. . . .

Our point of departure is Bourdieu’s theory of social structure, which we reconstruct using Weber’s distinction between rank and class societies.⁷ In this way, we hope to bring a comparative-historical dimension into Bourdieu’s framework, which was designed initially to explain reproduction rather than social change. (No criticism is implied here: Bourdieu’s research site is contemporary French society, where reproduction is the overwhelming trend; we study the recent social history of Central Europe where turbulence and large-scale social change have predominated.) From Bourdieu we borrow the ideas of three forms of capital, social space, and habitus. On this basis, we distinguish between social, economic, and cultural capital,

and we conceptualize social structures as ‘spaces’ which are differentially stratified by various distributions of these types of capital. Individuals ‘travel’ in these spaces, and if changes in the relative importance of one or another type of capital occurs in their lifetimes, they try to reshuffle their portfolio of different types of capital and convert devalued forms of capital into revalued forms in order to stay ‘on trajectory’. On this basis, there are several specific innovations we propose to explain the nature of social change in Central Europe.

1. *We conceptualize pre-communism, communism, and post-communism as three different stratification regimes defined by the dominance of different types of capital.* While Bourdieu used the three types of capital to describe contemporary French society, where economic capital is dominant and the other forms of capital are subordinate, in the ever-changing landscape of Central Europe during the past fifty years we distinguish three qualitatively different social spaces, each of which is defined by the dominance of a different form of capital.

2. *We claim that post-communism is a historically unique system of stratification in which cultural capital is dominant.* This enables us to develop a new theory of post-communist social structure which is consistent with our argument about the crucial roles played by the technocracy and former dissident intellectuals in the transition from socialism to post-communism. We define capitalism as a class-stratified system in which economic capital is dominant, and communism as a system in which social capital—institutionalized as political capital—was the major source of power and privilege. With the decay of state socialism and the rise of post-communism, the importance of political capital is declining, the role of cultural capital is increasing, and economic capital is sufficient only to locate its possessors in the middle of the social hierarchy.

3. *The transition to post-communism is a shift from socialist rank order to capitalist class strati-*

fication. Drawing on Weber, we conceptualize communism as a society based on rank order. Social capital was dominant, and resulted in a socialist form of patron-client relations. Compared to societies where economic capital was dominant—which we understand, with Weber, to be class-stratified societies—communist societies were examples of modern rank order. With this understanding, we can argue that the transition from communism to capitalism is, in principle, a transition from rank order to class society, and thereby introduce the dynamism of Weber’s historical sociology into Bourdieu’s structural analysis. Actually existing post-communism, however, as a system of ‘capitalism *without* capitalists’, occupies a middle position on this scale from rank to class. The *Bildungsbürgertum* [i.e., the educated middle classes] is neither a rank nor a class. Rather, it combines the characteristics and contains the possibilities of both logics of social stratification. . . .

The ‘New Class’ Project for the Fourth Time?

We have been careful not to call Central European post-capitalism a class society, and in particular we have explicitly avoided referring to technocrats and managers, or the intelligentsia as a whole, as a dominant class. Rather, the key to our theory of post-communist managerialism and the conception of ‘capitalism without capitalists’ is that the formation of a propertied bourgeoisie has been relatively slow: market institutions have developed much more rapidly in Central Europe than a capitalist class. It is precisely this weakness of a domestic propertied bourgeoisie which has made it possible for a power bloc, composed of different fractions of intellectual elites, to retain a hegemonic position. At the same time, our answer to the ‘Whodunit?’ question—that is to say, ‘Whose project is post-communism?’—is that it has been the project of the intelligentsia, or at least certain fractions of the intelligentsia. So it seems that the idea of the ‘New Class’ haunts

us—the post-communist power bloc is neither a new nor an old class, but it is composed of intellectuals, and these intellectuals are pursuing a power project.

During the last century—starting with Bakunin in the early 1870s, and stretching to Gouldner in the late 1970s—the idea of a ‘New Class’ of intellectuals dominating and leading society has haunted the social sciences. In the West, it was thought that intellectuals would replace the propertied bourgeoisie. In Central Europe, it was believed that they would replace the old-guard bureaucracy of communist regimes. And while successive predictions of the imminent rise of the intelligentsia to power proved false, social scientists never seemed to learn their lesson. They kept creating new theories about societies in which intellectuals would dominate. Such theorizing came in waves, was fashionable for a few years, and then discredited for a while, only to be reborn in a somewhat altered form some years later. In an essay published a few years ago, one of us suggested that the stubborn return of the idea of the New Class requires explanation.⁸ We argued that while neither the intelligentsia as a whole, nor any of its strata, ever succeeded in establishing itself as a new dominant class, it was probably historically true that intellectuals formulated such ‘projects’ for power. On this basis, we interpreted the different waves of New Class theorizing as critical or apologetic—albeit premature—generalizations of such power projects. It may be instructive to think about post-communist managerialism from this perspective, and to contrast it with earlier intellectual class projects.

Between 1870 and 1970 we can identify at least three distinct waves of New Class theorizing, and possibly three intellectual power projects.

We identify the first wave as the ‘intellectual class’ theory advanced by anarchists of the late nineteenth and early twentieth centuries. This theory contended that Marxism was really the ideology of an intellectual elite, which was try-

ing to use the working-class movement to smuggle itself into a position of class power. When a Marxism-inspired social order was established in the post-revolutionary Soviet Union—taking form as a dictatorship of the Stalinist bureaucracy rather than class rule by ‘socialist scholars’—this theory died a peaceful death. However, this does not mean that the anarchist analysis of Marxism as an intellectual ideology was completely without insight. From our perspective, it helps to explain why Central European intellectuals abandoned their *Bildungsbürgertum* project at the turn of the century, and why so many of them were attracted to left-wing radicalism.

The second wave of New Class theories—fashionable in a period roughly from 1930 to the 1950s—consisted of ‘bureaucratic-technocratic class’ theories. In the version which addressed the social conditions of advanced Western capitalism, New Class theorists argued that the technocratic and/or managerial stratum would fill the gap created by the decline of family capitalism and individual private property.⁹ We think these theorists were probably correct in identifying ‘New Class’ aspirations among managers and technocrats at this time, but they erred in taking the ideologues of managerialism at their word, and underestimating the power of the old moneyed class. Critics of managerialist New Class theories, like Zeitlin and Domhoff, were correct when they observed that managerialism in the United States and Western Europe failed because it confronted a strong propertied bourgeoisie.¹⁰ We tend to side with the critics in this debate, since there is strong evidence that the old propertied bourgeoisie is alive and well. Indeed, it seems that today the concentration of wealth in the United States and Western Europe is even more marked, and fewer people own more of the productive assets than ever before.

The third wave of New Class theories, created in the 1960s and 1970s, was comprised of theories of ‘the knowledge class’. In their right-wing version, these theories of the new

class interpreted the radical movements of the 1960s as a power plot by the counter-cultural intelligentsia to dominate society. By contrast, social scientists on the political Left emphasized that as science became an increasingly important factor of production, scientists would begin replacing property owners as the dominant group in advanced Western societies.¹¹ Some of these theorists, most notably Gouldner, were not thrilled by the prospect of a society ruled by a new class coalition of the technical intelligentsia and humanistic intellectuals merged together by the ‘culture of critical discourse’, but given the alternatives, they believed that it might be ‘our best card in history’. As the left radicalism of the 1960s faded away, however, and as counter-cultural intellectuals were replaced by Yuppies preaching neo-liberalism, this wave of theorizing also withered away.

What about post-communist managerialism? Is it possible that we are witnessing a fourth New Class project this century? Is it possible that intellectuals are indeed in power, even though it is ‘by default’ this time, since managers have no enemies able to resist their newly found dominance? Have the prophecies of Berle and Means actually turned out to be correct—not for the advanced West but ironically, for Central Europe? Or is post-communism, as we have analyzed it, simply a step in the transition to market capitalism as we know it from the advanced West? Or are we witnessing the transformation of existing elements into a new *Gestalt*, of which we have no adequate conception? We know that we are shooting at a rapidly moving target, and social scientists are poor snipers even when the target does not move.

In this context, it is difficult to answer the question of whether or not managerialism will last. It is difficult to predict the chances that the post-communist power bloc will reproduce itself and the social and economic order it governs. It is unclear whether or not the post-communist elite will resist the encroachments of international capital, and the

attempts of propertied middle classes to accumulate capital and take hold of the command posts of the economy. Reproduction is a possibility, however. It is conceivable that managers will consolidate their powerful institutional base by allying with national and international fiscal institutions and, with monetarism as their ideology, continue to govern post-communist capitalism as managers. The appropriate positions are ready, the adequate consciousness has been prepared, and the actors feel comfortable in their positions and are deeply committed to their ideology. Furthermore, in post-communist societies the alternatives seem to be difficult. The nationalization of property has been a process with far-reaching implications, and is possibly irreversible. As the old joke puts it: we know how to make fish soup from an aquarium, but how can we build an aquarium from a bowl of fish soup? East European economists enjoyed citing this joke before 1989, but they quickly forgot it after 1989. Indeed, one major underpinning of managers’ power is their claim to be the only ones with the knowledge of how to conduct the complex task of privatization. And the more complex the task, the more it is practically impossible to convert former public property into identifiable individual private property; thus the more power managers have. Thinking about post-communist managerialism as the ‘fourth wave’ of an intellectual project for power is useful, then, since it helps us to ask these intriguing questions. Moreover, thinking in this way means that we do not simply assume that managerial power is a transitional phenomenon; it allows us to ask: what will society be like if managerialism reproduces itself?

To push our luck even further, we may even wonder whether or not Central European post-communist developments have any relevance for the rest of the world. To put it differently: is Central Europe the future of the West? After all, Central European managerialism may not be all that different

from the way Western capitalism operates today: property rights are diffuse, managers exercise a lot of power, and monetarist ideologies are powerful and widespread. As we have already noted, however, these questions about the future are tricky ones, and social scientists are notoriously bad at answering them. Perhaps it would be best to pose the question in a different way and ask: 'If managerialism was a project during the interwar years, as Berle and Means suggested, and if it failed, why did it fail?' Approaching the issue in this way is likely to make us skeptical about the spread of managerialism around the world. Since we argued that managerialism was successful in Central Europe because it did not face powerful enemies like a large propertied bourgeoisie, we would predict that the chances for managerialism in the rest of the world, and especially in the countries of advanced Western capitalism, are slim indeed.

NOTES

1. Victor Nee (1989) 'A Theory of Market Transition: From Redistribution to Markets in State Socialism'. *American Sociological Review* 54(5): 663–81.

2. Iván Szelényi (1988) *Socialist Entrepreneurs*. Madison: University of Wisconsin Press.

3. Ákos Rona-Tas (1994) 'The First Shall Be The Last? Entrepreneurship and Communist Cadres in the Transition from Socialism'. *American Journal of Sociology* 100(1): 40–69.

4. Karl Marx (1919) *Das Kapital: Kritik der politischen Ökonomie*. Hamburg: Meissner; Adam Smith ([1779] 1976). *An Enquiry into the Nature and Causes of the Wealth of Nations*. Oxford: Oxford University Press.

5. For the classic statements on the growth of the large corporation and its consequences, see Adolf Berle and Gardiner Means ([1931] 1968) *The Modern Corporation and Private Property*. New York: Harcourt Brace Jovanovich; and Ralf Dahrendorf (1959) *Class and Class Conflict in Industrial Society*. Stanford, CA: Stanford University Press. For a discussion of the increasingly important role of financial institutions, see Neil Fligstein (1990) *The Transformation of Corporate Control*. Cambridge, MA: Harvard Uni-

versity Press; and for arguments which suggest that the role of capital is ending as the primary motor of economic growth in modern capitalism, see, for example, Daniel Bell ([1973] 1976) *The Coming of Post-industrial Society*. New York: Basic Books.

6. See, for example, Maurice Zeitlin (1974) 'Corporate Ownership and Control: The Large Corporation and the Capitalist Class'. *American Journal of Sociology* 79(5): 1073–1119; G. William Domhoff (1967) *Who Rules America?* Englewood Cliffs, NJ: Prentice Hall; G. William Domhoff (1970) *The Higher Circles: The Governing Class in America*. New York: Random House; and G. William Domhoff (1986) *Who Rules America Now? A View from the 1980s*. New York: Simon & Schuster.

7. Pierre Bourdieu ([1983] 1986) 'The Forms of Capital'. pp. 241–58 in *Handbook of Theory and Research for the Sociology of Education*, ed. John G. Richardson. New York: Greenwood Press; Pierre Bourdieu (1984) *Distinction: A Sociological Critique of the Judgement of Taste*. Cambridge, MA: Harvard University Press; Max Weber ([1915–21] 1978) *Economy and Society*. Berkeley: University of California Press, pp. 926–38.

8. Iván Szelényi and Bill Martin (1988) 'The Three Waves of New Class Theories'. *Theory and Society* 17(4): 645–67.

9. See, for example, Thornstein Veblen ([1919] 1963) *The Engineers and the Price System*. New York: Harcourt & Brace; Adolf Berle and Gardiner Means ([1931] 1968) *The Modern Corporation and Private Property*. New York: Harcourt Brace Jovanovich; James Burnham ([1941] 1962) *The Managerial Revolution*. Bloomington: Indiana University Press; Ralf Dahrendorf (1972) *Class and Class Conflict in Industrial Society*. Stanford, CA: Stanford University Press; John K. Galbraith (1967) *The New Industrial State*. New York: Houghton Mifflin.

10. Maurice Zeitlin (1974) 'Corporate Ownership and Control: The Large Corporation and the Capitalist Class'. *American Journal of Sociology* 79(5):1073–1119; G. William Domhoff (1967) *Who Rules America?* Englewood Cliffs, NJ: Prentice Hall; G. William Domhoff (1970) *The Higher Circles: The Governing Class in America*. New York: Random House.

11. Galbraith, in *The New Industrial State*, took this position, and to some extent, so did Daniel Bell ([1973] 1976) *The Coming of Post-Industrial Society*. New York: Basic Books. See also Alvin Gouldner (1979) *The Future of Intellectuals and the Rise of the New Class*. New York: The Continuum Publishing Corporation.

■ POST-MODERNITY AND HIGH MODERNITY ■

105 ■ John W. Meyer

The Evolution of Modern Stratification Systems

The sociological tendency, well represented in the chapters in this volume on “The Future of Social Stratification,” has been to see modern stratification systems in a realist vein. Real component individuals and groups, competing and cooperating in real interdependencies, create and change a system of inequalities. Real systems—and thus some sort of functionalism, broadly defined—are involved: To be sure, some theories may place particular stress on competitive interdependence (as in class-conflict models), and others may emphasize economic over other interdependencies, but nearly all share a vision of society as a system made up of interdependent parts.

For example, Kerr and his colleagues (Ch. 100) see the modern economy as evolving toward greater complexity and thus as functionally requiring changing forms of inequality (including emphases on education and merit). The core themes here are picked up and amplified by the other authors. Bell (Ch. 101) sees both economic and social complexity as creating (1) functional requirements for new levels of knowledge and competence and (2) a new set of strata linked to education. Although Gouldner (Ch. 32) has a similar vision, it is generated by class conflict as well as by differentiation, thus producing a recognizably similar, if darker, story. Like-

wise, this vision again appears in the work of Nee (Ch. 103) and of Eyal and his colleagues (Ch. 104), but now it plays out as communist regimes break down rather than as post-industrialism proceeds. Pakulski and Waters (Ch. 107), have similarly functional visions, but see the postmodern stratification system as built much more around the subjective and cultural than the realist model of industrial society. Culture, here, means taste, lifestyle, status, and subjective preferences. It is broadly linked to the expanding importance of education and the rise of the service sector, emphasized in all the chapters.

More critical writers, often on the left, use exactly the same lines of argument, but for them the functional requirements are those of some economic elites or political forces rather than society as a whole. Esping-Andersen (Ch. 102), and Western and Morris (Ch. 7), see these basic functional processes as operating to generate actual or potential increases in inequality and breakdowns in social welfare.

Despite their disagreements and conflicts, the authors of these lines of thought all share a general description and analysis of the evolution of modern stratification, a view rooted in a realist vision of modern society as a functioning system. Pakulski and Waters see

realist modernity as having evolved in more subjective and cultural ways, but retain the strong sense of an integrated functioning system. In all the arguments here, expansion produces differentiation and complexity—these require more reliance on education and ultimately new professional strata. Every story told in these chapters amounts, in essence, to either emphasizing or explaining the rising stratificational importance of education and its credentials. As the system expands around the world, modern forms of stratification rooted in education are found everywhere. The story is a standard sociological one, told with varying emphases on conflict and consensus; it is also very similar to the stories modern societies tell about themselves. We emphasize the latter point as of substantive importance in understanding modern societies, not so much as a criticism of stratification theories. If stratification theories function as cultural ideologies, this may help explain the spreading dominance of education.

The conventional story, conveyed repeatedly in this fashion, has some pronounced weaknesses: If expansion is the driving force, why does so much differentiation occur (e.g., in the modern Third World) even without expansion? If education and the new class are relied upon for technically efficient solutions, why can we not find better evidence for their actual efficiency (e.g., Berg 1971)? Above all, if stratification reflects real social requirements, why are the forms of stratification and mobility so strikingly similar across all the extreme variations of modern society?

There is a strong sociological tendency to leave cultural matters out of the equation, aside from some notions that individuals in society may have some socialized tastes in common. When culture is brought in, as in Pakulski and Waters, it is seen as subjective and individual in character—matters of taste and status—rather than institutionalized collective rules and models. Indeed, in a strain to

see education as functional, these authors almost all deemphasize it as a carrier of cultural content. The modern stratification theorist talks about education as human or cultural or social capital, not as culture per se. The idea that modern societies may with some collective self-consciousness be, through direct collective action, creating and changing their stratification system and rooting it more deeply in education, is not given much emphasis. “Class consciousness” is of interest as a dependent variable or as an intervening variable helping to account for various outcomes such as political mobilization. But this is class consciousness in the sense of Marx or Mannheim, who assume an underlying sociological reality and see some form of consciousness as its product. In research terms, it is class consciousness as a property of individuals and groups within society, not a property of the collective itself. More respect than imitation is given to Ossowski (1963), for whom the cultural aspects of stratification can be (1) independent variables and (2) collective or institutional models.

Inattention to collective and cultural aspects of stratification has been a substantial limitation in the field. It has led to extraordinarily mistaken theoretical analyses and predictions (as with the families of theories that have overemphasized the causal and descriptive role of economic dimensions in stratification). In addition, it has led to inattention to some obviously cultural aspects of stratification. In this chapter, corrective reflections are suggested. We offer cultural interpretations of modern differentiation, the rise of education and the service sector, and the overall homogeneity of modern stratification systems.

The cultural perspective we emphasize is an alternative to what we are calling *sociological realism* on several dimensions that in a broader discussion could be distinguished. We may note two of them here. First, modern sociological realists see society as made up of individuals and groups that have prior

and natural properties (e.g., motives and interests). Their strivings for equality and their inequality result from these properties and their interaction in social life. A more cultural view treats the existence and standing (and motives and interests) of these entities as highly collectively constructed—as the product of meaning and interpretation in modern schemes and scripts. The construction of modern citizens as formally equal, for example, is seen as an evolution in culture and theory rather than as either natural equalities or a result of interests in interaction (e.g., mystifying tactics or compromises of dominant elites or functionally necessary responses to social complexity).

Second, sociological realism tends to emphasize social systems as affected by high levels of interactive interdependence among their components—as really social systems of interdependent parts. A more collective and cultural view calls attention to the impact of cultural theories and ideologies on this putative interdependence: The high status and importance of many groups, such as the schooled professions, in modern stratification systems arise from cultural ideologies (which in the modern world commonly take the form of functional theories) as much as from “real” interactive dominance and dependence.

Conceptions of Culture

Part of the problem lies with a very limited conception of culture as individual attitudes or vague collective sentiments. In fact, most sociology deeply shares the contemporary bias that modern systems (unlike the traditional ones that anthropologists study) are so overwhelmingly articulated and real that they do not have very much culture: In other words, the rationalism and functionalism of the cultures of modern society are taken as true by the analysts. Pakulski and Waters see the postmodern system as hav-

ing greatly increased scope for cultural expression and stratification. They see culture as playing a more independent and autonomous role in the future, although the culture they envision tends to be individual and subjective rather than a system of collective rules.

Another way to put this is to note that social-scientific stratification theories are themselves core cultural elements of modern society. They are institutionalized as constitutional principles, but they also may provide cultural bases for criticism and opposition. Although this is obvious with Marxist theories, it is also the case with most others, which provide normative bases with which to defend, attack, or ignore inequalities of various sorts. Stratification theorists may understate the importance of the cultures of modern systems because stratification theories (though not necessarily only narrowly academic ones) are the central cultural elements involved. Recognizing the cultural character of stratification would weaken the realist vision of society as a real system. It would also undercut some of the scientific claims of stratification theorists: If stratification is cultural theory, the theorists (as classic phenomenological critics have had it) are studying themselves.

Our point here is not to support the often-noted special importance of some intellectuals and scientific professions in modern society. It is to note that explicit theories and ideologies about licit and illicit inequalities are deeply ingrained in the public discussions of the modern system. Most great political conflicts in this system are explicitly and articulately about equality and inequality: Contemporary opinion surveys suggest that even ordinary persons are able to go on at length about inequalities and the conditions under which they are legitimate. Indeed, it is for this reason that the system tends to pick out for great attention some of the ideas of intellectuals (e.g., Smith or

Marx) who might be quite obscure in other traditions. The culture of the system tends to make stratification theories important (rather than the theorists simply having some sort of mysterious dominance over the culture); carriers include everyone from ordinary members to political and economic elites and specialized intellectuals.

Our emphasis on a cultural reinterpretation of stratification can be pursued in either a strong or weak form; the difference between the two is immaterial here, for the most part. The strong form (Meyer 1989; Mann 1986; Hall 1986) treats the modern (Western and now worldwide) stratification system as having distinctive cultural roots in such themes as equality and progress. A weak form might argue that more narrowly “real” social forces produced such Western structures as high capitalism or political democracy, after which institutionalization turned such arrangements into cultural recipes to be repeated throughout the world. Either form is good enough for present arguments.

We briefly review the cultural postures built into the stratification theories of modern societies. That is, in discussing what seem to be the most prominent substantive themes in stratification theories, we recognize that modern systems are much affected by their cultures and conceive of stratification theories as prominent parts of these cultures. One important theme is that persons are formally and ultimately equal; another is that inequalities can be justified only by contributions to collective progress. Many lines of sociological thought incorporate such matters as aspects of social reality; we treat them as the cultural rules of modern collective actors. We go on to consider the consequences these cultural postures may have—the things that can be better explained when we recognize the cultural character of stratification theory in the modern system.

Sociological thinking routinely recognizes that stratification systems in the premodern

past were cultural constructions. Further, there is a tendency to imagine that in some future utopia (or dystopia) where inequalities are put right (or wrong), cultural principles could again hold sway. This is, obviously, a dramatic counterpoint to the aggressive realism about the present that is characteristic of Marxian theory. We apply the same perspective to contemporary modern society, seeing its stratification system as in part the realization of its cultural visions.

Stratification Principles

Clearly a first cultural principle built into the stratification models of modern theories and societies is the functional one. Society is a system—a rationally analyzable set of components organized as a purposive project—and inequalities among its parts are justified by inequalities in their contributions to collective goods and goals. For example, one can defend or criticize enhanced inequalities mainly by functional arguments. The chapters cited above (particularly those by Bell, Gouldner, Nee, Eyal, Kerr, and their colleagues) do just that: A greater return to private capital is required to promote more investment, or more pay and status are needed for teachers if we wish to select better ones and improve the performance of all, or the complex society requires more educated intellectuals at its center. The idea of social progress is obviously involved in the principle.

The idea that inequality is justified only or mainly by functional considerations is so firmly established (e.g., Rawls 1971) that sociological critics of particular inequalities must argue that they are not really functional but are rather products of tradition or power (as seen in Western and Morris or Esping-Andersen). In such analyses, socially useful achievement is set in opposition to two defective sources of inequality; discussing the issues involved is a modern (and a sociological)

obsession. One opposition is mainly with the past and with tradition; achievement is set against ascription in a drama that combines issues of social efficiency with those of social justice. Thus there is great attention to whether favorable evaluations or opportunities are provided to students or workers on merit or whether there is a direct effect of status background. Even if the analyses suggest merit is mainly involved, analysts go further—perhaps the standards of merit are status-biased, or perhaps the resources of those higher in status permit easier or more achievement. A second opposition is to set socially useful and legitimate achievement against power that may pervert goals or means (as when hegemonic capitalism is thought to substitute the requirements of capital for the good of society).

The issues involved here infuse the stratification literature as well as modern social discussion with much normative excitement. Minor details of the distribution of income or education become salient, and analyses of mobility give great attention to causal pathways distinguishable according to their justice and efficiency. On the one hand, there are many conflicts about what is really functional: Radical differences in the interpretation of the functions of private capital or of education can produce very different assessments. On the other hand, broad bands of agreement can be found: There are few defenders of extreme inequalities, of high levels of status consistency, or of high levels of status inheritance. No such defenses occur in the chapters here.

All this takes on more momentum given the second crucial principle of modern stratification (and theory): moral individualism. Individuals are to be ultimately morally equal, and justice is to be assessed in terms of inequality among individuals. The notion goes unquestioned in these chapters, and little reference is made to any possibility of natural inequalities between persons. There is considerable tension, naturally, between this

principle and the legitimation of inequality on functional grounds; we later suggest that this tension is a source of a good deal of modern social change.

The individualism of modern stratification theory and research is striking. The important inequality is, for instance, individual income inequality. We do not much care if the genders (or races or families or ethnic groups or groups of differing class origins) have equal incomes. In fact, calculations of the income of a gender, race, ethnic group, or family as a group are very rare, in contrast to stratificational thinking in societies in which corporate groups (e.g., clans or villages) are more real than individuals. We care only, on a per capita basis, if individuals of each gender (or class background or race or ethnicity) have equal incomes. When we attend to cross-national comparisons, we calculate not income differences among countries but income differences per capita. On some questions, other units have moral standing (as when each country has certain rights in international relations, or as when each family or town may be represented in certain activities), but the utterly dominant unit is the individual: Justice is assessed in terms of the equality or inequality of individuals.

The Impact of Stratification Ideology on the Development of the Modern System

Imagine, thus, that the social changes of the modern period are going on under the continuing scrutiny of collective actors and the intellectual advisors of those actors (such as the authors of the chapters in this book). Both actors and advisors, of course, are concerned about the principles of functionality for the collective good and the ultimate equality of individuals. This is not so unrealistic an image: Most modern stratification theorists, including the authors here, work with just these standards, and most (including most of the authors here) aspire to affect

public policy; more to the point, decision-making elites in modern systems operate in about the same way.

Theory of Progress

Some effects of the legitimation of inequality in terms of the rational pursuit of progress are noteworthy. First, it seems obvious that rationalistic and progressive ideas of the sort developed in the chapters here have played a role in supporting a great deal of modern social differentiation. Activities can be pulled out of social life and bundled into rationalized elements or roles; their inequalities thus are legitimated. The modern system has been profligate in its creation of new occupations, tasks, organizational forms, and so forth. Not all of this is easy to explain in terms of standard arguments about the real functional requirements of modern development: It is still utterly unclear what actual (as opposed to culturally defined) functional requirements call for such occupations as sociologist or such organizations as therapeutic ones. Rationalistic models not only helped justify the evolution of these minor institutions but also played a role in the historically crucial early Western institutionalization of the capitalist class. Thus, the modern tendency toward extreme differentiation of roles may reflect cultural legitimating pressures as much as instrumental functional requirements; this might help explain the exceptional differentiation characteristic of egalitarian U.S. society in contrast to some other modern systems.

Second, strong and highly culturally theorized notions of progress and the collective good help explain the homogeneity of modern stratification regimes (across country and time) despite great heterogeneity in immediate circumstances. The finding that occupational prestige systems and mobility regimes are relatively homogeneous has been an overpowering surprise in the research literature. One explanation has been a very broad functional image, but this makes little sense in light of the extreme variations in any plausi-

ble functional requirements among modern societies. A more realistic explanation is to see the cultural (functional) theories of modernity as widespread and as generating a good deal of stratificational isomorphism: National communities vary enormously, but modern stratification is legitimated in terms of ideal pictures of society, and these models are highly homogeneous (e.g., Anderson 1991 calls most modern societies “imagined communities”). This would also help explain why stratificational change (e.g., in the status of women or ethnic minorities, in the welfare correctives for the class structure, or in educational opportunity) tends to be global in character (Thomas et al. 1987).

Third, the institutionalization of ideas of rationality and progress can help explain the surprising worldwide rise of education as the critical dimension in all modern stratification systems. Education provides a clear legitimating account of sources of improved capacity and also locates this capacity in specific persons and groups (Meyer 1977). Thus, ascriptive and income/power considerations and even some of the functional ones of occupation tend to be replaced by education as the main dimension of status: Education tends to correlate more highly with more outcomes than the other dimensions do, and intergenerational educational correlations are usually higher than the others. Empirical researchers routinely build education into their measures, but both researchers and theorists are slow to let go of their nineteenth-century (cultureless) models of society and continue to use terms like *class* or *socioeconomic status* when education is the main phenomenon at issue in their measures. This peculiarity of labeling variables in modern stratification research is quite revealing of its ideological base in modern culture.

The worldwide rise in educational certification for all sorts of social positions has been discussed as resulting from functional requirements; this is the approach taken in the chapters above (especially those by Bell,

Gouldner, Nee, and Kerr and his colleagues). However, critics have noted the utter absence of evidence of such requirements and have seen the expansion as resulting instead from conflict processes (Collins 1979). The conflict explanation is incomplete. Why would so many forces make education, as opposed to other institutional settings, the new arena of competition and conflict? Education becomes the competitive arena in part because myths of rational progress have made educational institutions culturally central.

Finally, we may note the extraordinary—and in the main, unpredicted—rise of the service sector and its new classes: This, along with the rise of education, is the main theme of the chapters above, and in fact integrates them all. By and large, the essays attribute the expansion of the service sector to the functional requirements of the complex society. This is unconvincing because of the absence of empirical evidence that expanded service sectors are indeed functional in this sense. It becomes yet more unconvincing when one realizes that the expansion of the service sector is worldwide—little predicted by the evolutionary development of complexity. Some social theories attribute great feats of power and manipulation to the professions and the state but give no real explanation about where the power and capacity come from. The cultural commitments to rational progress previously noted may provide some explanation for why societies supposedly dominated by their economies (Marx) or their state bureaucracies (Weber) in fact turn out to have the professions (analogous to, of all things, a priesthood) as their leading groups.

Effects of Individualism

There have been several effects of the institutionalization of individualism and equality on stratification. A first consequence is the expansion of social citizenship and of the dimensions of social life it comprehends. Stratification theorists have been so engrossed in

the normative search for illicit inequality as to have avoided noticing the rise and expansion of a most important social status in terms of which all are formally equal. For example, Parsons (1970), until his last discussion of stratification, essentially omitted citizenship as an element. With the exception of Esping-Andersen and Pakulski and Waters (who treat the standard of equality as culturally constructed), the scholars represented here simply take for granted that persons are or should be naturally equal, and that inequalities are problematic and require justification. The chapters are written as if this point of view is special to the authors, rather than one of the most widespread cultural principles in the whole modern world, in which the banner of human rights has the highest standing.

Clearly, citizenship is now one of the more important dimensions of status throughout the world—and on such dependent variables as income perhaps the most important one. Obviously, claims for its expansion (in more welfare arenas) have been continuous throughout the modern period (and have tended to increase inequality among citizens of different countries and thus to redefine world stratification). Second, the importance of individual equality has tended to expand education. Not only is education the crucial certifier of inequality, but mass education is a most important constructor of equality as well. Third, the importance of individual equality has produced considerable expansion of protection of young and old persons (including extensions of personhood to prenatal periods). In this area, as in others, the rise of professional protectors and the growth of organizations with some type of protective agenda have been the most prominent developments.

Equal Persons, Unequal Roles

We next consider the impact of simultaneous emphases on equal individuals and unequal contributions to rational progress. Many ob-

servers have noted that the dualistic Western (and now worldwide) emphasis on both individual and collective good has been a mobilizing dynamic in modern history. We consider a few ways in which this works out for stratification.

Role Differentiation. The simultaneous stress on the inequality of activities (i.e., society-wide coordination) and the equality of persons generates a great deal of role differentiation. As already noted, theorists such as the authors of the chapters included here have tended to understate the cultural bases of such differentiation. However, differentiation clearly has served to protect central myths in the modern system in two ways: (1) Elaborate role differentiation and the tight organization of activity systems (e.g., in such arrangements as modern formal organizations) help to sustain functional myths and to define activity as primarily linked to social tasks rather than personal status, and (2) more important, the modern system notoriously differentiates activities and roles from persons.

The second point is critical because all sorts of functional roots of this in efficiency pressures (for the most part, unproven) are alleged by the theorists. It is more important to note how much such individual/activity differentiations help sustain the principle of individual equality. In a modern system, we can have a societywide coordination of activity (e.g., public tax policy or other political and economic decision making) implying the utter domination of the activities of millions or billions of people by the activities of a few, without any serious violation of the normative stratificational principle of equality. In a system of this kind, the activities control each other but not the persons: Differentiation is crucial to sustaining the principle of individual equality.

Consider how this works out in terms of measures of income inequality. As Lenski (1966) noted, agrarian societies (e.g., me-

dieval states) had very high levels of income inequality; the disparity is greatly lowered by the modern system. However, in the medieval world all the income of the state was attributed to the king personally, so that the king appeared to be very rich. In the modern world, a differentiation is established: A president gets only a modestly high salary (higher than the best sociologists but lower than the best economists) personally but may control a monumental budget as a property of the presidential role. If this budget were attributed to the president personally, as in medieval accounting, the modern system would show enormously high levels of inequality (a typical head of a modern state would be recorded as having up to half the national income). Neither the medieval accounting nor the modern differentiated one is wrong: The point is that because a great change in organizational accounting has occurred, enormous societywide political control can result with only modest levels of accounted interpersonal inequality.

As a result of the modern differentiation between person and role, decreasing proportions of monetary flows are recorded as individual income, and the well-known growth in all sorts of corporate actors occurs. Our point is that one pressure in this direction arises from the modern demand for individual equality, combined with the modern pressures for rationalization and progress. Differentiation permits increasingly equal individuals in a system of increasingly unequal and ordered activities: It involves, of course, an elaborate modern differentiation between the tastes and motives attributed to persons and those attributed to organizations (Meyer 1986).

Consider how both forms of differentiation—between persons and roles and between different roles—help sustain equality in a modern organization like a university. Some equalities in status and salary among professors are maintained, but some professors may control huge research grants, hiring

many servants and much equipment; in both social custom and in tax law, this is not attributed to their persons but to their roles. Further, if the inequalities among professorial roles become too great, a further differentiation occurs, such that the wealth of one professor is located in a specialized institute and controlled by the professor as director of the institute. This control can produce social accountings reflecting great equality of income and power among the professors as persons but great inequality among their roles or activities.

The role differentiation of the modern system (between roles and between roles and persons) solves cultural problems related to stratification theory and can be seen as having cultural or legitimating significance. A sociologically realist perspective treats all this differentiation as arising step-by-step out of the pressures of organizational life. Such a perspective has great difficulty explaining why similar differentiation occurs in many different places (regardless of variation in local problems or needs), is hard pressed to explain the fact that much of the differentiation seems only symbolic in character, and must resort to unsubstantiated arguments about the presumed efficiencies involved in differentiation. The idea employed here—that the functional theory involved is acting as a cultural principle—simplifies the problem.

What are the mechanisms by which cultural commitments produce this sort of differentiation? In our argument, they occur through the relatively articulate and explicit activity of collective actors: public political elites, intellectual and professional theorists, and the like. Modern systems are constantly under scrutiny by cultural theorists looking for precisely such illicit inequalities. Both tax laws and common opinion may penalize the insufficient differentiation of personal and organizational resources and reward instances of clear and “rational” differentiation.

The Rise of Educational Credentialism. The tension between equal individuals and unequal roles can be seen as one of the sources of ever-expanding educational credentialism and its related institutions. More and more social boundaries and mechanisms are installed to keep the structural inconsistency involved maintained. There are the educational credentials themselves, which can be used to legitimate the allocation of formally equal individuals to unequal roles. The legitimation problem then reduces to accounting for inter-individual variations in the “investment” in education. There is no shortage of ancillary social theories with this orientation. Moreover, there are also theories of variations in abilities as well as in tastes that treat role differentiation as a matter of individual choice. Such effects work through the mechanism of relatively articulate collective consciousness and action. Clearly, many groups advocate, and states often require, the replacement of less legitimated and more ascriptive criteria of role entry by educational credentials. This is argued to indicate merit, to deserve public respect and trust, and thus to enhance legitimate status.

The Personnel Society. All of these characteristics are institutionalized in an expanding modern personnel bureaucracy. A surprising feature of modern organizational development, from the point of view of most theories, has been the elaboration of the citizenship of persons within organizations. Many individual rights have expanded to provide organizational protection of individuals from potentially threatening inequalities. Such bureaucracy is seen as both more progressive and more just.

Conclusions

Two closely related themes arise from our discussion. One conclusion is about modern societies: Many aspects of their stratification

systems reflect cultural properties of the modern system, as with the expansion of education, the service sector, the personnel bureaucracy, and so forth. This was mostly unexpected by theorists emphasizing the economic or political realism of the modern system: From their point of view, much of the modern social structure reflects distortion, false consciousness, or mystification. The theories tend to be mistaken, precisely in their inattention to the modern system's sensitivity to and cultural analysis of its own stratification system.

A second conclusion is that the underlying culture of the modern stratification system is closely tied to the social-scientific analyses we call stratification theory. The obsessions of theory (e.g., with individual inequality and with the distinction between just and functional inequalities and unjust or power and ascription-ridden ones) are the main cultural themes of modern stratification.

These cultural commitments in fact drive social change. This can help explain many of the puzzling aspects of modern stratification regimes: their homogeneity across time and space, their constant shifts toward education and the service sector, their high levels of differentiation among roles (and especially between roles and persons), and the worldwide expansion in citizenship statuses.

We can illustrate these points by noting that in recent decades, the world itself has come to be conceptualized as a unitary social system or society. The shift to such a conception is rapidly changing public and scientific stratification analyses: Distinctions among individuals in different countries now require rectification (from the point of view of both dependency and human rights ideologies), and a differentiated set of world institutions is rapidly developing. The ultimate goal of stratification theories—both lay and scientific—is no longer the articulation of equal individual personhood combined

with social progress at the nation-state level; a bigger arena will presumably undercut some of the structures of the older one. For example, it may delegitimize some efforts to produce equality within developed nation-states at the cost of enhanced international inequality.

NOTE

This paper benefited from useful comments on earlier drafts by David Grusky, Ron Jepperson, David Strang, Marc Ventresca, and members of the Stanford Stratification Seminar.

REFERENCES

- Anderson, Benedict. 1991. *Imagined Communities*. 2nd ed. London: Verso.
- Berg, Ivar. 1971. *Education and Jobs: The Great Training Robbery*. Boston: Beacon.
- Collins, Randall. 1979. *The Credential Society: A Historical Sociology of Education and Stratification*. New York: Academic Press.
- Hall, John. 1986. *Powers and Liberties*. New York: Penguin.
- Lenski, Gerhard. 1966. *Power and Privilege*. New York: McGraw-Hill.
- Mann, Michael. 1986. *The Sources of Social Power*. Cambridge: Cambridge University Press.
- Meyer, John. 1977. "The Effects of Education as an Institution." *American Journal of Sociology* 83(July):55–77.
- . 1986. "The Self and the Life Course." Pages 199–216 in *Human Development and the Life Course*, edited by A. B. Sørensen, F. E. Weinert, and L. R. Sherrod. Hillsdale, NJ: Erlbaum.
- . 1989. "Conceptions of Christendom." Pages 395–413 in *Cross-National Research in Sociology*, edited by M. Kohn. Newbury Park: Sage.
- Ossowski, Stanislaw. 1963. *Class Structure in the Social Consciousness*. New York: Free Press.
- Parsons, Talcott. 1970. "Equality and Inequality in Modern Society, or Social Stratification Revisited." Pages 13–72 in *Social Stratification* edited by Edward Laumann. Indianapolis: Bobbs-Merrill.
- Rawls, John. 1971. *A Theory of Justice*. Cambridge: Belknap.
- Thomas, George, John Meyer, Francisco Ramirez, and John Boli. 1987. *Institutional Structure*. Newbury Park: Sage.

106 ■ Anthony Giddens

Social Justice and Social Divisions

The class structure and class divisions associated with the knowledge/service economy are quite different from those of the industrial age. The blue-collar working class was the largest class grouping in the old social order. Marx labelled the working class the ‘universal class’, and of course expected it to be the agent of revolutionary change. Today, the manual working class is very much a minority, and its numbers are set to reduce further as the proportion working in manufacture continues to diminish. The old working-class communities, that used to be the source of much local solidarity, have largely broken up. What used to be the ‘middle class’ has become much more differentiated, while the land-owning upper class has largely disappeared. Separate agrarian classes have more or less completely evaporated too.

New occupational divisions have come into existence, based upon the social and technological changes associated with the knowledge/service economy. About two-thirds of the jobs generated by the new economy are skilled—they demand a technical knowledge of information technology (IT) and other skills too. Such jobs are becoming more plentiful in relative terms. Over the period 1995–2004, the proportion of jobs in the EU15¹ demanding advanced qualifications in IT—‘Apple Mac’ jobs—went up from 20 percent to 24 percent. But many people have to work in much more routine ‘Big Mac’ jobs—serving in cafés, shops, supermarkets or petrol stations. The propor-

tion of such jobs may be declining compared to more skilled occupations, but it remains substantial.

The class structure of a post-industrial society looks as portrayed in figure 1. This picture is a generalized one, since different societies vary between themselves to some considerable degree. The percentages given are only rough guides, and refer to individuals, not households. The group at the top is an amorphous one of elites whose power and outlook are as much transnational as national, especially in Europe’s ‘global cities’. They may include substantial proportions of immigrants. The French banker working in the City of London leaves his office in the early evening. As he departs, a cleaner from the Philippines, who is regularly sending money to her family there, comes in to tidy up after him.

At least 50 percent of jobs in the knowledge/service economy demand a high level of cognitive and/or personal skills, stretching down into work performed by ‘wired workers’—those who use computers much of the day, but are not themselves IT specialists. All class categories can involve those who work for the state or in state-controlled enterprises. The distinction between those who work for the state and those who don’t tends to coincide with divisions in labour markets; but job security is normally highest for those in classes 1, 2 and 3.

One of the distinguishing features of the knowledge/service economy is the high pro-

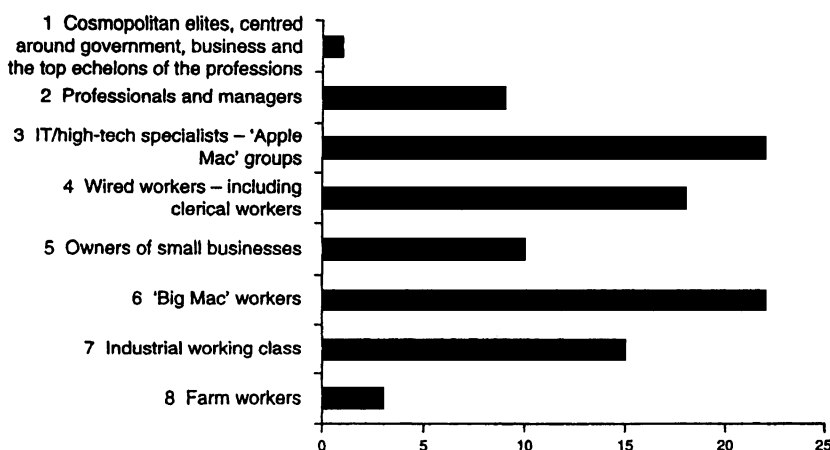


Figure 1. Classes in Post-Industrial Society (% of Population)

portion of women in the labour force—although few work in manual occupations. Women are heavily represented in ‘Big Mac’ jobs and among wired workers. They are also more likely to work part time than men are. Only 7 percent of men in the EU15 countries work part time, while 30 percent of women do. Women’s wages are on average about 15 percent lower than those of men doing the same or a comparable job.

These gross figures mask some important trends. The gender gap in pay is tending to become reduced almost everywhere. Moreover, at the bottom end women are actively doing better than men—declining wage levels and job security are much more marked among low-skilled working-class men than among women of comparable skill levels. Crucially, the role of women in providing household income is increasing dramatically. In Denmark, it is reaching parity with men, at 42 percent, although at the other end of the scale, in Spain and Italy, it is as low as 27 percent.²

Low pay does not necessarily mean that a person lives in hardship if he or she is a supplementary earner in the household. Household inequality is substantially lower than individual inequalities in most EU15 countries. Households most likely to be in poverty are single-earner households and es-

pecially households where no one is in work. The proportion of households in the EU15 with no one working ranges from 15 percent in the UK to 6 percent in Denmark.

The changing class structure has altered the nature of politics—and much else besides. In the old society, politics was largely shaped around the dividing lines between the manual working class and the rest. There were always ‘working-class Tories’, and social democratic parties needed to appeal to other groups besides the working class. But political activity correlated strongly with class, and voting preferences tended to be quite stable from election to election. Now the major political parties have to appeal to diverse constituencies, while many more voters are ‘dealigned’—they do not automatically stick to the same party.

The working class was (and is) very much divided along skill lines. Skilled workers, either craftsmen or those who worked in various areas of industrial technology, normally had quite stable jobs, and often were more affluent than clerical workers and others in the lower reaches of the middle class. Skilled workers in established trades that continue to thrive can still be in much demand—consider the fabled Polish plumber. However, the position of the unskilled has worsened. Unskilled workers—especially men—have

poor work opportunities, above all in areas where manufacturing industries have closed down. Even ‘Big Mac’ jobs involve face-to-face social skills that those from traditional working-class backgrounds may find it hard to master. Many men from such backgrounds are in any case unwilling to do what they see as ‘women’s work’. Unskilled men are thus highly vulnerable to spells of poverty or unemployment. This situation often applies with particular force to migrants or those from minority groups.

To refer to the situation of groups at the bottom of the socioeconomic scale, the term *social exclusion* has been widely popularized. The point of the concept is to recognize that factors other than poverty alone might prevent individuals or groups from playing a full part in society. The notion is usually traced to the work of the French author René Lenoir. The ‘excluded’ in his analysis included a variety of groups, making up about 10 percent of the population—not only the poor, but also the disabled, mentally ill, older people, drug users and so forth. The introduction of the concept was important, because it drew attention to the fact that the existing welfare provisions often did not reach the people in these categories; and because it stressed that it is not only sheer economic deprivation that prevents individuals from realizing their potential.³

At one time it became common to speak of the ‘50/40/10’ society in Europe: 50 percent of the working population are in stable jobs, 40 percent are in more insecure work, while the remaining 10 percent consists of the socially excluded, either unemployed or shifting in and out of the labour market (although the percentages varied in different versions). An underclass, in other words, has replaced the traditional working class, or the lower levels of it. A certain percentage of the population is cut adrift from the wider society.

However, this notion has proved to be incorrect.⁴ Sociologists are agreed that there is

no distinct underclass in this sense in the EU countries—although the notion may have more applicability in the US. Thus the government Strategy Unit in the UK studied four measures of social exclusion—those who are not in employment, education or training; those with a low income (below 60 percent of median income); those who have few established social ties with others; and those who perceive themselves as living in an area marked by high levels of crime, vandalism or material dilapidation. Only 1 percent of the UK population is excluded on all four of these measures. Multiple deprivation does exist, but it tends to be concentrated in specific neighbourhoods rather than affecting a ‘class’ of people.⁵ Social exclusion is a concept with some value, but it must not be used loosely and generically. It has essentially the same meaning as ‘multiple deprivation’, which tends to affect pockets of neighbourhoods and certain groups of individuals in particular (such as the homeless).

Class divisions in post-industrial society by and large are no longer behavioural, but are determined by differential life-chances. This change is a very significant one. At one time in most countries people in different class groups could be distinguished from one another quite visibly. Lifestyle variations still exist between those in different class groups, but they are often more influenced by taste and custom than by sheer financial constraints. Physical mobility, including foreign travel, is available to virtually everyone. The substantive freedoms most people have, centred around freedom of lifestyle choice, today are greater than for past generations. I shall call this phenomenon *everyday democratization*. Everyday democratization tends to stretch both ‘downwards’ and ‘upwards’ in the life-course. In the age of the Internet, childhood can no longer be as sheltered as it once was; and older people feel free to experiment with lifestyles just as much as younger ones do.

Everyday democratization does not necessarily bring greater security, or feelings of security. In fact a series of new insecurities come into being alongside it. Some are directly economic, others more social in nature. Most people want, and expect, more from their lives than previous generations did, leading to aspirations that cannot always be realized.

The degree of experienced security of different class groups tends to produce new ideological rifts in post-industrial societies. Such cleavages are partly based on rational anxieties (for example, fears of job loss) and partly on more free-floating worries. Those who flourish in the new economy tend to be happy with diversity, and embrace cosmopolitan lifestyles. Some may actively court what is seen negatively as insecurity. For instance, they may relish the prospect of moving on from job to job, and neither expect nor want to have a hierarchical career of the traditional sort. The creative industries, high-tech jobs, finance and banking and professional occupations are where such groups tend to be found in highest density.

What Richard Florida calls the 'creative class' now comprises well over 20 percent of the labour force, with a high concentration in certain metropolitan areas. It is mainly made up of people in classes 2 and 3 in my categories (see figure 1). The members of the creative class are diverse in terms of background—it includes people of all ages, from different ethnic groups and of differing sexual orientations.⁶

To measure the distribution of the creative class in the US, Florida developed a 'creativity index', based on four factors: the percentage of the creative class in the workforce; the percentage working in high-tech industry; innovation, as measured by patents per capita; and diversity, as assessed by the 'Gay Index', the proportion of gays being taken as a proxy of openness. Some cities, such as San Francisco, Austin or San Diego, rank high on all measures. Others are cities

bypassed by the creative class, such as Memphis or Pittsburgh.

Those cities near the top of the Florida list head the country in terms of economic prosperity and job creation. The liveliest local economies are characterized by the 'three Ts'—talent, technology and tolerance. They are above all *cosmopolitan*. The members of the creative class are very mobile, and will gravitate towards those cities in areas that offer what they want in terms of lifestyle. They prefer active, participatory recreation and street-level culture—a blend of cafés, restaurants, galleries and theatres. They are into a variety of active sports. They involve a high proportion of qualified migrants. According to one study, almost a third of all businesses set up in Silicon Valley in the 1990s were started by Indian and Chinese-born entrepreneurs.

Some cosmopolitans in Florida's sense may be found in other class groups too, given the available range of possible lifestyles. However, many who feel vulnerable to unwanted change are better characterized as *locals*, who rather want to stick to the existing order of things, or who look back nostalgically to the past—real or imagined. They might look for scapegoats, for example immigrants, upon whom to blame their troubles; and they might be attracted by political populism and economic protectionism. Immigrants themselves are not necessarily cosmopolitans. In their attitudes they may become locals, just as hostile to further immigration as groups in the native population.

Coupled to the other changes, especially in the sphere of the family, the changing class structure alters the distribution of 'at risk' groups, as well as the nature and form of inequalities. The conditions generating 'at risk' groups are structural, but how far they translate into real vulnerabilities depends upon the policy mix of a particular society, as well as specific blockages it might have. The other side of risk is *opportunity*. We must not make the mistake of supposing that risk is always a

negative factor. Some of the most important changes are the following:

1. There are on average fewer jobs with secure tenure for people in most work situations, although there are variations, depending upon the strength of insider/outsider divisions in labour markets.
2. Risk (and opportunity) are distributed differently across the life-span from how they were in the past. Risk and opportunity do not just 'happen' to people. More and more people think strategically about their lives in terms of future possibilities—including the decision to have a child/children. Transitions at different phases of life are both less predictable and less mechanical than they used to be. Heightened rates of divorce and separation mean that transition can happen at diverse times.
3. The intensity of technological change, conjoined to a more globalized division of labour, creates new vulnerabilities for some groups. Young men with no qualifications, as mentioned, are likely to fare especially poorly. Older workers in manufacture, whose jobs disappear, risk long spells of unemployment, or the prospect of never working again, unless appropriate policy interventions are made.
4. On average, older people hold a greater share of overall wealth and income than they did in the past, compared to the young. Many, however, are still at risk of poverty on retirement. Older women living on their own are the most vulnerable. However, risks have cascaded down more towards the young rather than the old.⁷ Child poverty has become commonplace where child care facilities and opportunities for work for women are underdeveloped.
5. 'Big Mac' jobs tend to offer little chance of a career. This situation may not matter where they are carried out by groups who have the capability to move into other sectors—such as students who work in coffee-shops during vacations or in gap years. It can matter a great deal for those who have few qualifications to allow them to move on.
6. In the knowledge/service society, credentials—certificates, diplomas, degrees—become of the first importance for career mobility. It is harder for affluent parents to pass on their advantages in a 'direct' way to their children. Hence they concentrate heavily on education. For example, university graduates on average earn significantly more over the course of their careers than those who do not go to university.
7. Ethnic minorities may be significantly at risk where they lack qualifications. Prejudices can be reinforced, and to some extent taken over by the members of the minority themselves. There can be additional problems for women where traditional beliefs confine their role to a domestic one.
8. Not only do many (most) women work, and for much of their lives, but their incomes are quite often crucial for sustaining the standard of living of a family. In an increasing proportion of cases they are the prime earners. This situation is not the cause of Europe's low birth rates. On the contrary, those families with the highest proportion of women in work are also ones with the most elevated birth rates. However, work/life issues become of great importance. Women are still the main carers and their careers are interrupted by having children much more than men's are.
9. There is a great deal of fluidity in contemporary society, but structural sources of mobility are different from a generation ago. There will not be as

clear a 'direction' to mobility between the generations as there used to be when many people from working-class backgrounds moved into white-collar and professional jobs. Structural mobility will depend upon a continuing overall upgrading of knowledge-based jobs at the expense of less-skilled service occupations. There are likely to be many more voluntary and involuntary career transitions, often involving a move sideways to a different job area, or taking time out from work for education or retraining.

10. Ageing is changing its nature, not just because more people will work longer, but also because there will not be the same discontinuities as before between working and non-working life. The old-age 'retirement ghetto'—a form of social exclusion if ever there was one—is being broken down. The very notions of retirement, and even pensions, might disappear in the future. In their place will come more flexible attitudes towards work and more orthodox sources of social support for those considered to be significant at-risk groups among older people, including especially many older women.

'Social justice' is a notoriously controversial notion. How far does it imply the redistribution of wealth and income, as compared to enhancing equality of opportunity? What can it actually mean in the context of a post-industrial society? The first of these questions can actually be quite easily answered on the level of principle. In a society that depends upon a highly dynamic marketplace for its prosperity, aspiration, ambition and opportunity have to be central. Equalizing opportunities is important because it makes the best use of available talents. However, reducing inequalities of opportunity necessarily involves redistribution, since otherwise those who are successful in one generation could

simply hold on to the fruits of their success. In the EU countries, post-tax inequality is significantly lower than pre-tax inequality, a desirable and necessary outcome since it sets a framework for other measures. Further redistribution occurs through the welfare system and through the direct effects of policy measures, depending on the welfare mix a society has.

How social justice should be defined is a difficult issue, about which large philosophical tomes have been written. However, a very useful working definition has been provided by the German political scientist, Wolfgang Merkel. He lists five priorities of social justice in post-industrial social conditions:⁸

1. The fight against poverty—not just because of economic inequality itself, but on the grounds that poverty (above all, enduring poverty) limits the individual's capacity for autonomy and self-esteem.
2. Creating the highest possible standards of education and training, rooted in equal and fair access for all.
3. Ensuring employment for those who are willing and able to work.
4. A welfare system that provides protection and dignity.
5. The limiting of inequalities of income and wealth if they hinder the realization of the first four goals or endanger the cohesion of society.

Obviously, the devil is in the detail, especially in respect of point 5. But the formula provides a down-to-earth scheme that is both simple and clear. It recognizes the role of equality of opportunity in a differentiated society, against the backdrop of current economic imperatives. It quite rightly gives pre-eminence to the struggle to reduce levels of poverty, since poverty limits life-chances and the capacity for self-realization. It follows from the formula that targeting child poverty is of especial significance. The higher the

proportion of those who suffer poverty as children, the more likely it is that all five goals will be compromised.

NOTES

1. EU15 comprises the following 15 countries: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden, United Kingdom. They are member countries of the European Union prior to the accession of 10 candidate countries on May 1, 2004.

2. Gøsta Esping-Andersen, 'Inequality of incomes and opportunities', in Anthony Giddens and Patrick Diamond (eds.), *The New Egalitarianism*. Cambridge: Polity, 2005, pp. 22–4.

3. René Lenoir, *Les exclus: un français sur dix*. Paris: Seuil, 1974.

4. See John Goldthorpe and Abigail McKnight, *The Economic Basis of Social Class*. London: Centre

for Analysis of Social Exclusion, London School of Economics, 2004.

5. Prime Minister's Strategy Unit, *Strategic Audit of the UK*. London, 2003.

6. Richard Florida, *The Rise of the Creative Class*. New York: Perseus, 2002. For the author's subsequent reflections, see Richard Florida, *The Flight of the Creative Class*. New York: HarperBusiness, 2005.

7. Esping-Andersen, 'Inequality of incomes and opportunities'.

8. Wolfgang Merkel, 'How the welfare state can tackle new inequalities', in Patrick Diamond and Matt Browne (eds.), *Rethinking Social Democracy*. London: Policy Network, 2004. Merkel's standpoint draws upon that of Amartya Sen. Sen argues that policies concerned with furthering equality should centre upon the 'capability set' of an individual—the overall freedom a person has to pursue his or her well-being. Amartya Sen, *Inequality Re-examined*. Oxford: Clarendon Press, 1992.

107 ■ Jan Pakulski and Malcolm Waters

The Death of Class

"Class" has always been a contentious concept. In its most famous Marxian version, it is an explanatory term linking the economic sphere of production with the political and ideological superstructure. Here it is also a label for collective social actors, especially for that new historical subject, the proletariat, whose destiny was to transform social consciousness and to turn class society into a classless one. This mixture of sociology and eschatology proved exceptionally potent in propelling the concept to celebrity status. Class became a master term for both sociology and political analysis. With the formation of class parties of the left, the concept was further implicated within the global and national, political and ideologi-

cal confrontations that became particularly acute during the hot and cold wars of the first three-quarters of the twentieth century. This was especially true in Europe where class became an ideological icon, a battle standard for the left and a bogey for the right. The intellectual left became preoccupied in identifying class divisions and condemning class inequalities, providing the central dogma of radicalized social science throughout the 1960s and 1970s. By contrast, denying the centrality of class became the hallmark of "bourgeois" sociology. Even if the term was used in an ideologically neutral context, the semantic-ideological halo it acquired made it inherently contentious and prone to political appropriation.

The degree of political appropriation has reduced in recent years. With the declining commitment to Marxism, the collapse of Soviet communism and the waning appeal of socialist ideologies in the West, class is losing its ideological significance and its political centrality. Both the right and the left are abandoning their preoccupation with class issues. The right is turning its attention to morality and ethnicity while the critical left is becoming increasingly concerned about issues of gender, ecology, citizenship and human rights. This rearrangement of political concerns coincides with a shift in intellectual fashion and a growing scepticism about the compatibility of class models with contemporary social reality. Class divisions are losing their self-evident and pervasive character. Class identities are challenged by “new associations” and new social movements. Class radicalism is no longer the flavour of the month in the intellectual salons and on university campuses. Like beads and Che Guevara berets, class is *passé*, especially among advocates of the postmodernist avant-garde and practitioners of the new gender-, eco- and ethno-centred politics.

Naturally, these shifts in public attention and intellectual fashion cannot be decisive in debates about the objective significance of class. However, they are important because they create an auspicious climate for reaching conclusions on its relevance. For the first time in over half a century, such a debate can be conducted outside the trenches of the Cold War and in a spirit of impartiality and mutual respect. In this new climate, arguments about the declining salience of class can no longer be dismissed as symptoms of ideological bias, intellectual weakness or moral corruption. . . .

The debate about classes combines issues of semantics and substance. Although our focus is firmly on the latter we must also pay some attention to the former. This is quite natural in a situation where a concept has been so stretched that it can be used to mean

almost any identifiable collectivity ranging from occupational position to political association, and any theoretical object from structural location to collective actor. It is impossible simultaneously to give justice to this semantic diversity and to present a consistent argument against the class paradigm. We therefore focus on analytic usages of “class” rather than the confused and vague applications of the term in popular discourse. In the context of the latter, “class” is bound to remain a routinized and vague classificatory term because it is a useful shorthand for a wide range of social entities and divisions. We respect this popular usage but refuse to accept it as evidence of, or an argument for, the relevance of class. We will also avoid defining class in such a restrictive way that it becomes a straw person, on one hand, or can be used to designate any structured inequality or conflict, on the other. The latter strategy transforms the class thesis into a virtual tautology because no known society is egalitarian and free of conflict. However, we do admit that there is a range of legitimate meanings of the term, each of which can be examined independently in terms of its theoretical utility and empirical relevance.

In the analytic vocabulary of contemporary sociology, “class” refers to a specific social location and causality, a specific pattern of groupness, and a specific form of identification. In the hands of some authors, as in classical Marxism, these aspects are combined, while for others, mainly those who subscribe to Weberian and action approaches, they are separated. However, we subscribe to the broad consensus that class is primarily about *economic-productive* location and determination; that is, it is based on property and/or market relations. People participate in class as producers rather than, say, as consumers, members of gender categories or organizational position-holders, although the impact of class may extend far beyond production roles. When stripped of this economic connotation and detached from its

classical Marxist and Weberian roots, “class” loses most of its explanatory power.

It is equally untenable to treat class as a mere substructure or as a statistical category, without at least drawing some implications for social relationships, social distance, patterns of interaction and patterns of association. Class is sociologically important as a social entity, as *social* class, and as such must transpire in detectable patterns of exploitation, struggle, domination and subordination, or closure. The number of classes, the extent of polarization, the clarity of class boundaries and the extent of conflict between classes may vary because they are not definitional elements. However, a minimum level of detectable clustering or groupness is essential if we are to say that classes exist. A society in which such economically caused social clustering occurs, and where these clusters are the backbone of the social structure, is a *class society*.

Lastly, a fully formed class involves a measure of self-recognition and self-identification on the part of participants, although this identification need not involve explicit class terminology or imply a consciousness of antagonism and struggle. Minimally, the members of a class have to be aware of their commonality and employ some recognized terms for collective self-description. A sense of difference between “them” and “us” is a necessary condition for the formation of class actors that marks the most developed examples of class articulation.

All of this implies that the articulation of class is a matter of degree. Its strength, and the extent to which it structures society, can vary from a minimal level, where economically based social clustering is weak but still predominant, to an advanced level. Here classes are well articulated in the social, cultural, and political-ideological domains, class identifications predominate, class consciousness is acute, and politics is dominated by struggles between class-based groups and organizations. Few historical instances of whole

societies approximate a fully fledged class configuration. More typically only a degree of “classness” can be observed. Moreover, the strength and the pattern of classness can vary historically and between contemporaneous societies.

The issue of the degree of classness of a society is linked to the issue of the relevance of “class analysis”, a type of social analysis that seeks to identify classes and to trace their social, cultural and political-ideological consequences. Class analysis makes sense only when applied to a class society. Its utility is proportional to the degree of classness. At one extreme, in a fully fledged class society, class analysis can be a legitimate substitute for social analysis. However, if classes are dissolving, a form of analysis that privileges class as an explanatory category has to give way to a more open-ended social analysis.

Historically there have been numerous examples of non-class societies. We embrace the convention that societies based on slave labour, such as ancient Rome and Greece, the estate societies of feudal Europe and modern state-socialist societies are non-class societies. In none of these societies are property and market relations the skeleton of the social structure or the predominant grid of social power. They are all unequal, stratified and conflictual but not made so predominantly by class. Class and class society are, in our vocabulary, distinctly modern phenomena inseparably linked to the market and its institutionalization within the early and mature forms of industrial capitalism.

This leads us to the central argument of this chapter, that classes are dissolving and that the most advanced societies are no longer class societies. The most important aspect is an attenuation of the class identities, class ideologies, and class organizations that framed West European corporatist politics in the middle of this century. Equally, the communal aspects of class, class subcultures and milieux, have long since disappeared. The issue of the decomposition of economic class

mechanisms is more controversial. We cite the following developments: a wide redistribution of property; the proliferation of indirect and small ownership; the credentialization of skills and the professionalization of occupations; the multiple segmentation and globalization of markets; and an increasing role for consumption as a status and lifestyle generator.

We must stress our concentration on the advanced societies of the capitalist West and, to a lesser extent, on the “newly industrializing countries” (NICs) of East and South East Asia and the rapidly democratizing and marketizing societies of Russia and Eastern Europe. Quite clearly, class remains salient in the “less developed countries” (LDCs) of Asia, Africa and Latin America that have not reached an advanced social, political and economic stage. In so far as such societies are structured by productive industrial property they clearly remain class societies.

The historical side of the argument charts the process of class decomposition in the industrialized West over the last century and a half. We argue that class societies are specific historical entities. They were born with industrial capitalism, changed their form under the impact of organized or corporate capitalism, and are disappearing in the face of post-industrialization and postmodernization. The early forms of class were chronicled in England by generations of social observers from Engels (1892) to Thompson (1980). They arose in parallel with a gradual decline of such pre-industrial collectivities as village communities and estates and the strengthening of nation-states. From the early twentieth century onward the fate of classes, nation-states and parties became intimately linked. While localized communities of fate were gradually marginalized, politically organized national classes were gradually institutionalized mainly as class parties and trade unions. These national classes were quite different from their predecessors. They were politically generated rather than spontaneous, occupa-

tionally heterogeneous, and reconstructed as imagined communities on a national level. Participation in these national classes involved little by way of interaction and communication. Their constitutive common interests were abstract political constructs developed by elites. What such classes lacked in social bonding and cultural cohesion, they compensated for in political organization and strategic coherence. It was these politically organized national classes, the power of which peaked with the establishment of corporatist structures, that became the most powerful social actors and agents of social transformation of the twentieth century. . . .

In summary, the stratification order of capitalist societies can be traced as a succession of three periods roughly demarcated by the nineteenth century, the first three-quarters of the twentieth century, and the contemporary period. In highly formalized and abstracted terms they are as follows.

Economic-class Society. This is a society arranged into patterns of domination and struggle between interest groups that emerge from the economic realm. In the familiar terms of Marx, the classes will be property owners and sellers of labour power, but they can be conceptualized as employers and employees. The dominant class can control the state and maintain itself as a ruling class either by capturing its apparatuses, or by rendering them weak. In so far as the subordinate class undertakes collective action, it will be rebellious or revolutionary in character, aimed at dislodging this ruling class by the abolition of private property. Culture is divided to match class divisions, into dominant and subordinate ideologies and into high and low cultures.

Organized-class Society. This type of society is dominated by a political or state sphere. The state is typically ruled by a single unified bloc, a political-bureaucratic elite, that exercises power over subordinated masses. The

bloc may be factionalized horizontally into formally opposed parties. The elite will comprise either a party leadership or a corporatized leadership integrating party leaders with the leaders of other organized interest groups including economic and cultural ones. The elite uses the coercive power of the state to regulate economics and culture. The state can dominate the economy by redistribution or by the conversion of private into public property, although this need not be a complete accomplishment. Masses, in turn, reorganize themselves in national-political classes rather than in industrial terms by establishing links with milieu parties. Meanwhile, the cultural realm can be unified under the state umbrella or under the aegis of state-sponsored monopolies. It can thus be turned into an industrialized or mass culture.

Status-conventional Society. In this type of society stratification emerges from the cultural sphere. The strata are lifestyle- and/or value-based status configurations. They can form around differentiated patterns of value commitment, identity, belief, symbolic meaning, taste, opinion or consumption. Because of the ephemeral and fragile nature of these resources, a stratification system based on conventional status communities appears as a shifting mosaic which can destabilize the other two spheres. The state is weakened because it cannot rely on mass support, and the economy is weakened (in its capacity for social structuring) because of the critical importance of symbolic values. Each order is de-concentrated by a prevailing orientation to values and utilities that are established conventionally rather than by reference to collective interests.

In failing to recognize these developments the class paradigm has made two errors that have seriously damaged the capacity of sociology to generate public debate about social and economic inequality. First, it has offered continued credence and legitimacy to class theory as a vehicle for the analysis of twentieth-

century developments. In an important but not absolutely critical sense, class died somewhere between the beginning of the twentieth century and the end of the Great Depression. It died with the absorption of class struggle into the democratic arena, with the emergence of the fascist and socialist states that so dominated civil society, with the institutionalization of corporatist deals that linked government, capital and labour into common projects, with the mitigation of the effects of the market through such institutions as citizenship and welfare, and with the domination of the planet by superpower politics. The development was not entirely critical because the effects of what we can call “organized class”, the successor to true economic class, could still be analysed sociologically. As a successor to class theory, class analysis is indeed therefore an appropriate and impressive intellectual development in which sociology can take just pride. However, we are now contemplating a second and more critical error. Sociology is continuing to offer class analysis as its main vehicle for the discovery of social inequality and struggle in the trans-millennial period. It is failing to recognize that oppression, exploitation, and conflict are being socially constructed around transcendent conceptions of individual human rights and global values that identify and empower struggles around such diverse focuses as postcolonial racism, sexual preferences, gender discrimination, environmental degradation, citizen participation, religious commitments and ethnic self-determination. These issues have little or nothing to do with class. In the contemporary period of history, the class paradigm is intellectually and morally bankrupt. . . .

Figure 1 is a kind of master diagram that summarizes the substantive argument of this chapter. It shows that the historical transformation from economic-class society to organized-class society to status-conventional society is not merely an issue of stratification but one of wide-ranging societal transforma-

| Axial principle | System level | | | | | | | |
|---------------------|---------------|---------------------|---------------------------------------|-------------------------|--------------------|-----------------------|-----------------------------------|------------|
| | World | Society | Politics | Economy | Community | Domesticity | Gender | Individual |
| Economic class | Colonialism | Laissez-faire state | Plutocratic and revolutionary parties | Owner capitalism | Property order | Reproduction site | Patriarchy | Worker |
| ↓ | | | | | | | | |
| Organized class | Imperialism | Corporatism | Mass parties | Fordism | Occupational order | Consumer sphere | Reorganized patriarchy (virarchy) | Citizen |
| ↓ | | | | | | | | |
| Conventional status | Globalization | Nation | Niche parties | Flexible specialization | Value order | Significant lifeworld | Hyperdifferentiation | Human |

Figure 1. The Registration of Stratification Orders Through System Levels

tion. The argument offered here is therefore part of the general theoretical effort that focuses on post-industrialization (Bell 1976), detraditionalization (Beck 1992; Giddens 1991), postmodernization (Crook et al. 1992; Harvey 1989; Lyotard 1984), societal disorganization (Lash and Urry 1987; 1994; Offe 1985), and globalization (Featherstone 1990; Robertson 1992; Waters 1995).

The emerging stratificational picture is represented in such work as that of Kornblum (1974) on Chicago steelworkers that reveals anything but structural and cultural homogeneity. He shows that dense social networks of primary groups cross-cut old class boundaries and establish salient non-class divisions along regional-residential, ethnic, racial and status lines. These are the main focuses for identity formation and local politics. A similar picture is emerging from British studies of communities (e.g. Pahl and Wallace 1985; Williams 1975), political behaviour (e.g. Rose and McAllister 1985; Dunleavy and Husbands 1985), and consumption (Saunders 1990; Featherstone 1991). Here, gender divisions, market fragmentation, housing and consumption-sector cleavages, and state dependency are the main influences on political identities and voting patterns.

Equally Phizacklea’s study (1990) of the fashion industry in Britain reveals deep divi-

sions created by combinations of economic inequality, race, gender and locality. She argues that a dual labour market is generated in the garment industry by an intersection between organizational factors (relations between large corporations and small subcontractors), racial and ethnic divisions and a gendered-labour pattern exerted by domestic norms and relations. Recession and cheap imports have pushed “ethnic entrepreneurs”, mainly of Asian extraction, into family businesses that rely on low-paid female labour. These women workers are vulnerable to exploitative conditions by dint of gendered social norms, i.e. notions of “women’s work”, modes of migration that often subject female dependants to indebtedness, and social discrimination that restricts their employment options. Such studies of multiple segmentation at the local level are becoming increasingly common.

This multiple socioeconomic and socio-cultural fragmentation is very distant from the kind of class differentiation identified by Lockwood (1958) or Goldthorpe et al. (1969) a generation earlier. They proposed differentiated sub-types of working-class and lower-middle-class orientation that created locally concentrated and culturally homogeneous class communities. By contrast contemporary studies emphasize cultural diversity, especially among younger people.

According to Willis (1990), this is the consequence not only of the decline in class-communal ties, but also of a drift away from the institutions that have engendered class identities in the past including community associations, trade unions, and schools. The mass unemployment of the 1980s that placed many young people beyond the class-formational impact of work situation and increasingly under the individualizing effects of commercialized consumer culture also contributed to the process.

We can now theorize this transformation more formally. The four propositions that class theory makes are: economism; groupness; behavioural linkage; and transformative capacity. A status-conventional theory would offer the following parallel propositions:

- The proposition of *culturalism*. Status-conventional stratification is primarily a cultural phenomenon. It is based on subscription to lifestyles that form around consumption patterns, information flows, cognitive agreements, aesthetic preferences and value commitments. Material and power phenomena are reducible to these symbolically manifested lifestyle and value phenomena.
- The proposition of *fragmentation*. Conventional statuses, like classes, are real phenomena. However they consist of a virtually infinite overlap of associations and identifications that are shifting and unstable. Status-conventional society is a fluid matrix of fragile formations that cycle and multiply within a globalized field.
- The proposition of *autonomization*. The subjective orientation and behaviour of any individual or aggregate of individuals are very difficult to predict by virtue of stratificational location. There is no central cleavage or single dimension along which preferences can be ordered. Such attributes as political preference, access to educational oppor-

tunity, patterns of marriage and income are self-referential rather than externally constrained.

- The proposition of *resignification* based on subjective interests. The stratification process is continuously fluid. Its openness allows a constant respecification and invention of preferences and symbolic dimensions that provide for continuous regeneration. The source of novelty is a process of restless subjective choice that seeks to gratify churning and unrepressed emotions that include anxiety and aggression as well as desire.

The proposition of culturalism specifies that symbolic dimensions will compete with each other in the field of social structure. This will produce the phenomenon of multiple status cleavages. The stratificational categories of status-conventional society constitute a complex mosaic of taste subcultures, “new associations”, civic initiatives, ethnic and religious revolutionary groups, generational cohorts, community action groups, new social movements, gangs, alternative lifestyle colonies, alternative production organizations, educational alumni, racial brotherhoods, gender sisterhoods, tax rebels, fundamentalist and revivalist religious movements, internet discussion groups, purchasing co-ops, professional associations, and so on. Many are ephemeral, some are continuous and stable.

A key feature of these multiple status cleavages is that because they are specialized and intersecting, membership in any one does not necessarily contradict membership in any other. From the subjective point of view the proposition of fragmentation ensures that individuals apprehend the stratification system as a status bazaar. Individuals can operate simultaneously as members of several status groups and have the potential to be members of any others. Their identities are reflexively self-composed as they move between status adherences. However, the fact of a status market does not imply an absolute voluntarism,

and indeed the freedoms in most cases are relevant to exit from status groups rather than entry. Closure processes remain effective in status-conventional society.

The proposition of autonomization nevertheless allows individuals to be profligate in their behaviour. They will tend to spend their resources of time, energy, money, influence, and power in the pursuit of symbolic attachments that tend to advance the interests, identities, values and commitments to which they subscribe and aspire. The very act of doing this will, by the proposition of resignification, tend to redefine and reorder the symbolic dimensions that reference the system. Indeed, a particular effect is the redefinition of some traditional status-membership dimensions, especially education, religion, and ethnicity, into a more ephemeral and conventional regime. So education becomes a marketplace for credentials, religion becomes a vehicle for handling this week's anxieties as one is born again and again and again, and ethnicity is something one rediscovers through community action and involvement.

We must stress that we are not arguing for a decline in inequality and conflict but for a decline in *class* inequality and conflict. One last illustration can confirm the point. The available research shows that, even allowing for recent reversals, during the twentieth century household wealth has become more widely distributed. This is an indicator of the decomposition of class, as normally defined in relation to property. However, at the present time there is also an increasing inequality of income (*The Economist* 5 November 1994). This is not evidence of increasing class inequality but it is evidence of increasing inequality of sumptuary capacity and this links directly with the kind of status-conventional stratification that is theorized here. The income-poor so-called "underclass" is not class defined but is rather status defined by the symbolizations attached to postcolonial migration, race, ethnicity, gender, age and pattern of family support. Exclu-

sionary closure based on these status attributes consigns people to an "underclass". The stigmatization that attaches to the "underclass" is a function not of its members' exploitation but of their incapacity to consume. An earlier generation of social scientists wrote of poverty as a culture (Lewis 1961; 1966; Valentine 1968). Perhaps it is time to do so once again.

All of this means that complexity is likely to increase. This calls for a theoretical stance that rejects any notion that there is a single conceptual or theoretical crowbar or even a magic word that can open the treasure cave of stratification, inequality and conflict. Contemporary sociology needs to be sensitive to and appreciative of diversity. It must be humble in the face of complexity and ready to accept contingency in relation to social attitudes and behaviours. It must not force a lively, engaging and ethically fraught reality into an inherited and stultifying conceptual straitjacket. If it continues on its procrustean path it will lose its audience.

REFERENCES

- Beck, U (1992) *Risk Society*. London: Sage.
 Bell, D (1976) *The Coming of Post-Industrial Society*. New York: Basic/Harper Torchbook (first published 1973).
 Crook, S, J Pakulski and M Waters (1992) *Post-modernization*. London: Sage.
 Dunleavy, P and C Husbands (1985) *British Democracy at the Crossroads*. London: Allen & Unwin.
 Engels, F (1892) *The Condition of the Working Class in England in 1844*. London: Hamden.
 Featherstone, M (1990) *Global Culture*. London: Sage.
 Featherstone, M (ed) (1991) *Consumer Culture and Postmodernism*. London: Sage.
 Giddens, A (1991) *Modernity and Self-Identity*. Cambridge: Polity.
 Goldthorpe, J, D Lockwood, F Bechhofer and J Platt (1969) *The Affluent Worker in the Class Structure*. Cambridge: CUP.
 Harvey, D (1989) *The Condition of Postmodernity*. Oxford: Blackwell.
 Kornblum, W (1974) *Blue Collar Community*. Chicago: Chicago UP.

- Lash, S and J Urry (1987) *The End of Organised Capitalism*. Cambridge: Polity.
- Lash, S and J Urry (1994) *Economies of Signs and Space*. London: Sage.
- Lewis, O (1961) *The Children of Sanchez*. New York: Random House.
- Lewis, O (1966) *La Vida: A Puerto Rican Family in the Culture of Poverty*. New York: Random House.
- Lockwood, D (1958) *The Blackcoated Worker*. London: Allen & Unwin.
- Lyotard, J (1984) *The Postmodern Condition*. Manchester: Manchester UP.
- Offe, C (1985) "New Social Movements: Challenging the Boundaries of Institutional Politics." *Social Research* 52(4): 817–68.
- Ossowski, S (1963) *Class Structure in the Social Consciousness*. London: Routledge (first published 1958).
- Pahl, R and C Wallace (1985) "Household Work Strategies in Economic Recession" in N Redclift and E Mingione (eds) *Beyond Employment*. Oxford: Blackwell.
- Phizacklea, A (1990) *Unpacking the Fashion Industry: Gender, Racism and Class in Production*. London: Routledge.
- Robertson, R (1992) *Globalization*. London: Sage.
- Rose, R and I McAllister (1985) *Voters Begin to Choose*. London: Sage.
- Saunders, P (1990) *A Nation of Home Owners*. London: Unwin Hyman.
- Thompson, E (1980) *The Making of the English Working Class* (2nd edn). Harmondsworth: Penguin (first edn 1968).
- Valentine, C (1968) *Culture and Poverty*. Chicago: Chicago UP.
- Waters, M (1995) *Globalization*. London: Routledge.
- Williams, R (1975) *The Country and the City*. London: Paladin.
- Willis, P (1990) *Common Culture*. Milton Keynes: Open UP.

108 ■ Ulrich Beck and Christoph Lau

Theoretical and Empirical Explorations in the 'Meta-Change' of Modern Society

All around the world, society is undergoing radical change—radical in the sense that it poses a challenge to Enlightenment-based modernity and opens up a space in which people choose new and unexpected forms of the social and the political. Sociological debates since the 1990s have sought to grasp and conceptualize this reconfiguration. Some authors, who lay great stress on the openness of the human project amid new contingencies, complexities and uncertainties, operate with the term 'post-modernity' (Bauman, Lyotard, Harvey, Haraway). However, we reject

the idea that so far this is a move from the modern to the postmodern. On theoretical as well as on empirical grounds our conclusion is that *all* Western societies are still 'modern' societies: there has been no movement beyond the realm of the modern to its opposite, because there has been no clear break with the basic principles of modernity but rather a transformation of basic institutions of modernity (for example, the nation-state and the nuclear family). We would suggest, therefore, that what we are witnessing is a *second* modernity.

Ulrich Beck and Christoph Lau. "Second Modernity as a Research Agenda: Theoretical and Empirical Explorations in the 'Meta-Change' of Modern Society," *British Journal of Sociology* 56 (2005), pp. 525–531, 556–557; Ulrich Beck, "Die uneindeutige Sozialstruktur," in *Individualisierung und Integration. Neue Konfliktlinien und neuer Integrationsmodus?*, edited by Ulrich Beck and Peter Sopp, pp. 183–197. Copyright © 1997 by Leske & Budrich.

We use the term 'first modernity' to describe the modernity based on nation-state societies; where social relations, networks and communities are essentially understood in a territorial sense. The collective patterns of life, progress and controllability, full employment and exploitation of nature that were typical of this first modernity have now been undermined by certain interlinked processes: globalization, individualization, the gender revolution, underemployment and global risks (such as the ecological crisis, the crash of global financial markets and the threat of transnational terrorist attacks). The real theoretical and political challenge of second modernity is the fact that society must respond to all these challenges *simultaneously*.

When these processes are considered more carefully, the thing they all have in common emerges more clearly: they are all unforeseen consequences, not of the crisis but of the *victory* of the first, simple, linear, industrial modernization based on the nation-state (the focus of classical sociology). This is what we mean when we speak of 'reflexive modernization': radicalized modernization consumes the foundations of first modernity and transforms its institutions and its frame of reference, often in a way that is neither desired nor anticipated. Or, in terms of systems theory: the unforeseen consequences of functional differentiation can no longer be controlled by further functional differentiation. In fact, the very idea of controllability, certainty or security—so fundamental to first modernity—collapses.

Basic Theoretical Assumptions

These basic assumptions include, first, the claim that there is a break between first and second, or reflexive, modernity. The term 'first modernity', far from constituting a denial of the different phases and developmental thrusts of modernization since the early modern era, instead encapsulates the claim that the transition to second modernity marks the end point of a specific logic of development, a logic

which, from the middle of the nineteenth century onwards, became more and more established in Europe and North America and culminated in the nation-state order of industrial society in the 1960s. This 'high modernity' of industrial society was—as we have mentioned—characterized by a configuration of institutions that mutually confirmed and supported one another, such as the nation-state, the Fordist company, the nuclear family, the system of industrial relations, the welfare state and unquestioned science. The foundation of this institutional structure was a logic of order and action that is only now plainly coming into view, as it comes to an end. This logic drew a strict boundary between categories of human beings, things and activities and made distinctions between modes of activity and ways of life, something which made it possible to assign areas of authority and responsibility unambiguously. This logic of non-ambiguity and strict differentiation is increasingly coming up against its own limitations at the present time. It is increasingly resistant to rational justification and, in some spheres, is completely ineffectual.

The logic of institutional action that held sway during first modernity worked according to the 'either/or' principle—either us or them, the organization or the market, family or not family, work or leisure, facts or values, war or peace. In the circumstances of reflexive modernity, by contrast, the either/or principle seems increasingly to be eclipsed or indeed replaced by the 'both/and' principle. Instead of there being either knowledge *or* not knowledge, nature *or* society, the organization *or* the market, there is instead both knowledge *and* not knowledge, nature *and* society, organization *and* the market, war *and* peace. As the boundaries and distinctions between categories become blurred, the institutions of first modernity that are based on them and depend on them for their existence, begin to encounter problems with regard to decision making. How are decisions to be made if, for example, it is no longer clear whether climate

change is a man-made or a natural phenomenon? How do authorities deal with migrants who belong to several societies and cultures at the same time? Where are the boundaries of 'patchwork families' to be drawn? What is apparent here is that as categorical boundaries become less clear and their rationalization more difficult, so it becomes increasingly necessary to decide on one or several new ways of drawing boundaries. In doing so, however, it is no longer possible to fall back on the tried and tested, usually scientific resources of rationalization, as these themselves have become ambiguous and uncertain. The pluralism of the newly discovered solutions is clearly a result of these kinds of pragmatic negotiation processes, in the course of which complex, plural demarcations come to take the place of unambiguous dualities, standard forms and distinctions.

How do these structural changes come about? The second basic assumption in the theory of reflexive modernization is that it is the process of modernization itself that calls the institutional order of first modernity into question, steals its clarity and forces it to make decisions. In other words, the crises of second modernity are triggered by processes of enforced modernization, which then take hold of modernization's own foundations. The 'protective zones' generated by first modernity to shield it against the dynamic of modernization are currently losing their unquestioned taken-for-grantedness. They are experienced as being contingent and open to influence, and are increasingly being forced to provide rationales for their existence. This also applies to the very legitimated sources of modernity themselves—and to science in particular. The latter may be producing more and more knowledge, but it is also producing an ever greater lack of knowledge, uncertainty and insufficient clarity. The rationalization process ultimately encompasses the model of hegemonic scientific rationality as well, which in many spheres is only of limited use for the task of making solid, unambiguous decisions. . . .

Programme of Empirical Research

The meta-change of modern societies, which has only been roughly outlined here, occurs in different ways in nearly every social sphere. The theory of reflexive modernization, translated into a programme of empirical research, makes it possible to compare different phenomena of the structural transcending of boundaries and to investigate whether they reveal any similar patterns or trends in three fields of inquiry: a) knowledge, science and technology, b) individuals, life situations, lifestyles and work, and c) economic politics and history. The results of these different endeavours have revealed remarkably consistent patterns of meta-change across very heterogeneous spheres of action and systems. For example, in the sphere of science and technology it has become clear that the elimination of boundaries from institutionalized nature-society distinctions and the elimination of the difference between scientific and non-scientific knowledge is leading to situations marked by crisis. Science is itself developing into a source of uncertainty, lack of knowledge and categorical ambiguity (Bösch and Wehling 2004). This means that those institutions that are still under obligation to make decisions have necessarily to make them in a pragmatic way, using political procedures and normative criteria (Bösch and Schulz-Schaeffer 2003).

The second field of inquiry (individuals, groups, work) has brought to light the different manifestations of a social structure in the process of becoming more fluid and individualized (Beck and Beck-Gernsheim 2001). Under conditions of reflexive modernization, research subjects are no longer able to take pre-given standard biographies, or the model of lifelong normal work, or the pattern of the stable nuclear family as their point of orientation. At the same time, however, rather than the various processes of increasing flexibility and boundary transcendence leading to an anomic breakdown of groups and communities, individuals themselves are finding ways

of linking conventional and new patterns of family ties and of locating themselves in communities of choice and working relationships. Here, too, it is a matter of new forms of boundary management, in which people set boundaries in a flexible and pragmatic way with regard to the family, friendship networks, working hours and individualization within couple relationships; this is a task they have to attend to over and over again (Keupp, Ahbe and Gmür 2002; Bonß, Kesselring and Vogl 2005; Wimbauer 2003).

The third empirical field (economics, politics and history) is concerned with processes of boundary transcendence related to corporations and nation-states and their consequences. Here, too, it has been found that the boundaries of transnational political regimes, for example, which can be understood as a political reaction to processes of denationalization, are precarious and contingent. Different criteria of inclusion compete with one another in this and have to be resolved (Beck and Grande 2004). However, globalization and denationalization not only lead to nation-state politics playing a reduced role, they also lead to a change in the structure of conflicts within nation-states (Grande and Kriesi 2004). Other research projects refer to the transnationalization of historical memory (Levy and Sznajder 2002; Beck, Levy and Sznajder 2004) and to the emergence of a 'world risk society' (Beck 1999). Last but not least, the transcendence of boundaries around national spaces makes it clear that society in second modernity can no longer be conceptualized using the concepts and categories of the nation-state, but that a new perspective of 'methodological cosmopolitanism' must be developed for this task (Beck 2004).

It is not only the nation-state but also the Fordist company that is affected by changes in its standard form. Processes of 'internal marketization' in particular constitute an important factor here, whereby the insulation of the market economy from the production economy is increasingly being done away with

(Kratzer 2003). This has consequences not only with regard to individuals' ever more arduous attempts at separating work from life, but also for the system of industrial relations, which is under threat due to the declining role of the trade unions (Heidling et al. 2004). A 'post-Fordist transitional regime' of economics and work is becoming apparent in which four features come together: (1) the limited feasibility of establishing the 'boundless company', the 'virtual workforce', 'flexible people'; (2) abrupt change in models of rationality; (3) institutionalized links between globalization, individualization and informatization; and (4) a change in legitimation towards cultural neoliberalism (Kratzer et al. 2004).

The cross-section these empirical studies draw through Western societies in transition to second modernity brings a surprisingly uniform picture to light: the institutions that are initially forced to cling to the old order are finding it hard to deal with the new fluid and hybrid forms, pluralizations and ambiguities that now exist. They encounter severe difficulties when it comes to taking action and making decisions, and they increasingly face problems in terms of attributing responsibility—whether it be in world politics or in couple relationships. . . .

Poverty and Wealth in A 'Self-Driven Culture'

Consider now the implications of these changes for inequality, poverty, and wealth.

Where individualization has been carried through, there arises what we would call an 'own life culture' or 'self-culture'. This should be understood in a twofold sense: it involves both recognition of the self (indeterminacy of the self and of the ensuing conflicts, crises and developmental opportunities) and a binding or bonding of self-oriented individuals to, with and against one another. 'Self-culture' thus denotes what was at first negatively addressed with the concept of a *post*-traditional lifeworld: that is, the compulsion and the

pleasure of leading an insecure life of one's own and co-ordinating it with the distinctive lives of other people.

The concept of a 'self-culture', understood as the adventurous search of the many for a 'life of their own', should first be distinguished from the concepts of *proletarian* and *bourgeois* culture that marked the features and the conflicts of the first (industrial) modernity. The latter's characteristic hopes and quirks, its habits in life, work, love and consumption, its styles of politics, its self-destructive rituals for sleeping, living, bathing, learning, relaxing and so on: all these are visible in the literature of every field, from the sciences through the novel to music. The self-culture develops to the extent that both proletarian culture and bourgeois culture fade and disappear. What then emerges is not, as sociologists have often suspected, a uniform middle-class culture and society (Schelsky 1965), but precisely a self-culture that is unpredictable both for oneself and for others, a cross between civil society, consumer society, therapy society and risk society.

One crucial difference from proletarian and bourgeois culture is that here it is no longer class categories but the very cultural and political dynamic of 'one's own life' which puts its stamp on society. The lines of conflict are more diffuse but no less profound. New imaginaries of morality and responsibility take shape and develop; poverty, marriage, youth and political commitment assume new countenances. And along with the basic distinction between fixed corporations and elective communities in which a reference to individuals is the norm, a number of other features may serve to define what is meant by a self-culture.

First there are external, *demographic* criteria: the growing number of single-person households (where individuals live the contradiction of being at once young and old, single, widowed and divorced etc.) and the correspondingly high value placed on separateness in every lifestyle. This is expressed in the basic need (developed and established historically)

for 'space' and 'time of one's own', and in the numerous consequences which this has for such things as the architectonics and infrastructure of late-modern cities. Another social-demographic feature of the self-culture is the high divorce figures and the related plethora, both open and concealed, of pre-marital, non-marital, extra-marital and post-marital lifestyles which we have grown used to lumping together under the veil of 'pluralization'.

In addition to these and other external characteristics, three further ones are of major significance for the self-culture. First, *the staging of the self in processes of aesthetic lifestyle creation*. The goal of making one's own life a work of art has become a guiding idea for the generation of 18 to 35 year olds; research findings about global (or 'glocal') culture are here materially relevant to the sociology of life histories.

Second, *an internalized, practising consciousness of freedom*. Only with difficulty, if at all, can this be circumscribed and directed from above into certain goals and forms of commitment (party, trade union or church membership etc.); this presupposes certain reforms—basic rights, citizenship, the rule of law—as well as educational facilities and much else besides, so that the cultural consciousness of freedom can develop in all its contradictions and be put to the test of practice.

Third, *self-organization geared to action and not just to the ballot box*. This differs from mere *participation*, which assumes a hierarchical division of jurisdiction and authority and sorts out who should be involved in what. Self-organization demands what is still denied in participation—namely, the right of citizens to take charge of matters they deem important.

We must be careful not to equate this with *emancipation* and all the other fine things proclaimed by theorists of democracy. For it *may* quite easily appear together with xenophobia, violence and all manner of panic movements in which people turn nasty and abandon not only their duty as 'good citizens' but also their role as election fodder for political parties.

Whereas, in the model of parliamentary democracy, the political activity of citizens is largely conceived as empowering someone else to represent them, self-organization is centred on the citizens' own activity. It is a question not just of occasionally voting people in and out, but of complaining, campaigning and acting about all things possible and impossible.

In this way a self-authorizing, self-referential sphere distinct from those of politics and economics, a kind of third sector of self-culture, develops through the centrifugal dynamic of 'a life of one's own'. Its autonomous logic of self-organization sets it apart from the money economy as much as from ballot-box democracy.

Does talk of a 'self-culture' perhaps correspond to how the winners see things, while the silent losers go downhill with the violence born of despair? Are today's winners anyway not tomorrow's losers? And is it not a fear of plunging downward which makes them too tremble for the future? In the 1970s and 1980s it was doubtless possible to speak of individualization based on affluence, but since the early 1990s the starting-point has rather been an individualization based on the precarious conditions of life in a *capitalism without work*. But the idea of a 'self-culture' anyway signifies not the overcoming but the *sharpening* of social inequality.

Self-culture means detraditionalization, release from pre-given certainties and supports. Your life becomes in principle a *risky venture*. A normal life story becomes a (seemingly) elective life, a *risk* biography, in the sense that everything (or nearly everything) is a matter for decision. And yet, faced with the opaque and contradictory character of modern society, the self-focused individual is hardly in a position to take the unavoidable decisions in a rational and responsible manner, that is, with reference to the possible consequences.

REFERENCES

Beck, Ulrich 1999 *World Risk Society*, Cambridge: Polity Press.

- Beck, Ulrich 2004 *Der kosmopolitische Blick oder: Krieg ist Frieden*, Frankfurt/M.: Suhrkamp/*The Cosmopolitan Vision*, Cambridge: Polity Press 2005.
- Beck, Ulrich and Beck-Gernsheim, Elisabeth 2001 *Individualization*, London: Sage.
- Beck, Ulrich and Grande, Edgar 2004 *Das kosmopolitische Europa. Gesellschaft und Politik in der Zweiten Moderne*, Frankfurt/M.: Suhrkamp.
- Beck, Ulrich, Levy, Daniel and Sznajder, Natan 2004 'Erinnerung und Vergebung in der Zweiten Moderne' in Ulrich Beck and Christoph Lau *Entgrenzung und Entscheidung. Was ist neu an der Theorie reflexiver Modernisierung*, Frankfurt/M.: Suhrkamp.
- Bonß, Wolfgang, Kesselring, Sven and Vogl, Gerlinde 2005 *Mobilitätsspioniere* (forthcoming).
- Böschen, Stefan and Schulz-Schaeffer, Ingo (eds) 2003 *Wissenschaft in der Wissensgesellschaft*, Wiesbaden: Westdeutscher Verlag.
- Böschen, Stefan and Wehling, Peter 2004 *Wissenschaft zwischen Folgenverantwortung und Nicht-Wissen*, Wiesbaden: VS Verlag.
- Grande, Edgar and Kriesi, Hanspeter 2004 'Nationaler politischer Wandel in entgrenzten Räumen' in Ulrich Beck and Christoph Lau (eds) *Entgrenzung und Entscheidung. Was ist neu an der Theorie reflexiver Modernisierung*, Frankfurt/M.: Suhrkamp.
- Heidling, Eckhard, Deiß, Manfred, Meil, Pamela and Schmierl, Klaus 2004 'Restrukturierung nationaler Interessenvertretung' in Ulrich Beck and Christoph Lau (eds) *Entgrenzung und Entscheidung*, Frankfurt/M.: Suhrkamp.
- Keupp, Reiner, Ahbe, Thomas and Gmür, Wolfgang 2002 *Identitätskonstruktionen*, Reinbek: Rowohlt.
- Kratzer, Nick, Boes, Andreas, Döhl, Volker, Marrs, Kira and Sauer, Dieter 2004 'Entgrenzung von Unternehmen und Arbeit—Grenzen der Entgrenzung' in Ulrich Beck and Christoph Lau (eds) *Entgrenzung und Entscheidung—Was ist neu an der Theorie reflexiver Modernisierung?*, Frankfurt/Main: Suhrkamp.
- Kratzer, Nick 2003 *Arbeitskraft in Entgrenzung*, Berlin: edition sigma.
- Levy, Daniel and Sznajder, Natan 2002 *Erinnerung im globalen Zeitalter. Der Holocaust*, Frankfurt/M.: Suhrkamp.
- Schelsky, H. 1965 *Auf der Suche nach Wirklichkeit. Gesammelte Aufsätze*, Düsseldorf: E. Duederichs Verlag.
- Wimbauer, Christine 2003 *Geld und Liebe. Zur symbolischen Bedeutung von Geld in Paarbeziehungen*, Frankfurt/M.: Campus.

Globalization and Inequality

109 ■ Joseph E. Stiglitz

Globalism's Discontents

Few subjects have polarized people throughout the world as much as globalization. Some see it as the way of the future, bringing unprecedented prosperity to everyone, everywhere. Others, symbolized by the Seattle protestors of December 1999, fault globalization as the source of untold problems, from the destruction of native cultures to increasing poverty and immiseration. In this [chapter], I want to sort out the different meanings of globalization. In many countries, globalization has brought huge benefits to a few with few benefits to the many. But in the case of a few countries, it has brought enormous benefit to the many. Why have there been these huge differences in experiences? The answer is that globalization has meant different things in different places.

The countries that have managed globalization on their own, such as those in East Asia, have, by and large, ensured that they reaped huge benefits and that those benefits were equitably shared; they were able substantially to control the terms on which they engaged with the global economy. By contrast, the countries that have, by and large, had globalization managed for them by the International Monetary Fund (IMF) and other international economic institutions

have not done so well. The problem is thus not with globalization but with how it has been managed.

The international financial institutions have pushed a particular ideology—market fundamentalism—that is both bad economics and bad politics; it is based on premises concerning how markets work that do not hold even for developed countries, much less for developing countries. The IMF has pushed these economics policies without a broader vision of society or the role of economics within society. And it has pushed these policies in ways that have undermined emerging democracies.

More generally, globalization itself has been governed in ways that are undemocratic and have been disadvantageous to developing countries, especially the poor within those countries. The Seattle protestors pointed to the absence of democracy and of transparency, the governance of the international economic institutions by and for special corporate and financial interests, and the absence of countervailing democratic checks to ensure that these informal and *public* institutions serve a general interest. In these complaints, there is more than a grain of truth.

Beneficial Globalization

Of the countries of the world, those in East Asia have grown the fastest and done most to reduce poverty. And they have done so, emphatically, via “globalization.” Their growth has been based on exports—by taking advantage of the global market for exports and by closing the technology gap. It was not just gaps in capital and other resources that separated the developed from the less-developed countries, but differences in knowledge. East Asian countries took advantage of the “globalization of knowledge” to reduce these disparities. But while some of the countries in the region grew by opening themselves up to multinational companies, others, such as Korea and Taiwan, grew by creating their own enterprises. Here is the key distinction: Each of the most successful globalizing countries determined its own pace of change; each made sure as it grew that the benefits were shared equitably; each rejected the basic tenets of the “Washington Consensus,” which argued for a minimalist role for government and rapid privatization and liberalization.

In East Asia, government took an active role in managing the economy. The steel industry that the Korean government created was among the most efficient in the world—performing far better than its private-sector rivals in the United States (which, though private, are constantly turning to the government for protection and for subsidies). Financial markets were highly regulated. My research shows that those regulations promoted growth. It was only when these countries stripped away the regulations, under pressure from the U.S. Treasury and the IMF, that they encountered problems.

During the 1960s, 1970s, and 1980s, the East Asian economies not only grew rapidly but were remarkably stable. Two of the countries most touched by the 1997–1998 economic crisis had had in the preceding three decades not a single year of negative growth; two had only one year—a better perfor-

mance than the United States or the other wealthy nations that make up the Organization for Economic Cooperation and Development (OECD). The single most important factor leading to the troubles that several of the East Asian countries encountered in the late 1990s—the East Asian crisis—was the rapid liberalization of financial and capital markets. In short, the countries of East Asia benefited from globalization because they made globalization work for them; it was when they succumbed to the pressures from the outside that they ran into problems that were beyond their own capacity to manage well.

Globalization can yield immense benefits. Elsewhere in the developing world, globalization of knowledge has brought improved health, with life spans increasing at a rapid pace. How can one put a price on these benefits of globalization? Globalization has brought still other benefits: Today there is the beginning of a globalized civil society that has begun to succeed with such reforms as the Mine Ban Treaty and debt forgiveness for the poorest highly indebted countries (the Jubilee movement). The globalization protest movement itself would not have been possible without globalization.

The Darker Side of Globalization

How then could a trend with the power to have so many benefits have produced such opposition? Simply because it has not only failed to live up to its potential but frequently has had very adverse effects. But this forces us to ask, why has it had such adverse effects? The answer can be seen by looking at each of the economic elements of globalization as pursued by the international financial institutions and especially by the IMF.

The most adverse effects have arisen from the liberalization of financial and capital markets—which has posed risks to developing countries without commensurate rewards. The liberalization has left them prey to hot

money pouring into the country, an influx that has fueled speculative real-estate booms; just as suddenly, as investor sentiment changes, the money is pulled out, leaving in its wake economic devastation. Early on, the IMF said that these countries were being rightly punished for pursuing bad economic policies. But as the crisis spread from country to country, even those that the IMF had given high marks found themselves ravaged.

The IMF often speaks about the importance of the discipline provided by capital markets. In doing so, it exhibits a certain paternalism, a new form of the old colonial mentality: “We in the establishment, we in the North who run our capital markets, know best. Do what we tell you to do, and you will prosper.” The arrogance is offensive, but the objection is more than just to style. The position is highly undemocratic: There is an implied assumption that democracy by itself does not provide sufficient discipline. But if one is to have an external disciplinarian, one should choose a good disciplinarian who knows what is good for growth, who shares one’s values. One doesn’t want an arbitrary and capricious taskmaster who one moment praises you for your virtues and the next screams at you for being rotten to the core. But capital markets are just such a fickle taskmaster; even ardent advocates talk about their bouts of irrational exuberance followed by equally irrational pessimism.

Lessons of Crisis

Nowhere was the fickleness more evident than in the last global financial crisis. Historically, most of the disturbances in capital flows into and out of a country are not the result of factors inside the country. Major disturbances arise, rather, from influences outside the country. When Argentina suddenly faced high interest rates in 1998, it wasn’t because of what Argentina did but because of what happened in Russia. Argentina cannot be blamed for Russia’s crisis.

Small developing countries find it virtually impossible to withstand this volatility. I have described capital-market liberalization with a simple metaphor: Small countries are like small boats. Liberalizing capital markets is like setting them loose on a rough sea. Even if the boats are well captained, even if the boats are sound, they are likely to be hit broadside by a big wave and capsize. But the IMF pushed for the boats to set forth into the roughest parts of the sea before they were seaworthy, with untrained captains and crews, and without life vests. No wonder matters turned out so badly!

To see why it is important to choose a disciplinarian who shares one’s values, consider a world in which there were free mobility of skilled labor. Skilled labor would then provide discipline. Today, a country that does not treat capital well will find capital quickly withdrawing; in a world of free labor mobility, if a country did not treat skilled labor well, it too would withdraw. Workers would worry about the quality of their children’s education and their family’s health care, the quality of their environment and of their own wages and working conditions. They would say to the government: If you fail to provide these essentials, we will move elsewhere. That is a far cry from the kind of discipline that free-flowing capital provides.

The liberalization of capital markets has not brought growth: How can one build factories or create jobs with money that can come in and out of a country overnight? And it gets worse: Prudential behavior requires countries to set aside reserves equal to the amount of short-term lending; so if a firm in a poor country borrows \$100 million at, say, 20 percent interest rates short-term from a bank in the United States, the government must set aside a corresponding amount. The reserves are typically held in U.S. Treasury bills—a safe, liquid asset. In effect, the country is borrowing \$100 million from the United States and lending \$100 million to the United States. But when it borrows, it

pays a high interest rate, 20 percent; when it lends, it receives a low interest rate, around 4 percent. This may be great for the United States, but it can hardly help the growth of the poor country. There is also a high *opportunity* cost of the reserves; the money could have been much better spent on building rural roads or constructing schools or health clinics. But instead, the country is, in effect, forced to lend money to the United States.

Thailand illustrates the true ironies of such policies: There, the free market led to investments in empty office buildings, starving other sectors—such as education and transportation—of badly needed resources. Until the IMF and the U.S. Treasury came along, Thailand had restricted bank lending for speculative real estate. The Thais had seen the record: Such lending is an essential part of the boom-bust cycle that has characterized capitalism for 200 years. It wanted to be sure that the scarce capital went to create jobs. But the IMF nixed this intervention in the free market. If the free market said, “Build empty office buildings,” so be it! The market knew better than any government bureaucrat who mistakenly might have thought it wiser to build schools or factories.

The Costs of Volatility

Capital-market liberalization is inevitably accompanied by huge volatility, and this volatility impedes growth and increases poverty. It increases the risks of investing in the country, and thus investors demand a risk premium in the form of higher-than-normal profits. Not only is growth not enhanced but poverty is increased through several channels. The high volatility increases the likelihood of recessions—and the poor always bear the brunt of such downturns. Even in developed countries, safety nets are weak or nonexistent among the selfemployed and in the rural sector. But these are the dominant sectors in developing countries. Without adequate safety nets, the recessions

that follow from capital-market liberalization lead to impoverishment. In the name of imposing budget discipline and reassuring investors, the IMF invariably demands expenditure reductions, which almost inevitably result in cuts in outlays for safety nets that are already threadbare.

But matters are even worse—for under the doctrines of the “discipline of the capital markets,” if countries try to tax capital, capital flees. Thus, the IMF doctrines inevitably lead to an increase in tax burdens on the poor and the middle classes. Thus, while IMF bailouts enable the rich to take their money out of the country at more favorable terms (at the overvalued exchange rates), the burden of repaying the loans lies with the workers who remain behind.

The reason that I emphasize capital-market liberalization is that the case against it—and against the IMF’s stance in pushing it—is so compelling. It illustrates what can go wrong with globalization. Even economists like Jagdish Bhagwati, strong advocates of free trade, see the folly in liberalizing capital markets. Belatedly, so too has the IMF—at least in its official rhetoric, though less so in its policy stances—but too late for all those countries that have suffered so much from following the IMF’s prescriptions.

But while the case for trade liberalization—when properly done—is quite compelling, the way it has been pushed by the IMF has been far more problematic. The basic logic is simple: Trade liberalization is supposed to result in resources moving from inefficient protected sectors to more efficient export sectors. The problem is not only that job destruction comes before the job creation—so that unemployment and poverty result—but that the IMF’s “structural adjustment programs” (designed in ways that allegedly would reassure global investors) make job creation almost impossible. For these programs are often accompanied by high interest rates that are often justified by a single-minded focus on inflation. Sometimes that

concern is deserved; often, though, it is carried to an extreme. In the United States, we worry that small increases in the interest rate will discourage investment. The IMF has pushed for far higher interest rates in countries with a far less hospitable investment environment. The high interest rates mean that new jobs and enterprises are not created. What happens is that trade liberalization, rather than moving workers from low-productivity jobs to high-productivity ones, moves them from low-productivity jobs to unemployment. Rather than enhanced growth, the effect is increased poverty. To make matters even worse, the unfair trade-liberalization agenda forces poor countries to compete with highly subsidized American and European agriculture. . . .

An Unfair Trade Agenda

The trade-liberalization agenda has been set by the North, or more accurately, by special interests in the North. Consequently, a disproportionate part of the gains has accrued to the advanced industrial countries, and in some cases the less-developed countries have actually been worse off. After the last round of trade negotiations, the Uruguay Round that ended in 1994, the World Bank calculated the gains and losses to each of the regions of the world. The United States and Europe gained enormously. But sub-Saharan Africa, the poorest region of the world, lost by about 2 percent because of terms-of-trade effects: The trade negotiations opened their markets to manufactured goods produced by the industrialized countries but did not open up the markets of Europe and the United States to the agricultural goods in which poor countries often have a comparative advantage. Nor did the trade agreements eliminate the subsidies to agriculture that make it so hard for the developing countries to compete.

The U.S. negotiations with China over its membership in the WTO displayed a double

standard bordering on the surreal. The U.S. trade representative, the chief negotiator for the United States, began by insisting that China was a developed country. Under WTO rules, developing countries are allowed longer transition periods in which state subsidies and other departures from the WTO strictures are permitted. China certainly wishes it were a developed country, with Western-style per capita incomes. And since China has a lot of “capitas,” it’s possible to multiply a huge number of people by very small average incomes and conclude that the People’s Republic is a big economy. But China is not only a developing economy; it is a low-income developing country. Yet the United States insisted that China be treated like a developed country! China went along with the fiction; the negotiations dragged on so long that China got some extra time to adjust. But the true hypocrisy was shown when U.S. negotiators asked, in effect, for developing-country status for the United States to get extra time to shelter the American textile industry.

Trade negotiations in the service industries also illustrate the unlevel nature of the playing field. Which service industries did the United States say were *very* important? Financial services—industries in which Wall Street has a comparative advantage. Construction industries and maritime services were not on the agenda, because the developing countries would have a comparative advantage in these sectors.

Consider also intellectual-property rights, which are important if innovators are to have incentives to innovate (though many of the corporate advocates of intellectual property exaggerate its importance and fail to note that much of the most important research, as in basic science and mathematics, is not patentable). Intellectual-property rights, such as patents and trademarks, need to balance the interests of producers with those of users—not only users in developing countries, but researchers in developed countries.

If we underprice the profitability of innovation to the inventor, we deter invention. If we overprice its cost to the research community and the end user, we retard its diffusion and beneficial effects on living standards.

In the final stages of the Uruguay negotiations, both the White House Office of Science and Technology Policy and the Council of Economic Advisers worried that we had not got the balance right—that the agreement put producers' interests over users'. We worried that, with this imbalance, the rate of progress and innovation might actually be impeded. After all, knowledge is the most important input into research, and overly strong intellectual-property rights can, in effect, increase the price of this input. We were also concerned about the consequences of denying lifesaving medicines to the poor. This issue subsequently gained international attention in the context of the provision of AIDS medicines in South Africa. The international outrage forced the drug companies to back down—and it appears that, going forward, the most adverse consequences will be circumscribed. But it is worth noting that initially, even the Democratic U.S. administration supported the pharmaceutical companies.

What we were not fully aware of was another danger—what has come to be called “biopiracy,” which involves international drug companies patenting traditional medicines. Not only do they seek to make money from “resources” and knowledge that rightfully belong to the developing countries, but in doing so they squelch domestic firms who long provided these traditional medicines. While it is not clear whether these patents would hold up in court if they were effectively challenged, it is clear that the less-developed countries may not have the legal and financial resources required to mount such a challenge. The issue has become the source of enormous emotional, and potentially economic, concern throughout the developing world. This fall, while I was in

Ecuador visiting a village in the high Andes, the Indian mayor railed against how globalization had led to biopiracy. . . .

Trickle-Down Economics

We recognize today that there is a “social contract” that binds citizens together, and with their government. When government policies abrogate that social contract, citizens may not honor their “contracts” with each other, or with the government. Maintaining that social contract is particularly important, and difficult, in the midst of the social upheavals that so frequently accompany the development transformation. In the green eye-shaded calculations of the IMF macroeconomics there is, too often, no room for these concerns.

Part of the social contract entails “fairness,” that the poor share in the gains of society as it grows, and that the rich share in the pains of society in times of crisis. The Washington Consensus policies paid little attention to issues of distribution or “fairness.” If pressed, many of its proponents would argue that the best way to help the poor is to make the economy grow. They believe in trickle-down economics. *Eventually*, it is asserted, the benefits of that growth *trickle down* even to the poor. Trickle-down economics was never much more than just a belief, an article of faith. Pauperism seemed to grow in nineteenth-century England even though the country as a whole prospered. Growth in America in the 1980s provided the most recent dramatic example: while the economy grew, those at the bottom saw their real incomes decline. The Clinton administration had argued strongly against trickle-down economics; it believed that there had to be active programs to help the poor. And when I left the White House to go to the World Bank, I brought with me the same skepticism of trickle-down economics; if this had not worked in the United States, why would it work in developing countries?

While it is true that sustained reductions in poverty cannot be attained without robust economic growth, the converse is not true: growth need not benefit all. It is not true that “a rising tide lifts all boats.” Sometimes, a quickly rising tide, especially when accompanied by a storm, dashes weaker boats against the shore, smashing them to smithereens.

In spite of the obvious problems confronting trickle-down economics, it has a good intellectual pedigree. One Nobel Prize winner, Arthur Lewis, argued that inequality was good for development and economic growth, since the rich save more than the poor, and the key to growth was capital accumulation. Another Nobel Prize winner, Simon Kuznets, argued that while in the initial stages of development inequality increased, later on the trend was reversed.¹

The history of the past fifty years has, however, not supported these theories and hypotheses. East Asian countries—South Korea, China, Taiwan, Japan—showed that high savings did not require high inequality, that one could achieve rapid growth without a substantial increase in inequality. Because the governments did not believe that growth would automatically benefit the poor, and because they believed that greater *equality* would actually enhance growth, governments in the region took active steps to ensure that the rising tide of growth did lift most boats, that wage inequalities were kept in bounds, that some educational opportunity was extended to all. Their policies led to social and political stability, which in turn contributed to an economic environment in which businesses flourished. Tapping new reservoirs of talent provided the energy and human skills that contributed to the dynamism of the region.

Elsewhere, where governments adopted the Washington Consensus policies, the poor have benefited less from growth. In Latin America, growth has not been accompanied by a reduction in inequality, or even a reduction in poverty. In some cases poverty has actually increased, as evidenced by the urban

slums that dot the landscape. The IMF talks with pride about the progress Latin America has made in market reforms over the past decade (though somewhat more quietly after the collapse of the star student Argentina in 2001, and the recession and stagnation that have afflicted many of the “reform” countries during the past five years), but has said less about the numbers in poverty.

Clearly, growth alone does not always improve the lives of all a country’s people. Not surprisingly, the phrase “trickle-down” has disappeared from the policy debate. But, in a slightly mutated form, the idea is still alive. I call the new variant *trickle-down-plus*. It holds that growth is necessary and *almost* sufficient for reducing poverty—implying that the best strategy is simply to focus on growth, while mentioning issues like female education and health. But proponents of trickle-down-plus failed to implement policies that would effectively address either broader concerns of poverty or even specific issues such as the education of women. In practice, the advocates of trickle-down-plus continued with much the same policies as before, with much the same adverse effects. The overly stringent “adjustment policies” in country after country forced cutbacks in education and health: in Thailand, as a result, not only did female prostitution increase but expenditures on AIDS were cut way back; and what had been one of the world’s most successful programs in fighting AIDS had a major setback.

The irony was that one of the major proponents of trickle-down-plus was the U.S. Treasury under the Clinton administration. Within the administration, in domestic politics, there was a wide spectrum of views, from New Democrats, who wanted to see a more limited role for government, to Old Democrats, who looked for more government intervention. But the central view, reflected in the annual Economic Report of the President (prepared by the Council of Economic Advisers), argued strongly against trickle-down economics—or even trickle-down-plus. Here

was the U.S. Treasury pushing policies on other countries that, had they been advocated for the United States, would have been strongly contested *within the administration*, and almost surely defeated. The reason for this seeming inconsistency was simple: The IMF and the World Bank were part of Treasury's turf, an arena in which, with few exceptions, they were allowed to push their perspectives, just as other departments, within their domains, could push theirs. . . .

Global Social Justice

Today, in much of the developing world, globalization is being questioned. For instance, in Latin America, after a short burst of growth in the early 1990s, stagnation and recession have set in. The growth was not sustained—some might say, was not sustainable. Indeed, at this juncture, the growth record of the so-called post-reform era looks no better, and in some countries much worse, than in the widely criticized import-substitution period of the 1950s and 1960s when Latin countries tried to industrialize by discouraging imports. Indeed, reform critics point out that the burst of growth in the early 1990s was little more than a “catch-up” that did not even make up for the lost decade of the 1980s.

Throughout the region, people are asking: “Has reform failed or has globalization failed?” The distinction is perhaps artificial, for globalization was at the center of the reforms. Even in those countries that have managed to grow, such as Mexico, the benefits have accrued largely to the upper 30 percent and have been even more concentrated in the top 10 percent. Those at the bottom have gained little; many are even worse off. The reforms have exposed countries to greater risk, and the risks have been borne disproportionately by those least able to cope with them. Just as in many countries where the pacing and sequencing of reforms has resulted in job destruction outmatching job creation, so too has the exposure to risk outmatched the abil-

ity to create institutions for coping with risk, including effective safety nets.

In this bleak landscape, there are some positive signs. Those in the North have become more aware of the inequities of the global economic architecture. The agreement at Doha to hold a new round of trade negotiations—the “Development Round”—promises to rectify some of the imbalances of the past. There has been a marked change in the rhetoric of the international economic institutions—at least they talk about poverty. At the World Bank, there have been some real reforms; there has been some progress in translating the rhetoric into reality—in ensuring that the voices of the poor are heard and the concerns of the developing countries are listened to. But elsewhere, there is often a gap between the rhetoric and the reality. Serious reforms in governance, in who makes decisions and how they are made, are not on the table. If one of the problems at the IMF has been that the ideology, interests, and perspectives of the financial community in the advanced industrialized countries have been given disproportionate weight (in matters whose effects go well beyond finance), then the prospects for success in the current discussions of reform, in which the same parties continue to predominate, are bleak. They are more likely to result in slight changes in the shape of the table, not changes in who is *at* the table or what is on the agenda.

September 11 has resulted in a global alliance against terrorism. What we now need is not just an alliance *against* evil, but an alliance *for* something positive—a global alliance for reducing poverty and for creating a better environment, an alliance for creating a global society with more social justice.

NOTE

1. See W. A. Lewis, “Economic Development with Unlimited Supplies of Labor,” *Manchester School* 22 (1954), pp. 139–91, and S. Kuznets, “Economic Growth and Income Inequality,” *American Economic Review* 45(1) (1955), pp. 1–28.

110 ■ Glenn Firebaugh

The New Geography of Global Income Inequality

The new geography of inequality—not to be confused with the “new economic geography” that arose in economics in the 1990s (Fujita, Krugman, and Venables 1999)—refers to the new pattern of global income inequality caused by the recent phenomenon of declining inequality across nations accompanied by (in many places) rising inequality within nations. This phenomenon, which began in the last third of the twentieth century and continues today, results in a “new geography” because it represents the reversal of trends that trace back to the early stages of Western industrialization. Put in the perspective of an individual, the new geography of global income inequality means that national location—while still paramount—is declining in significance in the determination of one’s income.

Despite a recent surge of interest in global inequality, researchers have largely overlooked its changing contour. Studies of global income inequality over the last decades of the twentieth century have been preoccupied with the problem of global divergence, that is, the presumed problem of worsening income inequality for the world as a whole (for example, Wade 2001; Milanovic 2002). This preoccupation with global divergence is misguided, first, because global income inequality almost certainly declined over this period and, second, because the focus on the level of global income inequality has diverted attention from the changing nature of global income inequality in recent decades. Global income inequality is no

worse today than it was in the 1960s and 1970s, but global income inequality is nevertheless changing—it is gradually shifting from inequality across nations to inequality within nations. The rising importance of within-nation inequality and declining importance of between-nation inequality represents a historic change, since it involves the reversal of a trend that began with the uneven industrialization of the world that started more than two centuries ago.

This chapter contrasts the New Geography Hypothesis with the popular view that globalization—by which I mean the increased interconnectedness of localities, particularly the deepening of economic links between countries—has led to growing global income inequality:

Globalization → global inequality.

For short, I call this popular view the Trade Protest Model, because the protests against the World Trade Organization in Seattle and elsewhere were driven at least in part by the assumption that global trade is exacerbating global inequality. To place the New Geography Hypothesis in context, it is useful first to examine five myths that underlie the globalization → global inequality model.

Myths of the Trade Protest Model

Under the heading “Siege in Seattle,” the December 13, 1999, issue of the U.S. mag-

azine *Newsweek* gave this account of the protests surrounding the meeting of the World Trade Organization in Seattle, Washington:

Until last week, not so many Americans had even heard of the WTO. Fewer still could have identified it as the small, Geneva-based bureaucracy that the United States and 134 other nations set up five years ago to referee global commerce. To Bill Clinton, it is a mechanism that can allow America to do well and good at the same time. But to many of the 40,000 activists and union members who streamed into Seattle—a clean, scenic city that has grown rich on foreign trade—the WTO is something else again: a secretive tool of ruthless multinational corporations. They charge it with helping sneaker companies to exploit Asian workers, timber companies to clear-cut rain forests, shrimpers to kill sea turtles and a world of other offenses.

Media accounts grappled with the sheer diversity of the protesters, from leaders of U.S. labor to members of environmental groups to a leading Chinese dissident. The common thread seemed to be, as the *New York Times* (1999) put it, the view that the WTO is a “handmaiden of corporate interests whose rulings undermine health, labor and environmental protections around the world.” According to *The Economist* (1999), “The WTO has become a magnet for resistance to globalisation by both old-fashioned protectionists and newer critics of free trade.”

Some of the protest groups emphasized rising inequality as among the most noxious consequences of increasing trade globalization. For example, Ralph Nader’s Public Citizen group portrays the WTO as a tool of big business “which is harming the environment and increasing inequality” (*The Economist* 1999), and representatives of 1,448 non-governmental organizations protesting the

WTO signed a statement claiming that “globalisation has three serious consequences: the concentration of wealth in the hands of the multinationals and the rich; poverty for the majority of the world’s population; and unsustainable patterns of production and consumption that destroy the environment” (*New Scientist* 1999). Note that two of the three consequences—concentration of wealth and impoverishment of the majority—tie globalization directly to growth in inequality.

Global income inequality is the result of the interplay of multiple causes, of course, so serious analyses are unlikely to give an unqualified endorsement to the notion that globalization has automatically resulted in an explosion in global income inequality. Global inequality existed before the recent growth in world trade, and it would persist if nations suddenly stopped trading. Nonetheless, popular literature on globalization has tended to fuel the belief in a globalization-led explosion in global income inequality by making claims that purport to be grounded in the findings of serious scholarly analyses. Upon closer inspection, however, many of the claims fly in the face of available empirical evidence. This section examines the key myths that underlie the globalization → global inequality model.

Myth 1. The Myth of Exploding Global Income Inequality. A steady drumbeat of reports and articles claims that the world’s income is becoming more and more unequally distributed. Here is a sample:

- “Globalization has dramatically increased inequality between and within nations” (Jay Mazur, 2000, in *Foreign Affairs*).
- “The very nature of globalization has an inherent bias toward inequality. . . . One would have to be blind not to see that globalization also exacerbates the

disparity between a small class of winners and the rest of us” (Paul Martin, Canada’s prime minister, June 1998, quoted in Eggertson 1998).

- “Along with ecological risk, to which it is related, expanding inequality is the most serious problem facing world society” (Anthony Giddens, 1999).
- “Thus, overall, the ascent of informational, global capitalism is indeed characterized by simultaneous economic development and underdevelopment, social inclusion and social exclusion. . . . There is polarization in the distribution of wealth at the global level, differential evolution of intra-country income inequality, and substantial growth of poverty and misery in the world at large” (Manuel Castells, 1998, p. 82, emphasis omitted).

What I have found (Firebaugh 2003) is that global income inequality has not exploded but in fact leveled off and then declined in the last part of the twentieth century. Although income inequality rose somewhat in the average nation, income inequality declined across nations. Since between-nation inequality is the larger component of global income inequality, the decline in between-nation income inequality more than offset the rise in within-nation income inequality. As a result, global income inequality declined in the last years of the twentieth century. Sherlock Holmes was right: it *is* a capital mistake to theorize in advance of the facts (Doyle 1955, p. 507). With respect to global income inequality, much mischief has been done by theorizing about global income inequality on the basis of the views expressed above. Theorizing based on the widespread view of exploding global income inequality is theorizing based on facts that aren’t.

Myth 2. The Myth of Growing Income Inequality Across Nations, as Rich Nations Surge

Ahead and Poor Nations Fall Further Behind. The first myth—exploding global inequality—is based on a second myth, the myth that inequality is growing across nations. The second myth is as widespread as the first, and it has been fueled by widely circulated reports of international agencies:

- “Figures indicate that income inequality between countries has increased sharply over the past 40 years” (World Bank 2000b, *World Development Report 2000/2001*, p. 51).
- “The average income in the richest 20 countries is 37 times the average in the poorest 20—a gap that has doubled in the past 40 years” (International Monetary Fund 2000, p. 50).
- “Gaps in income between the poorest and richest people and countries have continued to widen. In 1960 the 20% of the world’s people in the richest countries had 30 times the income of the poorest 20%—in 1997, 74 times as much. . . . Gaps are widening both between and within countries” (United Nations Development Program 1999, *Human Development Report 1999*, p. 36).
- “It is an empirical fact that the income gap between poor and rich countries has increased in recent decades” (Ben-David, Nordström, and Winters 1999, World Trade Organization special study, p. 3).
- “In 1960, the Northern countries were 20 times richer than the Southern, in 1980 46 times. . . . [I]n this kind of race, the rich countries will always move faster than the rest” (Sachs 1992, *The Development Dictionary*, p. 3).

The myth of growing income inequality across nations is based in large part on a misinterpretation of the widely cited finding (for example, World Bank 2000b, p. 50) that income growth has tended to be slower in poor

nations than in rich nations. As we shall see, this positive cross-country association between income level and income growth rate conceals the critical fact that the poor nations that are falling badly behind contain no more than 10 percent of the world's population, whereas the poor nations that are catching up (largely in Asia) contain over 40 percent of the world's population. When nations are weighted by population size—as they must be if we want to use between-nation inequality to draw conclusions about global income inequality—we find that income inequality across nations peaked sometime around 1970 and has been declining since. This peaking of between-nation income inequality circa 1970 is particularly interesting in light of Manuel Castells's (1998, p. 336) well-known claim that a “new world” originated in the late 1960s to mid-1970s. Ironically, though, Castells characterizes the world born in this period as a world of sharply increasing global inequality, and many other globalization writers make the same error.

I have debunked myth number 2 first by replicating the United Nations/World Bank results and then by demonstrating how they have been misinterpreted (Firebaugh 2003). In addition to the weighting problem just mentioned, the claims of growing inequality often ignore nations in the middle of the income distribution, focusing instead on selected nations at the tails. When the entire income distribution is used, and when individuals are given equal weight, we find that—far from growing—income inequality across nations declined in the late twentieth century.

Myth 3. The Myth That Globalization Historically Has Caused Rising Inequality Across Nations. Contrary to this myth, the trend in between-nation inequality historically has not followed changes in the trend in world economic integration. First, although it is true that between-nation income inequality increased dramatically over the nineteenth and

early twentieth centuries, Peter Lindert and Jeffrey Williamson (2000) argue that the period of rising inequality across nations began *before* the period of true globalization started, so globalization apparently did not cause the upturn. Second, the sharp decline in globalization between World War I and World War II did not result in declining inequality across nations (to the contrary, between-nation income inequality shot up rapidly over the period). Finally, income inequality across nations has declined in recent decades, during a period when globalization has presumably reached new heights. (I say “presumably” because globalization is itself a contentious issue: see Guillén 2001. Nonetheless virtually all agree that the world has become more economically integrated over recent decades, even if the degree of globalization is often overstated, as Chase-Dunn, Kawano, and Brewer 2000, among others, have noted.) In short, the rise in global inequality predates the rise in globalization, global inequality has risen while globalization was declining, and currently global inequality is declining while globalization is rising. It is hard then to make the case historically that globalization is the cause of rising income inequality across nations (O'Rourke 2001).

Myth 4. The Myth of a Postindustrial World Economy. In reading the globalization literature it is easy to lose sight of the fact that, until recently, most of the world's people were engaged in agriculture. So the world's workforce is barely postagricultural, much less postindustrial. This book makes the point that the primary engine still driving the growth in world production is more manufacturing. A new information age might be on the way, but it is not here yet—at least it is not here for most of the world's people. It is important to look ahead, of course, and it is hard to argue against the view that the world will eventually be

postindustrial. The death of industrialization is nonetheless much exaggerated, as is the view that we are rapidly approaching an information-based global economy (Quah 1997). Estimates of the composition of global output, albeit rough approximations, rule out the claims of some globalization writers that we live in a new economic era quite unlike the era of the last generation. Industrialization was important in the nineteenth century, it was important in the twentieth century, and it remains important in most regions of the world in the twenty-first century. A preoccupation with postindustrialization in the face of the continuing diffusion of industrialization results in an incomplete and distorted story of global income inequality that deemphasizes the critical role of the continuing spread of industrialization to all regions of the world. Computers are important, but they are not all-important. In accounting for recent trends in global income inequality, the bigger story is industrial growth in Asia, not technological growth in the West.

Myth 5. The Myth of International Exchange as Inherently Exploitative. Globalization involves increased exchange over national boundaries. One might posit that increased exchange worsens global inequality under some historical conditions and reduces it under other conditions. Until those historical conditions are identified and understood, the effect of globalization on global income inequality at any point in time is an open question to be settled empirically.

But if international exchange is inherently exploitative, as some theories of world stratification insist, then rising exchange implies rising exploitation, and the Trade Protest Model is true virtually by definition. The Trade Protest Model then becomes:

Globalization → more exploitation of poor nations by rich nations → greater global inequality

Note that this elaboration of the Trade Protest Model reveals how high the theoretical stakes are with regard to empirical tests of the globalization → global income inequality model, since the failure of globalization to lead to rising global income inequality would undermine exploitation theories (for example, dependency theories) as well as undermining the Trade Protest Model.

I have presented evidence that increasing international exchange over recent decades has been accompanied by declining—not rising—income inequality across nations (Firebaugh 2003). Other studies also document declining between-nation income inequality over recent decades. The decline in between-nation inequality has significant theoretical implications. The assumption of inherent exploitation favoring rich nations in international exchange is the linchpin of some theoretical schools. But if international exchange were inherently exploitative, we would not expect to observe declining inequality across nations during a period of rising international trade. Yet the assumption persists, suggesting that in some theories the notion of inherent exploitation is so essential that it enjoys creedal status as a doctrine to be believed rather than as a hypothesis to be tested.

Causes of the Reversal: An Overview

I argue that the world's spreading industrialization and growing economic integration in the late twentieth century and the early twenty-first have reversed the historical pattern of uneven economic growth favoring richer nations. The conventional view, just elaborated, is that globalization has exacerbated global income inequality. The evidence challenges that view. In reality globalization has offsetting effects—by spurring industrialization in poor nations, globalization raises inequality within many nations and compresses inequality across nations. The net effect has been a reduction in global income

inequality in recent decades, since the reduction in between-nation income inequality has more than offset the growth in within-nation income inequality.

The new pattern of rising within-nation and falling between-nation income inequality has multiple causes. The most important cause is spreading industrialization—the diffusion of industrialization to the world's large poor nations. The diffusion of industrialization to poor regions compresses inequality across nations and boosts inequality within them. The effect of spreading industrialization on between-nation inequality is reinforced by the effect of the growing integration of national economies. Growing economic integration tends to dissolve institutional differences between nations. The convergence of institutional economic goals and policies compresses inequality across nations by (in some instances at least) removing impediments to growth in poor nations.

There are at least four other significant causes of the new geography of global inequality. The first is technological change that reduces the tyranny of space in general and more particularly reduces the effect of labor immobility across national boundaries. This technological change works to reduce inequality across nations. The second is a demographic windfall that has benefited some poor Asian nations in recent decades and promises to benefit other poor nations in the near future. This effect also operates to compress global income inequality, by reducing between-nation inequality. The third is the rise of the service sector, especially in richer nations. Growth in this sector has boosted income inequality within nations, and it is likely to do so in the future as well. The fourth is the collapse of communism, which also boosted within-nation inequality. This is a nonrecurring event, however, so its effect on within-nation inequality is limited to a specific point in history, the 1990s.

In short, the decline in between-nation income inequality that began in the late twen-

tieth century was caused by deepening industrialization of poor nations, by growing economic integration that dissolves institutional differences between nations, by technological change that reduces the effects of labor immobility across national boundaries, and by a demographic windfall that has benefited some poor nations and promises to benefit others in the future. The growth in within-nation income inequality was caused by the deepening industrialization of poor nations, by the growth of the service sector, and by the collapse of communism.

The Inequality Transition

The industrialization of richer nations in the nineteenth century and first half of the twentieth caused income inequality across nations to explode. As a result, global income inequality shifted from inequality within nations to inequality across nations. Now, however, poorer nations are industrializing faster than richer ones are, and between-nation inequality is declining while within-nation inequality appears to be rising. If this turnaround continues, future historians will refer to an inequality transition that accompanied world industrialization. That transition is from within-nation inequality to between-nation inequality back to within-nation inequality, with the late twentieth century as the period when the shift back to within-nation income inequality began.

Phase 1 of the Transition: From Within- to Between-Nation Inequality

Phase 1 of the inequality transition coincides with the period of Western industrialization that began in the late eighteenth century and ended in the second half of the twentieth century. The first phase of the inequality transition was characterized by unprecedented growth in income inequality across nations. As Lant Pritchett (1996, p. 40) puts it, “the overwhelming feature of modern economic

history is a massive divergence in per capita incomes between rich and poor countries.” The evidence is incontrovertible. First, it is clear that current levels of between-nation income inequality would not have been possible earlier in human history. Again quoting Pritchett (1997, pp. 9–10): “If there had been no divergence, then we could extrapolate backward from present income of the poorer countries to past income assuming they grew at least as fast as the United States. However, this would imply that many poor countries must have had incomes below \$100 in 1870 [in 1985 U.S. dollars]. Since this cannot be true, there must have been divergence.”

Second, Pritchett’s conclusion that “there must have been divergence” is supported by estimates of between-nation income inequality in the nineteenth century. Consider the recent estimates of Bourguignon and Morrisson (1999). Bourguignon and Morrisson use the Maddison (1995) data to estimate changes in the level of between-nation income inequality from 1820 to 1992. Because their objective is to estimate total world income inequality—not just between-nation income inequality—Bourguignon and Morrisson begin by disaggregating national income data into vintiles (5 percent groups, that is, twenty income groups per nation). National boundaries have changed over the past two centuries, of course, and nations have come and gone over the past two centuries. Even in nations where boundaries remained constant, we do not always have income data for the entire period. To overcome these problems, Bourguignon and Morrisson grouped the 199 nations with income data in 1992 into 33 homogeneous groups, each of which represented at least 1 percent of the world population or world GDP in 1950. The 33 groups include single nations (such as China and the United States) as well as large groups of small nations and small groups of medium-sized nations. From these 660 data points (33 nation groups x 20) it is a straightforward matter to apply the population-weighted formulas for the Theil index and the

mean logarithmic deviation (MLD)—two measures of inequality—to calculate summary measures of the world’s total inequality for different years. By collapsing the 199 nations into 33 groups, Bourguignon and Morrisson are able to extend their inequality series back to 1820. Note that, to the extent that their grouping strategy introduces bias, the bias is in the direction of underestimating between-nation inequality and inflating within-nation inequality, since some of the inequality within the nation groups is actually between-nation inequality. But that bias should not affect our basic conclusions about the relative growth in between-nation and within-nation income inequality over the past two centuries.

The results are striking (Figure 1). Two facts stand out. First, the *B/W* ratio—the ratio of between-nation to within-nation income inequality—is much higher now than it was in the early stages of Western industrialization. The increase in the *B/W* ratio reflects both a rise in between-nation income inequality and a decline in within-nation income inequality since 1820 (Table 1). By far the greater change is in between-nation income inequality, however. The Theil index for between-nation income inequality (actually, inequality between nation *groups*) shot up from 0.061 in 1820 to 0.513 in 1992, and the MLD shot up from 0.053 in 1820 to 0.495 in 1992. As anticipated, then, the Industrial Revolution of the past two centuries has increased income inequality across nations, but the magnitude of the increase is stunning. There has been a metamorphosis from a world where poverty was the norm in all nations to a richer world with much lower poverty rates (Bourguignon and Morrisson 1999, Table 1) but also with much greater income inequality across nations. Because the steep rise in between-nation income inequality was not accompanied by an increase in inequality within nations, where you live—your nation—is much more important in determining your income in today’s world than it was in the preindustrial world.

The Reversal of Historical Inequality Trends

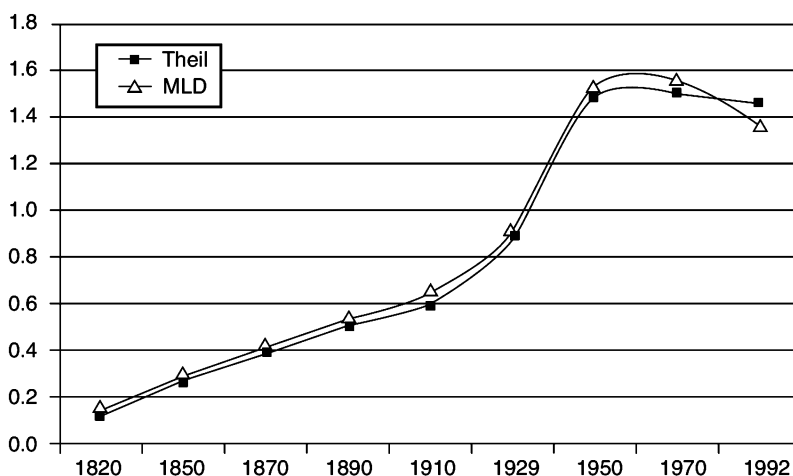


Figure 1. Ratio of Between-Nation to Within-Nation Income Inequality for Thirty-three Nation Groups, 1820–1992

Based on Table 1. Theil and MLD are measures of inequality

Table 1. Trends in Income Inequality Between and Within Thirty-three Homogeneous Nation Groups, 1920–1992

| Year | Between-nation groups | | Within-nation groups | |
|------|-----------------------|------|----------------------|------|
| | Theil | MLD | Theil | MLD |
| 1820 | .061 | .053 | .472 | .388 |
| 1850 | .128 | .111 | .477 | .393 |
| 1870 | .188 | .162 | .485 | .399 |
| 1890 | .251 | .217 | .498 | .408 |
| 1910 | .299 | .269 | .500 | .413 |
| 1929 | .365 | .335 | .413 | .372 |
| 1950 | .482 | .472 | .323 | .309 |
| 1970 | .490 | .515 | .324 | .330 |
| 1992 | .513 | .495 | .351 | .362 |

Source: Bourguignon and Morrisson (1999), table 3.

Estimates: *Between-nation groups*: From Bourguignon and Morrisson (1999), based on Maddison (1995) data set.

Within-nation groups: From Bourguignon and Morrisson (1999), based on updating of Berry, Bourguignon, and Morrisson (1983a,b) for the post–World War II period and on various sources (for example, Lindert 1999; Morrisson 1999) for the pre–World War II period. Bourguignon and Morrisson (2002) published modestly revised estimates for inequality. By collapsing vintiles to deciles, they report slightly lower estimates of inequality within the nation groups, but the fundamental conclusions are the same.

Note: Income measures are adjusted for purchasing power parity, and inequality measures are based on income vintiles (see Bourguignon and Morrisson 1999 for elaboration).

The second fact that stands out is that the growth in the B/W ratio stalled in the second half of the twentieth century. The B/W ratio stopped growing in the second half of the

twentieth century because growth in inequality across the nation groups has slowed dramatically since 1950. In the four decades after 1950, income inequality across the nation

groups increased by 6 percent using the Theil index and by 5 percent using the MLD. Over the four decades prior to 1950, the Theil had grown about 60 percent and the MLD had grown about 75 percent. Apparently the most dramatic effects of Western industrialization on between-nation inequality are over. After more than a century of sharp divergence in national incomes, the trend has been much more stable in recent years. This finding is in line with the findings of others (for example, Schultz 1998; Firebaugh 1999; Melchior, Telle, and Wiig 2000; Goesling 2001) who, drawing on data for individual nations instead of nation groups, find that between-nation income inequality is no longer rising. As we shall see later, between-nation income inequality declined in the late twentieth century when income data for the 1990s are added to data for the 1970s and 1980s.

Finally, it should be emphasized that the results here are so strong that the historical story they tell of increasing inequality across nations and of a rising B/W ratio cannot be dismissed as due to error in the data. To be sure, income estimates for the nineteenth century are gross approximations for many nations. But even if we make the extreme assumption that incomes are so drastically overstated for poorer nations in 1820 (or so drastically understated for richer nations in 1820) that the Theil and the MLD estimates understate between-nation income inequality in 1820 by a factor of three, that would still mean that between-nation income inequality tripled from 1820 to 1992 (from 0.18 to 0.51 based on the Theil index and from 0.16 to 0.50 for the MLD), and the B/W ratio still would have more than tripled for both inequality measures.

Phase 2 of the Transition: From Between-Nation Back to Within-Nation Inequality

The second phase of the inequality transition began in the second half of the twentieth century, with the stabilization of between-

nation income inequality in the 1960–1990 period and the decline in between-nation inequality beginning in earnest in the 1990s. Social scientists hardly have a stellar track record for predictions, especially with regard to sweeping predictions such as the one made here. Nonetheless there is sufficient theory and evidence that we can plausibly forecast that the B/W ratio will continue to decline in the twenty-first century.

The prediction of a declining B/W ratio is based on two separate conjectures. The first conjecture is that between-nation income inequality will decline, and the second conjecture is that within-nation income inequality will rise, or at least will not decline. These conjectures are based on the causes of the current trends, which I expect to continue and in some cases to intensify. (I am assuming that there will be no cataclysmic upheaval in the twenty-first century, such as a global war or a worldwide plague.) Recall the causes listed earlier for the inequality turnaround in the late twentieth century. I expect the major causes to continue, so between-nation income inequality will decline because of the continued industrialization of poor nations, because of the continued convergence of national economic policies and institutions arising from growing economic integration of national economies, because of the declining significance of labor immobility across national borders, and because of a demographic windfall for many poor nations. Within-nation income inequality will rise—or at least not decline—because of the continued industrialization of poor nations and because of continued growth in the service sector. Because between-nation inequality is the larger component, global income inequality will decline.

The conjecture of declining global income inequality is out of step with much of the globalization literature. A recurring theme in that literature is that we have entered a new information-based economic era where productive activity is becoming less dependent

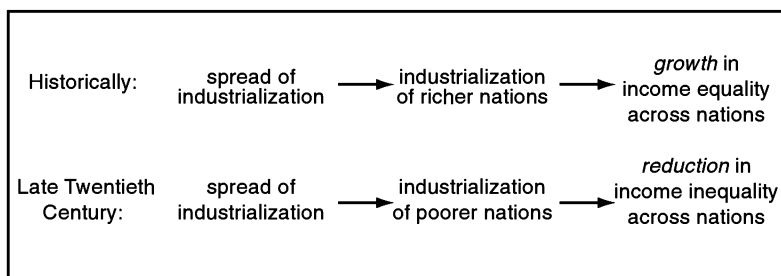


Figure 2. Industrialization and Between-Nation Income Inequality: Historically and in the Late Twentieth Century

on physical space, as a rising share of the world's economic output is produced in electronic space that knows no national borders (Sassen 2000). This phenomenon is possible because of the emergence of a global economy where income—and hence income inequality—is becoming increasingly rooted in knowledge rather than in capital goods (Reich 1991). What do these developments imply for global income inequality? For many globalization writers, the answer is clear: global inequality is bound to worsen because of the growing “global digital divide” that enlarges the gap between the “haves” and the “have-nots” (Campbell 2001; Ishaq 2001; Norris 2001).

The theoretical argument that a shift to a knowledge-based global economy would worsen global income inequality is shaky. It is not hard to think of reasons why the shift from an industrial-based to a knowledge-based global economy would reduce, not increase, inequality across nations. Knowledge is mobile, especially with today's telecommunication technologies that permit virtually instant worldwide codification and distribution of knowledge. In addition, because knowledge can be given away without being lost, the notion of property rights is more problematic in the case of knowledge, so it is harder to concentrate and monopolize knowledge across nations than it is to concentrate and monopolize capital goods across nations. Hence the switch to a knowledge-based global economy should mean that one's income is increasingly

determined by how much knowledge one obtains and uses as opposed to where one lives. The tighter link between knowledge and income in turn implies declining income inequality across nations and rising inequality within nations since—absent institutional barriers . . . —the variance in individuals' ability to obtain and benefit from knowledge is greater within nations than between them.

The issue of how the new information age will affect global inequality in the near term is not as decisive as often imagined, since as already noted, the coming of the information age is often much exaggerated. What is still most important in today's world is industrialization for the many, not digitization for the few. Historically the spread of industrialization has been the primary force driving the growth in between-nation income inequality. The initially richer nations of the West were the first to industrialize, and the poorer nations of Asia and Africa lagged behind. The new geography of inequality is also driven by the spread of industrialization, but the effects are different today: now the spread of industrialization means the diffusion of manufacturing technology to the world's largest poor regions. In recent decades inequality has declined across nations as industrialization has been an engine of growth in the most populous poor regions of the world, especially East Asia. That growth has worked both to compress inequality across nations (Figure 2) and to boost inequality within the industrializing nations.

The significance of this continuation of world industrialization has been lost in much of the literature on globalization, because of preoccupation with the idea that we are witnessing the emergence of a new knowledge-based technology regime. To be sure, in the categories used to classify world production, the output of the so-called service sector is estimated to exceed the output of the industrial sector for the world as a whole (World Bank 1997, table 12). Yet much of the service sector—an amorphous sector that includes wholesale and retail trade, the banking industry, government, the transportation industry, the commercial real estate industry, and personal services (including health care and education)—has arisen to grease the wheels of industry. Aside from the growth in personal services and government, much of the growth in the service sector has been for services for producers, not consumers—for example, the rise of an engineering industry to design better machines, and the growth of a banking industry and a commercial real estate industry for commercial transactions. In addition, many of the other so-called service industries—the transportation industry that distributes manufactured goods, for example, and the specialized retailing industry that sells the goods—benefit producers as well as consumers. In short, a significant portion of the growth in service industries over the past century can be seen as ancillary to the industrialization process.

With regard to income inequality within nations, we expect the continuing spread of world industrialization to boost inequality along the lines of the classic argument, from Simon Kuznets (1955), that industrialization boosts inequality (at least initially) as workers move from the lower-wage but larger agricultural sector to the higher-wage industrial sector. Importantly, this argument suggests that income inequality increases in poor nations as they industrialize because of income gains, not income losses. In other words, inequality

grows not because some people are becoming poorer but because some people are becoming richer, so the growth in income inequality in poor nations as they industrialize reflects rising rather than declining fortunes. If the industrialization of large poor nations does boost income inequality in those nations, at this juncture in history the continued spread of industrialization implies growth in income inequality in the average nation, because those nations are home to such a large fraction of the world's population. And there are no obvious counterforces on the horizon. Although the jump in income inequality following the collapse of communism in Eastern Europe and the former USSR is not likely to be repeated, there is no good reason to expect income inequality in those nations to fall in the near future. Nor is there any good reason to expect income inequality to decline notably in the West in the near future, either. We lack reliable inequality data for many African nations, but even if inequality were falling rapidly (which is unlikely), the effect of falling within-nation inequality in Africa would not offset the effect of rising within-nation inequality in Asia, given the relative sizes of the two regions. Few expect within-nation income inequality to decline sharply in Latin America in the near future, even though Latin America currently exhibits the highest level of within-nation income inequality of the world's major regions. Indeed, because of the advantage enjoyed by North American farmers with respect to some types of produce, economic integration in the Americas could exacerbate inequality by removing protections for farmers in Latin America. If that reasoning is correct, there is merit in the concern of WTO protesters that globalization may exacerbate income inequality within some poor nations by driving down incomes in the lower-income agricultural sector. But the bigger story is that industrialization (not the collapse of farm prices) will tend to drive up inequality within

nations, at least initially. Because many nations are still on the part of the Kuznets (1955) curve where migration from farm to factory boosts inequality, we can anticipate further growth in within-nation income inequality in the early decades of the twenty-first century.

REFERENCES

- Ben-David, Dan, Håkan Nordström, and L. Alan Winters. 1999. *Trade, Income Inequality, and Poverty*. WTO Special Study no. 5. Geneva: World Trade Organization.
- Berry, Albert, François Bourguignon, and Christian Morrisson. 1983a. "Changes in the world distribution of income between 1950 and 1977." *Economic Journal* 93:331–350.
- _____. 1983b. "The level of world inequality: How much can one say?" *Review of Income and Wealth* 29:217–241.
- Bourguignon, François, and Christian Morrisson. 1998. "Inequality and development: The role of dualism." *Journal of Development Economics* 57:233–257.
- _____. 1999. "The size distribution of income among world citizens: 1820–1990." Draft. June.
- _____. 2002. "Inequality among world citizens: 1820–1992." *American Economic Review* 92:727–744.
- Campbell, Duncan. 2001. "Can the digital divide be contained?" *International Labour Review* 140:119–141.
- Castells, Manuel. 1998. *End of Millennium*. Malden, Mass.: Blackwell.
- Chase-Dunn, Christopher, Yukio Kawano, and Benjamin D. Brewer. 2000. "Trade globalization since 1795: Waves of integration in the world-system." Special millennium issue, edited by Glenn Firebaugh, of the *American Sociological Review* 65:77–95.
- Doyle, Arthur Conan. 1955. *A Treasury of Sherlock Holmes*. Selected and with an introduction by Adrian Conan Doyle. Garden City, N.Y.: Hanover House.
- The Economist*. 1999. "The battle in Seattle." U.S. edition. November 27.
- Eggertson, Laura. 1998. "Rich-poor gap next issue for Martin." *Toronto Star*. June 3, p. A6.
- Firebaugh, Glenn. 1999. "Empirics of world income inequality." *American Journal of Sociology* 104:1597–1630.
- _____. 2003. *The New Geography of Global Income Inequality*. Cambridge: Harvard University Press.
- Fujita, Masahisa, Paul Krugman, and Anthony J. Venables. 1999. *The Spatial Economy: Cities, Regions, and International Trade*. Cambridge: MIT Press.
- Giddens, Anthony. 1999. "Globalization: An irresistible force." *The Daily Yomiuri*, June 7, p. 8.
- Goesling, Brian. 2001. "Changing income inequalities within and between nations: New evidence." *American Sociological Review* 66:745–761.
- Guillén, Mauro F. 2001. "Is globalization civilizing, destructive, or feeble? A critique of five key debates in the social science literature." *Annual Review of Sociology* 27:235–260.
- International Monetary Fund (IMF). 2000. "How we can help the poor." *Finance and Development* (December). (<http://www.imf.org/external/pubs/>)
- Ishaq, Ashfaq. 2001. "On the global digital divide." *Finance and Development* 38:44–47.
- Kuznets, Simon. 1955. "Economic growth and income inequality." *American Economic Review* 45:1–28.
- Lindert, Peter, and Jeffrey G. Williamson. 2000. "Does globalization make the world more unequal?" Paper presented at the "Globalization in Historical Perspective" preconference, National Bureau of Economic Research, Cambridge, Mass. November.
- Maddison, Angus. 1995. *Monitoring the World Economy, 1820–1992*. Paris: OECD.
- Mazur, Jay. 2000. "Labor's new internationalism." *Foreign Affairs* 79:79–93.
- Melchior, Arne, Kjetil Telle, and Henrik Wiig. 2000. *Globalization and Inequality: World Income Distribution and Living Standards, 1960–1998*. Studies on Foreign Policy Issues. Oslo: Royal Norwegian Ministry of Foreign Affairs.
- Milanovic, Branko. 2002. "True world income distribution, 1988 and 1993: First calculation based on household surveys alone." *Economic Journal* 112:51–92.
- Morrisson, Christian. 1999. "Historical perspectives on income distribution: The case of Europe." In *Handbook of Income Distribution*, ed. A. B. Atkinson and F. Bourguignon. Amsterdam: Elsevier.
- New Scientist*. 1999. Editorial. December 4.
- New York Times*. 1999. "National guard is called to quell trade talk protests." December 1.
- Norris, Pippa. 2001. *Digital Divide: Civic Engagement, Information Poverty, and the Internet Worldwide*. Cambridge and New York: Cambridge University Press.

- O'Rourke, Kevin H. 2001. "Globalization and inequality: Historical trends." National Bureau of Economic Research. Draft. April.
- Pritchett, Lant. 1996. "Forget convergence: Divergence past, present, and future." *Finance and Development* (June):40–43.
- . 1997. "Divergence, big time." *Journal of Economic Perspectives* 11:3–17.
- Quah, Danny T. 1997. "Increasingly weightless economies." *Bank of England Quarterly Bulletin* (February):49–55.
- Reich, Robert B. 1991. *The Work of Nations: Preparing Ourselves for Twenty-first-Century Capitalism*. New York: Knopf.
- Sachs, Wolfgang, ed. 1992. *The Development Dictionary: A Guide to Knowledge as Power*. London: Zed Books.
- Sassen, Saskia. 2000. "Territory and territoriality in the global economy." *International Sociology* 15:372–393.
- Schultz, T. Paul. 1998. "Inequality in the distribution of personal income in the world: How it is changing and why." *Journal of Population Economics* 11:307–344.
- United Nations Development Program. 1999. *Human Development Report 1999*. New York: Oxford University Press.

111 ■ Alan B. Krueger and Jitka Malečková

Does Poverty Cause Terrorism?

The Economics and the Education of Suicide Bombers

I
That investment in education is critical for economic growth, improved health, and social progress is beyond question. That poverty is a scourge that the international aid community and industrialized countries should work to eradicate is also beyond question. There is also no doubt that terrorism is a scourge of the contemporary world. What is less clear, however, is whether poverty and low education are root causes of terrorism.

In the aftermath of the tragic events of September 11, several prominent observers and policymakers have called for increased aid and educational assistance as a means for ending terrorism. "We fight against poverty," President George W. Bush has declared, "because hope is an answer to terror." But a careful review of the evidence provides little reason for optimism that a reduction in poverty

or an increase in educational attainment would, by themselves, meaningfully reduce international terrorism. Any connection between poverty, education, and terrorism is indirect, complicated, and probably quite weak. Instead of viewing terrorism as a direct response to low market opportunities or lack of education, we suggest it is more accurately viewed as a response to political conditions and long-standing feelings of indignity and frustration (perceived or real) that have little to do with economics.

An understanding of the causes of terrorism is essential if an effective strategy is to be crafted to combat it. Drawing a false and unjustified connection between poverty and terrorism is potentially quite dangerous, as the international aid community may lose interest in providing support to developing nations when the imminent threat of terrorism

recedes, much as support for development waned in the aftermath of the Cold War; and connecting foreign aid with terrorism risks the possibility of humiliating many people in less developed countries, who are implicitly told that they receive aid only to prevent them from committing acts of terror. Moreover, premising foreign aid on the threat of terrorism could create perverse incentives in which some groups are induced to engage in terrorism to increase their prospects of receiving aid. In our view, alleviating poverty is reason enough to pressure economically advanced countries to provide more aid than they are currently giving. Falsely connecting terrorism to poverty serves only to deflect attention from the real roots of terrorism.

To make any headway investigating the determinants of terrorism, one must have a working definition of terrorism. This is a notoriously difficult task. More than one hundred diplomatic and scholarly definitions of the term exist. . . .

The State Department, which acknowledges that no single definition of terrorism has gained universal acceptance, seems to have captured what is considered terrorism by many governments and international organizations. Since 1983, it has employed this definition for statistical and analytical purposes: “The term ‘terrorism’ means premeditated, politically motivated violence perpetrated against noncombatant targets by subnational groups or clandestine agents, usually intended to influence an audience. The term ‘international terrorism’ means terrorism involving citizens or the territory of more than one country.” The State Department also specifies that “the term noncombatant is interpreted to include, in addition to civilians, military personnel who at the time of the incident are unarmed and/or not on duty. . . . We also consider as acts of terrorism attacks on military installations or on armed military personnel when a state of military hostilities does not exist at the site, such as bombings against U.S. bases in the Persian Gulf, Europe, or else-

where.” The rub, of course, is that the definitions of “subnational” and “military hostilities” leave much latitude for disagreement.

The definitions of terrorism used by scholars, by contrast, tend to place more emphasis on the intention of terrorists to cause fear and terror among a targeted population that is considerably larger than the actual victims of their attacks, and to influence the views of that larger audience. The actual victim of the violence is thus not the main target of the terrorist act. Scholarly definitions often also include nation-states as potential perpetrators of terrorism.

Rather than dogmatically adhere to one definition, we have analyzed involvement in or support for activities that, at least when judged by some parties, constitute terrorism. Still, in the incidents that we have analyzed, the line between terrorism and resistance is often blurred. At the least, all of the cases we examine could be thought of as involving politically motivated violence. Moreover, it is reassuring that our main conclusions appear to hold across a varying set of circumstances, cultures, and countries. (We do not examine state terrorism because we suspect that the process underlying participation in state terrorism is quite different from the process underlying sub-state terrorism, and would involve a different type of analysis. We do not dispute that state terrorism exists, and that it has at times generated sub-state terrorism as a response.)

In economics, it is natural to analyze participation in terrorism in the framework of occupational choice. As is conventional in economics, involvement in terrorism is viewed as a rational decision, depending on the benefits, costs, and risks involved in engagement in terrorism compared with other activities. Not surprisingly, the standard rational-choice framework does not yield an unambiguous answer to the question of whether higher income and more education would reduce participation in terrorism.

In this context, we have also reviewed evidence on “hate crimes,” which can be viewed

as a close cousin to terrorism in that the target of an offense is selected because of his or her group identity, not because of his or her individual behavior, and because the effect of both is to wreak terror in a greater number of people than those directly affected by the violence. A consensus is emerging in the social science literature that the incidence of hate crimes, such as lynchings of African Americans or violence against Turks in Germany, bears little relation to economic conditions.

Most significantly, we have considered data from a public-opinion poll conducted in the West Bank and Gaza Strip by the Palestinian Center for Policy and Survey Research (PCPSR). In December 2001, Palestinians were asked whether they supported attacks on Israeli civilian and military targets, and about whether they considered certain incidents acts of terrorism. Breaking down the data by education and occupation indicates that support for violence against Israeli targets is widespread in the Palestinian population, and at least as great among those with higher education and higher living standards as it is among the unemployed and the illiterate. Similarly, a review of the incidence of major terrorist acts over time in Israel, and an analysis that relates the number of terrorist acts each year to the rate of economic growth in that year or in the recent past, yields the same skepticism about the idea that poverty is a cause of terrorism.

The data on participation in and support for political violence, militancy, and terrorism that we have examined are meager, often indirect, and possibly nonrepresentative. In addition, participation in terrorist activities may be highly context-specific, and we have examined terrorism, militancy, and politically motivated violence in a small number of settings primarily in the Middle East. Consequently, our results must be considered tentative and exploratory. Yet we are not aware of compelling evidence that points in the opposite direction from what we have found. In light of our results, we would urge intellectuals and policymakers to exercise caution in presum-

ing that poverty and education have a direct and causal impact on terrorism.

II

A simple view of terrorism is that participation in terrorism is akin to participation in crime in general. Economists have a well-developed and empirically successful theory of participation in criminal activities. As emphasized by Gary Becker, individuals should choose to allocate their time between working in the legal job market or working in criminal activities in such a way that maximizes their utility. After accounting for the risk of being caught and penalized, the size of the penalty, and any stigma or moral distress associated with involvement in crime, those who receive higher income from criminal activities would choose involvement in crime. According to this model, crime increases as one's market wage falls relative to the rewards associated with crime, and decreases if the risk of being apprehended after committing a crime, or the penalty for being convicted of a crime, rises. Available evidence suggests that individuals are more likely to commit property crimes if they have lower wages or less education; but the occurrence of violent crimes, including murders, is typically found to be unrelated to economic opportunities.

Some economists, notably William Landes, Todd Sandler, and Walter Enders, have applied the economic model of crime to transnational terrorism. They focus on how an increase in penalties and law enforcement influences the incentive to partake in terrorist activities. But the economic model yields few concrete predictions insofar as the relationship between market opportunities and participation in terrorism is concerned, because participation in terrorist acts by individuals with different characteristics depends on the probability that participation will bring about the desired political change, as well as the differential payoff for the various groups associated with achieving the terrorists' desired aims versus the penalties as-

sociated with failure. It is possible, for example, that well-educated individuals will disproportionately participate in terrorist groups if they think that they will assume leadership positions if they succeed, or if they identify more strongly with the goals of the terrorist organization than less-educated individuals.

Other important considerations include the relative pay of skilled and unskilled individuals for participation in terrorist organizations and how it compares to relative pay in the legal sector, and the selection of particular terrorists by terrorist organizations. Bill Keller recently reported in *The New York Times* that Iraq decided in March to increase the payment to families of suicide bombers in the West Bank and Gaza from \$10,000 to \$25,000. In the month after that decision, suicide bombings increased, though it is unclear whether the connection is causal.

Even before the increase in the payment to families of suicide bombers, there was a large supply of willing suicide bombers, as Nasra Hassan, a relief worker for the United Nations, reported last year in *The New Yorker*. Between 1996 and 1999, Hassan interviewed nearly two hundred fifty militants and associates of militants involved in the Palestinian cause, including failed suicide bombers, the families of deceased bombers, and those who trained and prepared suicide bombers for their missions. One Hamas leader whom Hassan interviewed remarked: "Our biggest problem is the hordes of young men who beat on our doors, clamoring to be sent [on suicide missions]. It is difficult to select only a few." A senior member of the al-Qassam Brigades said: "The selection process is complicated by the fact that so many wish to embark on this journey of honor. When one is selected, countless others are disappointed." Thus, the demand side is also part of the equation.

With a queue of willing participants, how do terrorist or militant groups choose their suicide bombers? A planner for Islamic Jihad explained to Hassan that his group scrutinizes the motives of a potential bomber to be sure

that the individual is committed to carrying out the task. Now, a high level of educational attainment is probably a signal of one's commitment to a cause, as well as of one's ability to prepare for an assignment and carry it out. For this reason, the stereotype of suicide bombers being drawn from the ranks of those who are so impoverished that they have nothing to live for may be wildly incorrect. This interpretation is also consistent with another of Hassan's observations about suicide bombers: "None of them were uneducated, desperately poor, simple-minded, or depressed. Many were middle class and, unless they were fugitives, held paying jobs. More than half of them were refugees from what is now Israel. Two were the sons of millionaires."

Suicide bombers clearly are not motivated by the prospect of their own individual economic gain, although it is possible that the promise of larger payments to their families may increase the willingness of some to participate in these lethal missions. We suspect their primary motivation instead results from their passionate support for the ideas and the aims of their movement. "Over and over," Hassan reported, "I heard them say, 'The Israelis humiliate us. They occupy our land and deny our history.'" The eradication of poverty and the attainment of universal high school education are unlikely to change these feelings. Indeed, it is even possible that those who are well-off and well-educated experience such feelings more acutely. . . .

III

The Palestinian Center for Policy and Survey Research (PCPSR) is an independent, non-profit research organization in Ramallah that conducts policy analysis and academic research in the West Bank and Gaza. In December 2001, PCPSR conducted a public-opinion poll of 1,357 Palestinians eighteen years old or older in the West Bank and Gaza. The survey, which was conducted by in-person interviews, covered topics including the participants' views

toward the September 11 attacks in the United States, the participants' support for an Israeli-Palestinian peace agreement, and their opinions about armed attacks against Israel. Under trying circumstances in the middle of one of the worst periods of the Israeli-Palestinian conflict, a researcher at the center kindly provided us with tabulations of key questions broken down by the educational level and the occupational status of the respondents.

Although public-opinion polls are subject to multiple interpretations, such data can provide indirect information about which segments of the population support terrorist or militant activities. The PCPSR poll reveals several things. First, the support for armed attacks against Israeli targets by the Palestinian population is widespread, though it is important to emphasize that there is a distinction between support for armed attacks expressed in a public-opinion poll at a particular point in time and participation in or active support for such attacks. Second, a majority of the Palestinian population believes that armed attacks against Israeli civilians have helped to achieve Palestinian rights in a way that negotiations could not have achieved. This finding raises obvious implications concerning the difficulty of ending the attacks, and may partially account for the Palestinian public's opposition to a United Nations initiative to fight terrorism, which was also found in the poll. If the Palestinian public believes the attacks are efficacious, they are unlikely to cease supporting additional attacks unless their demands are met. Another question asked was: "To what extent do you support or oppose the position taken by President Arafat and the [Palestinian Authority] regarding the U.S. campaign against terror?" Thirty-six percent supported or strongly supported the position of Arafat in this case, and 50.4 percent opposed it.

Moreover, a majority of the Palestinian population did not consider suicide bombings, such as the one that killed twenty-one Israeli youths at the Dolphinarium disco in Tel Aviv, terrorist events. Toward the end of the ques-

tionnaire, respondents were also asked whether they thought the international community considered the Dolphinarium bombing a terrorist event. Ninety-two percent responded yes. These results highlight important differences in interpreting the meaning of the word "terrorism."

Most important for our purposes, there is no evidence in these results that more highly educated individuals are less supportive of violent attacks against Israeli targets than are those who are illiterate or poorly educated. Consider the percentage of individuals who say they support or strongly support armed attacks against Israeli targets less those who say they oppose or strongly oppose such attacks. By a margin of 68 points, those with more than a secondary school education support armed attacks against Israeli targets, while the margin is 63 points for those with an elementary school education and 46 points for those who are illiterate.

A survey conducted by PCPSR in November 1994, before the current intifada, asked respondents whether they supported a dialogue between Hamas and Israel. Responses were reported by educational attainment. More highly educated respondents were less supportive of a dialogue with Israel: 53 percent of those with a B.A. and 40 percent of those with an M.A. or a Ph.D. supported a dialogue, compared with 60 percent of those with nine years of schooling or less. (Based on other questions, it is clear that supporters of dialogue generally favored a more peaceful coexistence with Israel.)

The PCPSR study in 2001 showed also that support for armed attacks against Israeli targets is widespread across all Palestinian occupations and groups, but particularly strong among students (recall that respondents are age eighteen or older) and merchants and professionals. Notably, the unemployed are somewhat less likely to support armed attacks against Israeli targets. If poverty was indeed the wellspring of support for terrorism or politically motivated violence, one would

have expected the unemployed to be more supportive of armed attacks than merchants and professionals, but the public-opinion evidence points in the other direction. . . .

We also performed a detailed analysis of participation in Hezbollah in Lebanon, which has reportedly instructed Palestinian extremist groups on the use of suicide bomb attacks. We compared the background characteristics of 129 members of Hezbollah's militant wing who died in action mostly in the late 1980s to those of a random sample of 121,000 young people in the Lebanese population. Many of these militants died fighting Israeli occupation, while others died in suicide bomb attacks or while planting booby traps. The Hezbollah militants in this sample are likely to be representative of those who engaged in terrorist acts, and some were carrying out terrorist acts when they died. Compared to the general Lebanese population from the same age group and region, the Hezbollah militants were actually slightly less likely to come from impoverished households, and were more likely to have attended secondary school or higher. These results suggest that the militants were not particularly drawn from those with the least opportunities in society.

And this conclusion is ratified by political violence on the other side as well. In the late 1970s and early 1980s, numerous violent attacks against Palestinians were conducted by Israeli Jews in the West Bank and Gaza. These attacks included attempts to kill three Palestinian mayors of West Bank cities and attempts to blow up the Dome of the Rock mosque in Jerusalem. From 1980 to 1984, a total of twenty-three Palestinians were killed in attacks by what became known among Israelis as the Jewish Underground, and 191 Palestinians were injured. The *International Encyclopedia of Terrorism* (1997) refers to these attacks as acts of terrorism. In a ruling in 1985, an Israeli court convicted three Israeli settlers of murder and found others guilty of violent crimes in cases involving attacks in the West Bank.

What were the biographical backgrounds of those involved in these violent attacks by Israeli Jewish extremists? A list may be compiled of the name, the age, the occupation, and the nature of underground activity for twenty-seven individuals involved in the Jewish Underground in the early 1980s, based mainly on a memoir of the Jewish Underground by Haggai Segal, one of its members. It is clear from such a chart that these Israeli extremists were overwhelmingly well-educated and in high-paying occupations. The list includes teachers, writers, university students, geographers, an engineer, a combat pilot, a chemist, and a computer programmer. As Donald Neff observed in 1999 about the three men convicted of murder, "all were highly regarded, well-educated, very religious." Although we have not statistically compared the backgrounds of these extremists to the wider Israeli population, these twenty-seven individuals certainly do not appear to be particularly underprivileged or undereducated.

IV

The evidence that we have assembled and reviewed suggests that there is little direct connection between poverty, education, and participation in or support for terrorism. Indeed, the available evidence indicates that compared with the relevant population, participants in Hezbollah's militant wing in the late 1980s and early 1990s in Lebanon were at least as likely to come from economically advantaged families and to have a relatively high level of education as they were to come from impoverished families without educational opportunities. We should caution, however, that the evidence we have considered is tentative due to data limitations. In addition, our focus has been primarily on the Middle East, so our conclusions may not generalize to other regions or circumstances. . . .

While economic deprivation may not be associated with participation in terrorism and political violence at the individual level, it may

nonetheless matter at the national level. If a country is impoverished, a minority of the relatively well off in that country may turn to terrorism to seek to improve the conditions of their countrymen. One might question, though, whether the goal of many terrorist organizations is to install a political regime that is likely to reduce poverty. In addition, there are well-documented cases of homegrown terrorism in economically advanced countries (remember Timothy McVeigh?), so it is far from clear that poverty at a national level is associated with support for terrorism. Of course, this question can only be addressed by cross-country analyses.

In addition, poverty may indirectly affect terrorism through the apparent connection between economic conditions and the proclivity for countries to undergo civil wars. James Fearon and David Laitin have found that GDP per capita is inversely related to the onset of civil war, and Paul Collier and Anne Hoeffler have found that the growth rate of GDP per capita and the male secondary-school enrollment rate are inversely related to the incidence of civil war. Lebanon, Afghanistan, and Sudan are high-profile examples of countries where civil war provided a hospitable environment for international terrorists to operate. But there are other situations in which countries undergoing a civil war did not provide a breeding ground for international terrorism, so it is unclear how much one should extrapolate from the relationship between economic development and civil war. And terrorism has arisen in many countries that were not undergoing a civil war.

Enough evidence is accumulating that it is fruitful to begin to conjecture why participation in terrorism and political violence is apparently unrelated—or positively related—to individuals' income and education. The standard economic model of crime suggests that those with the lowest value of time should engage in criminal activity. But we would hypothesize that in most cases terrorism is less like property crime and more like a violent

form of political engagement. More-educated people from privileged backgrounds are more likely to participate in politics, probably in part because political involvement requires some minimum level of interest, expertise, commitment to issues, and effort, all of which are more likely if people are educated enough and prosperous enough to concern themselves with more than economic subsistence. These factors could outweigh the effect of opportunity cost on individuals' decisions to become involved in terrorism.

The demand side for terrorists must be considered as well as the supply side. Terrorist organizations may prefer highly educated individuals over less-educated ones, even for suicide bomb attacks. In addition, educated middle-class or upper-class individuals are better suited to carry out acts of international terrorism than are impoverished illiterates, because the terrorists must fit into a foreign environment to be successful. This consideration suggests that terrorists who threaten economically developed countries will disproportionately be drawn from the ranks of the relatively well off and highly educated.

On the whole, we must conclude that there is little reason to be optimistic that a reduction in poverty or increase in educational attainment will lead to a meaningful reduction in the amount of international terrorism without other changes. Jessica Stern has observed that many madrasahs, or religious schools, in Pakistan are funded by wealthy industrialists, and that those schools deliberately educate students to become foot soldiers and elite operatives in various extremist movements around the world. She further reported that "most madrasahs offer only religious instruction, ignoring math, science, and other secular subjects important for functioning in modern society." These observations suggest that, in order to use education as part of a strategy to reduce terrorism, the international community should not limit itself to increasing years of schooling, but should consider very carefully the content of education.

ABOUT THE EDITORS

David B. Grusky is Professor of Sociology at Stanford University, Director of the Center for the Study of Poverty and Inequality at Stanford University, and coeditor (with Paula England) of the *Social Inequality Series*. He received his B.A. at Reed College, his M.S. and Ph.D. at the University of Wisconsin-Madison, and has held positions at the University of Chicago, Cornell University, and Stanford University. He is a Fellow of the American Association for the Advancement of Science, recipient of the 2004 Max Weber Award (with Maria Charles), founder of the Cornell University Center for the Study of Inequality, and a former Presidential Young Investigator. His research addresses issues of inequality and takes on such questions as whether and why gender, racial, and class-based inequalities are growing stronger, why they differ in strength across countries, and how such changes and differences are best measured. His recent books are *Poverty and Inequality* (coedited with Ravi Kanbur, Stanford University Press, 2006), *Mobility and Inequality* (coedited with Stephen Morgan and Gary Fields, Stanford University Press, 2006), *Occupational Ghettos* (coauthored with Maria Charles, Stanford University Press, 2004), *The Declining Significance of Gender?* (coedited with Francine Blau and Mary Brinton, Russell Sage, 2006), *The Inequality Reader: Contemporary and Foundational Readings in Race, Class, and Gender* (coedited with Szonja Szelényi, Westview Press, 2007), and *Inequality: Classic Readings in Race, Class, and Gender* (coedited with Szonja Szelényi, Westview Press, 2006).

Mamwai C. Ku is a Ph.D. candidate in the Department of Sociology at Stanford University. She received her B.A. from the University of Pennsylvania and her M.A. from Stanford University. Her research focuses on issues of race and gender and addresses such topics as racial identification and life outcomes among the multiracial population, gender differences in career paths within the professions, and intra-occupational sex segregation.

Szonja Szelényi is Associate Director of the Center for the Study of Poverty and Inequality at Stanford University. Her books are *Equality by Design* (Stanford University Press, 1998), *The Inequality Reader: Contemporary and Foundational Readings in Race, Class, and Gender* (coedited with David B. Grusky, Westview Press, 2007), and *Inequality: Classic Readings in Race, Class, and Gender* (coedited with David B. Grusky, Westview Press, 2006).

ABOUT THE BOOK AND OUR NEW SERIES OF INEQUALITY READERS

The publication of the third edition of *Social Stratification* completes the recently undertaken expansion of our series of inequality anthologies into a full three-volume set. This expansion is a response to the burgeoning interest in poverty and inequality, the increasingly diverse research literature within the poverty and inequality field, and the increasingly diverse needs of consumers of this literature. The resulting series now includes separate books for the advanced student, the introductory student, and the student needing a classics-only compilation.

An advanced reader: The original book in our three-part series, *Social Stratification*, combines a full roster of classic readings with the most advanced, cutting-edge scholarship on poverty and inequality. Not for the faint of heart, *Social Stratification* is appropriate for anyone requiring full exposure to advanced scholarship in the field, including the most influential quantitative and technical pieces. Now fully revised and updated, *Social Stratification* remains the definitive industry standard, the most complete introduction available.

An introductory reader: *The Inequality Reader* is intended for readers requiring a more streamlined and less technical introduction to the field. This volume draws broadly from classic theoretical and conceptual pieces, qualitative and ethnographic studies, quantitative analyses of the sources of poverty and inequality, and “experiential pieces” that convey the lived experience of poverty and inequality. In assembling this volume, a special effort was made to include articles that cast light on recent changes in the structure of poverty and inequality, not only the spectacular takeoff in income inequality but also equally dramatic developments in contemporary race, class, and gender inequalities.

A compilation of the classics: Published in 2006, *Classic Readings in Race, Class, and Gender* is a yet slimmer volume of essential readings in poverty and inequality, readings that define the conceptual and empirical foundations of the field. This book, which represents the classical core alone, may be supplemented with contemporary articles that are tailored to particular topics or interests.

INDEX

- Absolute deprivation hypothesis, 904
- Acland, Henry, on effect of family/schooling in America, 498–504
- Age
- power of old in socializing young, 46–47
 - sibling correlations and social ascription, 596–608
- Agrarian society. *See* Horticultural/agrarian stratification
- Agricultural populations, 958
- Alienation
- class conflict in capitalist world economy, 111–113
 - class structure analysis, 98–110
 - class structure of advanced societies, 132–142
 - class/class conflict, 91–97
 - class/status/party, 114–124
 - ideology, 89–90
 - Marxism and class theory, 143–158
 - open/closed relationships, 128–132
 - pre-capitalism/capitalism, 79–88
 - social classes, 74–78
 - status/groups/classes, 124–128
- Allocation and inequality, 12–13
- An American Dilemma* (Myrdal), 351
- Anomie, and poverty in Chicago neighborhoods, 380
- Archer Park (Chicago) and race/ethnicity, 721–722
- Arrow impossibility theorem, 54, 59n2
- Arts and gender discrimination, 747–758
- Arts, Chan and Goldthorpe on stratification of, 916–925
- Ascriptive processes, and inequality, 6
- Asiatic stratification, 13–15
- Assimilation. *See* Immigration
- Associational change and social classes, 171–172
- Attainment
- educational/early occupational, 516–526
 - and Granovetter on strength of weak ties, 576–579
 - Lin on social networks and, 580–582
 - new social psychological model of, 542–549
- Attribution error, employment discrimination and, 774–775
- Audit studies, 12
- Authority and legitimate power, 94–97
- Baltzell, Digby, 306
- Bane, Mary Jo, on effect of family/schooling in America, 498–504
- Beck, Ulrich, on meta-change of modern society, 1030–1035
- Belkin, Lisa, “Opt-Out Revolution,” 735–738
- Bell, Daniel, on post-industrialist society, 966–978
- Beltway (Chicago) and race/ethnicity, 719–720
- Bertrand, Marianne, on labor market discrimination, 673–677
- Bielby, William T., on sex segregation, 786–796
- Blalock, Hubert, 717–718
- Blau, Francine D., on gender wage gap, 843–858
- Blau, Peter M.
- on status of occupations, 187–190
 - on stratification process, 486–497
- Bonacich, Edna, on ethnic antagonism, 632–644
- Bourdieu, Pierre, on judgement of taste, 870–892
- Bourgeoisie
- and class conflict in world economy, 113
 - class conflicts and, 308–309
 - and Gouldner on rise of a New Class, 299–300
 - market economy and, 80–83
- Boushey, Heather, and women opting-out from the workforce, 739–746
- Breen, Richard
- on rational action theory of educational differentials, 610–621
 - on social mobility in Europe, 465–480
- Brezhnev, Leonid, 35, 39n4

- Brooks, David, on new upper class, 304–310
- Burt, Ronald S., on structural holes, 583–585
- Canadian Self-Sufficiency Project (SSP), 402
- Capitalism, 78
- as built by non-capitalists, 1000–1005
 - class awareness and, 294–295
 - class conflict in world economy (Wallerstein), 111–113
 - decomposition due to separation of
 - ownership/control, 91–92
 - and feudalism, 110n11
 - Marx on, 79–88
 - post-communist managerialism and, 311–316
- Caste society stratification, 13, 16, 27n17. *See also* Stratification/stratification systems
- Categorization, employment discrimination and, 771–773. *See also* Race/ethnicity
- Ceaușescu, Nicolae, 34–35
- Center for Economic and Policy Research (CEPR), 740–741
- Chan, Tak Wing, on social stratification of arts, 916–925
- Change and social classes
 - associational, 171–172
 - sociotechnical, 169–171*See also* Social classes
- Chicago Gautreaux Housing Program, 393–399
- Chicago race/ethnicity, 716–726
- Civil assets, and stratification, 6
- Clark-Kauffman, Elizabeth, on children and welfare policies, 400–406
- Class analysis, rent-based
 - class and wealth, 219–223
 - class and wealth as life conditions, 223–224
 - exploitation classes/collective action in modern capitalism, 229–232
 - wealth/rent/exploitation, 224–229
- Class conflict
 - in capitalist world economy (Wallerstein), 111–113
 - institutionalization of, 93–94
 - labor markets and, 20–21
 - merging cultures and, 308–309
 - power/authority and, 94–97
- Class differentials, rational action theory of, 614–617
- Class persistence. *See* Rigidity
- Class structure
 - absolute mobility and, 469–473
 - American apartheid, 349–359
 - Brooks on new upper class, 304–310
 - class mobility, 420–425
 - Giddens on advanced societies
 - structuration of class, 138–142
 - theory of class, 135–138
 - Weberian critique, 132–135
 - Gouldner on rise of a New Class, 295–303
 - Hout on realignment of U.S. presidential voting, 945–951
 - judgement and taste
 - distinction, 881–882
 - food space, 882–887
 - form/substance, 878–881, 887
 - habitus, 871, 874–877
 - homologies between class conditions, 877–878
 - petit-bourgeois experience and, 888–890
 - social space and, 870–873
 - Lareau on class/race/family life, childrearing strategies, 926–927
 - leisure class
 - conspicuous consumption, 865–869
 - conspicuous leisure, 864–865
 - pecuniary canons of taste, 869–870
 - pecuniary emulation, 862–864
 - Mills on power elite, 275–285
 - Mosca on ruling class, 268–274
 - neo-Marxist analyses
 - domination vs. exploitation, 99
 - polarization/nonpolarization of class, 98–99
 - postcapitalist classes, 99–100
 - Pakulski and Waters on death of, 1022–1025
 - Roemer on class exploitation
 - assets/exploitation/classes, 105–108
 - exploitation concept, 100–101
 - game-theory approach, 102–103
 - general class analysis, 103–105
 - labor-transfer approach, 101–102
 - middle class in contemporary locations, 108–109
 - in terms of domination within production, 103
 - and stratification order of capitalistic societies, 1025–1029*See also* Mobility
- Class system stratification, 13
- Class theories
 - alienation
 - class conflict in capitalist world economy, 111–113
 - class structure analysis, 98–110

- class structure of advanced societies, 132–142
- class/class conflict, 91–97
- class/status/party, 114–124
- ideology, 89–90
- Marxism and class theory, 143–158
- open/closed relationships, 128–132
- pre-capitalism/capitalism, 79–88
- social classes, 74–78
- status/groups/classes, 124–128
- Marx on constitution of class, 88
- Marx on ideology and, 89–90
- Marxism
 - boundaries of property ownership, 144–146
 - class reproduction, 153–156
 - manual/nonmanual groups, 143–144
 - social closure, 146–153
 - possessing vs. proletarian self-alienation, 78
 - types of class struggle, 153–156
- Class-specific cultures, inequality and, 8–10
- Class/status/party, Weber on
 - communal action from class interest, 116–117
 - determination of class by market, 114–116
 - economically determined, 114
 - effects of status stratification on economic conditions, 121–123
 - ethnic segregation/caste, 119–121
 - parties/power, 123–124
 - status honor, 118–119
 - status privileges, 121
 - status stratification guarantees, 119
 - types of class struggle, 117–118
 - See also* Status groups/classes
- Closure, income inequality
 - closure mechanisms/strategies, 177–179
 - closure theory, 176–177
 - closure/earnings, 182–183
 - data/variables, 179–182
 - methods/models, 179
- Cohen, David, on effect of family/schooling in America, 498–504
- Collective efficacy, and poverty in Chicago neighborhoods, 379–385
- Collectivist exclusion, 155
- Communist Party
 - economic inequality and, 34–35
 - and Giddens on power elite, 287
 - and Gouldner on rise of a New Class, 297–298
 - post-communist managerialism, 311–316
 - See also* Socialism
- Community Survey of the Project on Human Development in Chicago Neighborhoods (PHDCN), Chicago neighborhoods and, 380–384
- Comparable worth, 834
- Competition, labor and development of, 86
- Conley, Dalton, on sibling correlations, 596–608
- Conspicuous consumption, leisure class and, 865–869
- Conspicuous leisure, leisure class and, 864–865
- Constant Flux: A Study of Class Mobility in Industrial Societies* (Erikson and Goldthorpe), 465, 476
- Consumer credit, income inequality and, 66
- Corporate community, Domhoff on power and, 291
- Correll, Shelley J., on workplace penalties for motherhood, 759–768
- Crime, and Pager on finding work after incarceration, 683–689
- Crystallization
 - and examples of ideal-typical systems, 13
 - and stratification, 6
 - Structuralist Phase of social class measurement and, 9
- Cultural assets, and inequality, 6
- Cultural environment
 - and evolution of modern stratification systems, 1008–1009
 - and Gouldner on rise of a New Class, 300–303
 - inequality of opportunity and, 20
 - neighborhoods/poverty and, 360–368
 - and poverty in America, 311–316
 - and poverty in America comparatively, 311–316
- Cultural Phase of social class measurement, 9–11
- Culture of critical discourse (CCD), 300–301
- Current Population Survey's Outgoing Rotation Group, 739
- Cynicism and anomie, 381–385
- Dahrendorf, Ralf, 91–97
- Davis, Kingsley, 30–33, 41
- Dealignment in U.S. presidential voting, 945–951
- DeLuca, Stefanie, and Chicago Gautreaux Housing Program, 393–399
- Demand side accounts, income inequality and, 63–64

- Denton, Nancy, on American apartheid, 349–359
- Di Carlo, Matthew, on labor market gradation, 249–264
- Digital inequality, 936–944
- Disaggregation of social classes
 class identification, 166–167
 class outcomes, 168
 collective action, 167–168
 social closure, 167
- Discrimination. *See* Gender inequality;
 Race/ethnicity
- Division of labor, Durkheim on, 159–164
- Domhoff, G. William, on power and politics, 290–295
- Domination
 vs. exploitation (class), 99
 within production, 99
See also Authority
- Dover (Chicago) and race/ethnicity, 720–721
- Duncan, Greg J., on children and welfare policies, 400–406
- Duncan, Otis Dudley
 on status of occupations, 187–190
 on stratification process, 486–497
- Dunlop, John T., on industrialism, 954–965
- Durkheim, Emile
 anomie and, 380
 “Division of Labor in Society,” 159–164
- Earnings. *See* Income inequality and earnings
- East European countries, socialism and, 33–39
- East Germany, socialism and, 33–39
- Economic assets, and inequality, 6
- Economic inequalities
 and declining significance of race, 693–695
 employment discrimination and, 771–777
 globalization and, 1036–1043
 and influence of system of production, 695–698
 and poverty as cause of terrorism, 1056–1062
 and poverty in America, 311–316
 and poverty in America comparatively, 311–316
See also Labor
- Economic inequality
 class conflicts and, 97
 education/training as solution and, 58–59
 and income inequality
 illustrations of contrast, 237–240
 interpersonal utility comparisons, 240–242
 money distribution, 235
 need for analysis, 235–237
 quality of life/functionings/capabilities, 242–245
 weights/valuations/explicitness, 245–246
- intentional policies and, 50–52
- Marx on alienation and social classes, 74–78
- Marx on proletarian class, 84–87
See also Inequality
- Education, 297
 and Brooks on new upper class, 304–310
 class mobility and, 420–425
 and early occupational attainment process, 516–526
 effect on families in America of, 498–504
 industrialism/industrial man and, 956–957
 institutional accounts and, 64
 need for minimum to offset inequality, 58
 new social psychological model of attainment
 commitment/preparation, 542–544
 decision tree model (stochastic) for
 commitment, 546–549
 decision tree model for choices, 544–546
 and poverty as cause of terrorism, 1056–1062
 and prestige of occupations, 194
 and rational action theory of educational
 differentials, 610–621
 stereotype threat and African-American
 student achievement, 678–682
 and welfare’s effect on children, 403
- Ehrenreich, Barbara, on poverty in America, 317–326
- Elites. *See* Power
- Employees. *See* Labor
- England, Paula, on gender wage gap, 834–837
- Enlightenment
 inequality trend and, 18–20
 working class stratification and, 17
- Erikson, Robert, on class mobility, 437–462
- Esping-Andersen, Gøsta, on postindustrial economies, 978–992
- Essentialism and sex segregation, 816–817, 820–824
- Ethnic antagonism
 exclusion movements/caste systems and, 632–633
 expensive labor group and, 640–643
 and factors affecting initial price of labor, 634–637
 and price of labor, 637–638
 split labor market and, 633–634, 638–640
- Ethnicity. *See* Race/ethnicity
- Exclusion, collectivist/individual, 155

- Exploitation
 assets/exploitation/classes, 105–108
 and distribution of consumption,
 110–111n18
 domination vs., 99
 and domination within production, 99
 exploitation concept, 100–101
 game-theory approach, 102–103
 general class analysis, 103–105
 labor-transfer approach, 101–102
 middle class, 108–109
 and rent-based class analysis, 229–232
 in terms of domination within production,
 103
- Eyal, Gil
 on capitalism, 1000–1005
 on post-communist managerialism, 311–316
- Fair Housing Act, 351–352, 357
- Families. *See also* Gender inequality; Labor
 annual household food expenditures,
 884–885
 and effect of schooling in America, 498–504
 immigration, segmented assimilation,
 658–664
 and income in adulthood, 505–514
 jobless poverty and, 345
 Lareau on class/race/family life
 children's time use, 928–930
 language use, 930–932
 parental intervention, 932–935
 study participants/observations, 927–928
 neighborhoods/poverty and, 360–368
 and new social psychological model of
 attainment, 516–526
 and welfare's effect on children, 400–406
 and women on welfare, 407–416
- Family Transition Program (FTP), 402
- Farley, Reynolds, on racial identities in 2000,
 624–631
- Feagin, Joe R., on antiblack discrimination,
 703–709
- Featherman, David L.
 on occupational achievement, 202–204
 on occupational mobility, 425–435
- Feminism, and women opting-out from the
 workforce, 735–746. *See also* Gender
 inequality
- Fernandez, Roberto M., on
 networks/race/hiring, 587–595
- Fernandez-Mateo, Isabel, on
 networks/race/hiring, 587–595
- Feudalism
 bourgeoisie and, 86
 and Roemer on class/exploitation, 104
 Feudalism as type of stratification, 13, 15–16,
 27n13–16
 pre-capitalism/capitalism and, 79–80
See also Stratification/stratification systems
- Firebaugh, Glenn, on global income inequality,
 1044–1055
- Fischer, Claude S., 49–53
- Functionalists, inequality and, 33
- Game-theory approach and exploitation,
 102–103
- Gautreaux Housing Program, 393–399
- Gemeinschaftlich* character of social classes,
 165–172
- Gender inequality, 4
 Bielby on sex segregation
 allocation of effort, 787–789
 noneconomic bases of discrimination,
 789–792
 neoclassical views, 786–787
 reconciliation of perspectives, 792–799
 and England on gender wage gap, 834–87
 and Giddens on social justice, 102
 Goldin on gender discrimination in
 performing arts, 747–758
 Grusky and Levanon on sex segregation,
 812–824
- Income inequality
 between-class (BC) component, 258–264
 between-occupation/within-class (BO/WC)
 component, 258–264
 within-occupation variance (WO),
 258–264
 and Jacobs on women's careers, 797–801
 Kahn and Blau on gender wage gap
 future prospects, 855–858
 role of qualifications/discrimination,
 845–851
 role of wage structure, 851–852
 trends, 843–845, 852–855
 labor force participation and, 730–734
 and Petersen on discrimination, 780–785
 and Petersen on wage gap, 826–832
 and rational action theory of educational
 differentials, 610–621
 and Reskin on employment discrimination,
 770–777
 and Reskin on occupational sex composition,
 802–810

- Gender inequality (*continued*)
 and Tam on gender wage gap, 838–842
 and women opting-out from the workforce, 735–738
 workplace penalties for motherhood
 experiment/results, 761–769
 motherhood as status characteristic, 761
 productivity/discrimination, 760
 status characteristics theory, 760–761
 wage penalties, 759–760
- Geneology of Morals* (Nietzsche), 120
- German Democratic Republic, socialism and, 35–37
- Giddens, Anthony, 132–142
 on power elite, 285–289
 on social justice/social divisions, 1016–1022
- Gintis, Herbert, on effect of family/schooling in America, 498–504
- Globalization and income inequality, 24, 64–65
 comparative poverty and, 327–338
 Firebaugh on, 1044–1055
 Stiglitz on, 1036–1043
See also Income inequality; Inequality; Poverty
- Goldin, Claudia, on gender discrimination, 747–758
- Goldthorpe, John H.
Constant Flux: A Study of Class Mobility in Industrial Societies, 465, 476
 on occupational grading/prestige, 195–201
 on rational action theory of educational differentials, 610–621
 social class schema of, 467
 on social stratification of arts, 916–925
- Gouldner, Alvin, 110n13
 “Future of Intellectuals and the Rise of the New Class,” 295–303
- Granovetter, Mark S., on strength of weak ties, 576–579
- Great Depression, effect on top income, 70–71
- Groveland (Chicago) and race/ethnicity, 722–724
- Growth and income inequality, 56–57
- Grusky, David B.
 “Are There Big Social Classes?,” 165–175
 and inequality, 2–27
 on labor market gradation, 249–264
 on sex segregation, 812–824
- Haller, Archibald O., on educational/early occupational attainment process, 516–526
- Hankiss, Elemér, 311
- Harbison, Frederick H., on industrialism, 954–965
- Harding, David J., on family background/income in adulthood, 505–514
- Hargittai, Eszter, on digital reproduction of inequality, 936–944
- Hauser, Robert M.
 on occupational achievement, 202–204
 on occupational mobility, 425–435
 “Socioeconomic Indexes for Occupations,” 213–217
- Hays, Sharon, on women and welfare, 407–416
- Health and income inequality
 Mullahy, Robert and Wolfe on, 904–912
- Heyns, Barbara, on effect of family/schooling in America, 498–504
- Hierarchy, power elite and, 289. *See also* Class structure; Stratification/stratification systems
- Hiring practices, inequality of opportunity and, 19–20
- Historical perspective on top incomes, 67–72
- Hochschild, Arlie Russell, on labor force participation, 730–734
- Hodge, Robert W., on occupational status measurement, 205–217
- Honecker, Erich, 34–35
- Honorific assets, and inequality, 6
- Hope and inequality, 528–545
- Hope, Keith, on occupational grading/prestige, 195–201
- Horticultural/agrarian stratification, 13–14, 16, 27n11
- Hout, Michael, 49–53
 on realignment of U.S. presidential voting, 945–951
- Human assets, and inequality, 6
- Hunting-gathering society, 13
- Ideology, Marx on class and, 89–90
- Immigrant enclave, 664–667
 Japanese on West Coast, 651–652
 Jews in Manhattan, 649–651
 modes of incorporation, 648–649
 and process of immigrant adaptation to new society, 646–648
 typology of process of incorporation, 652–656
- Immigration
 and Portes on segmented assimilation, 658–667

- Waters on West Indian immigrants, 669–672
See also Immigrant enclave
- Incarceration
 Pager on finding work after, 683–689
 Western on, 388–392
- Incentive systems. *See* Motivation
- Income inequality and earnings, 24–25
 class conflict and, 97
 closure/earnings, 182–183
 data/methods, data processing/variances, 255–257
 data/variables, 179–182
 and economic inequality
 illustration of contrasts, 237–240
 interpersonal utility comparisons, 240–242
 money distribution, 235
 need for analysis, 235–237
 quality of life/functionings/capabilities, 242–245
 weights/valuations/explicitness, 245–246
 and England on gender wage gap, 834–837
 family background/income in adulthood
 changes in background/inheritance, 509–511
 genetic advantages, 505–507
 gifts/bequests, 507–508
 human capital, 507
 impact of children's education, 508–509
 taste for market goods, 507
- Firebaugh on globalization
 inequality across nations, 1049–1055
 myths of trade protest model, 1044–1048
 reversal of uneven economic growth, 1048–1049
- future of inequality, 65–66
- increase
 data, 61–63, 66n3
 demand-side accounts, 63–64
 institutional accounts, 64–65
 supply-side accounts, 63
- and Kahn and Blau on gender wage gap, 843–858
- MacLeod on aspirations in low-income neighborhood, 528–545
- methods/models, 179
- Mullahy, Robert and Wolfe on health and, 904–912
- occupational closure
 closure mechanisms/strategies, 177–179
 closure theory, 176–177
- and Petersen on gender wage gap, 826–832
 and poverty in America, 311–316
- Stiglitz on, 1036–1043
 and Tam on gender wage gap, 838–842
- top incomes
 construction of new database (historically), 67–68
 analysis, 70–72
 shares results, 68–70
- trend sources
 between-class (BC) component, 250–253, 258–264
 between-occupation/within-class (BO/WC) component, 250–251, 253–254, 258–264
 within-occupation variance (WO), 250–251, 254–255, 258–264
- and welfare's effect on children, 403–406
 and workplace penalties for motherhood, 759–768
See also Labor
- Income inequality hypothesis, 905
- Income mobility, 481–484
- Income pooling, income inequality and, 66
- Income tax, top incomes and, 68
- Individuals in society
 and evolution of modern stratification systems, 1012
 exclusion and, 155
 and inequality due to nature, 51
 power in socializing young, 46–47
 stratification, functionally important positions, 43
 and women on welfare, 407–416
- Industrial society stratification, 13, 16–17
- industrialism/industrial man
 large scale of society, 957–959
 pluralistic industrialism, 960–965
 society's consensus, 959–960
 work force, 954–957
- postindustrial economies' social foundations
 cost-disease/service expansion, 984–986
 globalization/technological change, 979–980
 microfoundations of employment, 986–988
 pessimism, 981–984
 problems of new service economy, 980–981
 service occupations, 981
 social democracy, 990–992
 welfare regimes, 988–990
- post-industrialist society
 components, 967–969
 meritocracy, 975–977

- Industrial society stratification (*continued*)
- power, 969–975
 - social structure changes, 966–967
 - postsocialist, 994–998
 - post-war European experience
 - absolute mobility rates, 443–453
 - class mobility trends, 437–441
 - relative mobility rates, 453–456
 - relevance, 441–443
- Inequality
- allocation analysis and, 12–13
 - approaches to studying
 - contemporary analyses, 23–24
 - income inequality, 24–25
 - intransigence of poverty/inequality, 22–23
 - multidimensionalism/new inequalities, 22
 - post-war, 21
 - assets/examples of, 4–5
 - consequences of, 4
 - examples of ideal-typical systems, 13–14
 - future of, 4
 - Hargittai on digital reproduction of, 936–944
 - history of, 13–18
 - impact on growth, 71–72
 - incarceration and, 388–392
 - in income
 - data, 61–63, 66n3
 - demand side accounts, 63–64
 - future, 65–66
 - institutional accounts, 64–65
 - supply side accounts, 63
 - as man-made in America
 - intentional policies, 50–51
 - political environment, 50, 52–53n6
 - social environment, 49–50
 - Marxism, functionalists' inequality views, 33
 - measuring, 6
 - minimum wage and, 24
 - nonempirical interpretations of growing concern
 - “Moral Credentials” account, 25–26
 - Newsworthiness account, 25
 - “Obsession with Small Differences” account, 26
 - of opportunity, trend in, 19–20
 - principles of stratification
 - functional necessity of stratification, 30–31
 - positional rank, 31–32
 - as providing opportunities for economic growth, 52
 - reasons to reduce in America
 - education/training as solution, 58–59
 - efficient policy changes, 56
 - enlightened self-interest, 55
 - growth/income inequality, 56–57
 - health/income inequality, 57
 - importance of balance of goals, 58
 - inefficiency in superstar markets, 57
 - influence from money, 56
 - low wages/imperfect monitoring/public safety, 55
 - market failure, 55–56
 - philosophy, 54
 - public preference, 57–58
 - religion, 54–55
 - socialism, and worker motivation, 35–37
 - sources of concern for, 2
 - stratification systems, and dysfunctions, 47–48
 - top incomes
 - construction of new database, 67–68
 - shares results, 68–70
 - trends in, 18–19
 - working class stratification and, 17
- Information age
- and Hargittai on digital reproduction of, 936–944
 - working class stratification and, 17
- Information and communication technologies (ICT), 936–944
- Institutional accounts, income inequality and, 64–65
- Intellectuals, Gouldner on future of, 295–303
- Intergenerational inheritance, 510–512
- International Monetary Fund (IMF), 1036–1043. *See also* Economic inequalities
- International perspective on top incomes, 67–72
- International Standard Classification of Occupations (ISCO)*, 192
- Jacobs, Jerry A., on sex segregation, 797–801
- Jankowski, Martín Sánchez, 49–53
- Jencks, Christopher
 - on effect of family/schooling in America, 498–504
 - on family background/income in adulthood, 505–514
- Jobs-First, 402
- Kahn, Lawrence M., on gender wage gap, 843–858
- Kalleberg, Arne L.
 - on matching persons to jobs, 553–561

- “Nonstandard Employment Relations and Labour Market Inequality,” 562–573
- Kerr, Clark, on industrialism, 954–965
- Kim, Young-Mi, on labor market gradation, 249–264
- Kohn, Melvin, on job complexity/adult personality, 893–902
- Krueger, Alan B., 53–60
on poverty and terrorism, 1056–1062
- Ku, Manwai C., and inequality, 2–27
- Labor
- alienation of worker, 74–78
 - and aspirations in low-income neighborhood, 528–545
 - bourgeoisie and, 80–83
 - decomposition of, 92–93
 - demand-side accounts and, 63–64
 - division of, 159–164
 - dynamics of split markets of, 638–640
 - and England on gender wage gap, 834–837
 - ethnic antagonism and, 633–634
 - expensive labor group, 640–643
 - factors affecting initial price of, 634–637
 - and Fernandez on networks/race/hiring and, 587–595
 - and future of income inequality, 65–66
 - gender gap, and full vs. part-time employment, 858–859n4
 - and Goldin on gender discrimination in performing arts, 747–758
 - Hochschild on labor force participation, 730–734
 - incarceration and, 388–392
 - industrialism/industrial man
 - large scale of society, 957–959
 - pluralistic industrialism, 960–965
 - society’s consensus, 959–960
 - work force, 954–957
 - institutional accounts and, 64–65
 - and Jacobs on women’s careers, 797–801
 - jobless poverty and, 340–349
 - and Kahn and Blau on gender wage gap, 843–858
 - and Kohn on job complexity/adult personality, 893–902
 - market discrimination, 673–677
 - Marx on class and, 89–90
 - matching persons to jobs
 - basic concepts (Weber’s), 555–556
 - economic theory, 553–555
 - growth in earnings constraints, 560–561
 - neoclassical theory of earnings
 - determination, 556–558
 - vacancy competition, 558–560
 - nonstandard employment relations
 - consequences for labour market outcomes, 569–573
 - cross-national patterns, 566–569
 - standard/nonstandard work arrangements, 562–566
 - occupational mobility and, 429–433
 - Pager on finding work after incarceration, 683–689
 - and Petersen on discrimination, 780–785
 - and Petersen on gender wage gap, 826–832
 - Piore on dual market, 550–553
 - price of ethnicity and, 637–638
 - as reason to reduce inequality, 55
 - and Reskin on employment discrimination, 770–777
 - and Reskin on occupational sex composition, 802–810
 - socialism, and worker motivation, 35–37
 - supply side accounts and, 63
 - and Tam on gender wage gap, 838–842
 - wage disparities in America, 502
 - and women opting-out from the workforce, 735–746
 - workplace penalties for motherhood
 - experiment/results, 761–769
 - motherhood as status characteristic, 761
 - productivity/discrimination, 760
 - status characteristics theory, 760–761
 - wage penalties, 759–760
- Labor force participation rates (LFPR) of women, 740–746
- Labor market gradations
- data/methods
 - data processing, 255–257
 - results, 258–264
 - trend sources
 - between-class (BC) component, 252–253
 - between-occupation/within-class (BO/WC) component, 253–254
 - within-occupation component (WO), 254–255
- Labor-transfer approach and exploitation, 101–102
- Lareau, Annette, on class/race/family life, 926–936
- Lau, Christoph, on meta-change of modern society, 1030–1035
- Lenski, Gerhard, 33–40

- Levanon, Asaf, on sex segregation, 812–824
- Lewis, Oscar, 352
- Leisure class
 - conspicuous consumption, 865–869
 - conspicuous leisure, 864–865
 - pecuniary canons of taste, 869–870
 - pecuniary emulation, 862–864
- Lin, Nan, on social networks/status attainment, 580–582
- Lopoo, Leonard M., on family
 - background/income in adulthood, 505–514
- Low-income, MacLeod on aspirations in neighborhood of, 528–545
- Lucas, Samuel R., 49–53
- Luxembourg Income Study (LIS), 327–338
- MacLeod, Jay, on aspirations in low-income neighborhood, 528–545
- Macro-level externalities account, poverty/inequality and, 2–3
- Managerialism
 - post-communist, 311–316
 - and poverty in America, 311–316
- Market economies
 - inequality of opportunity and, 19–20
 - pre-capitalism/capitalism and, 79–80
 - as reason to reduce inequality, 55–57
- Marx, Karl
 - “Alienation and Social Classes,” 74–78
 - “Classes in Capitalism and Pre-capitalism,” 79–88
 - “Ideology and Class,” 89–90
- Marxism
 - alienation of worker
 - ideology, 89–90
 - pre-capitalism/capitalism, 79–88
 - social classes, 74–78
 - class theory and, 143–158
 - functionalists’ inequality views and, 33
 - income inequality and, 24
 - motivation and, 35–37
 - power elite and, 281, 285–289
 - social classes, and associational change, 171–172
 - socialist countries, and economic inequality, 33–39
 - working class stratification and, 17
- Massey, Douglas S., on American apartheid, 349–359
- Mayer, Susan E., on family background/income in adulthood, 505–514
- Meritocracy, post-industrialist society and, 975–977
- Merkel, Wolfgang, 1021
- Meyer, John W., on modern stratification systems, 1006–1015
- Michelson, Stephen, on effect of family/schooling in America, 498–504
- Micro-level externalities account, poverty/inequality and, 2
- Middle class, in contemporary society, 108–109
- Mills, C. Wright, on power elite, 275–285
- Minimum wage, income inequality and, 24
- Mobility
 - in Europe
 - absolute mobility/class structure, 469–473
 - data/methods, 465–469
 - explanations/conclusions, 478–480
 - previous research, 476–478
 - social fluidity, 473–476, 479
 - income, 481–484
 - industrialism/industrial man and, 955–956
 - nonstandard employment relations and, 572
 - occupational, 425–435
 - post-war European experience
 - absolute rates, 443–453
 - class mobility trends, 437–441
 - relative rates, 453–456
 - relevance, 441–443
- Modernity, Beck and Lau on meta-change of modern society, 1030–1035
- Moore, Wilbert E., 30–33
- “Moral Credentials” account of inequality, 25–26
- Morenoff, Jeffrey D., on social processes in Chicago, 371–385
- Morgan, Stephen L., on new social psychological model of attainment, 542–549
- Morris, Martina, 61–66
- Morris, Pamela, on children and welfare policies, 400–406
- Mosca, Gaetano, on the ruling class, 268–274
- Motivation
 - necessity of differential inducement value, 45–46
 - socialism and, 35–37
 - successful, 38–39
- Moving to Opportunity (MTO), 366–368, 397–398
- Mullahy, John, on health/income/inequality, 904–912

- Mullainathan, Sendhil, on labor market discrimination, 673–677
- Music and gender discrimination, 747–758
- Myers, Charles A., on industrialism, 954–965
- Myrdal, Gunnar, 351
- National Evaluation of Welfare to Work Strategies (NEWWS), 402
- National Longitudinal Surveys (NLS), 483
- Nazism, power elite and, 281
- Nee, Victor, on postsocialist stratification, 994–998
- Neo-Marxism
 - class structure
 - collective action, 167–168
 - domination vs. exploitation, 99
 - polarization/nonpolarization of class, 98–99
 - postcapitalist society classes, 99–100
 - working class stratification and, 17
- Networks. *See* Social networks
- New inequalities account, poverty/inequality and, 2
- Newsworthiness account of inequality, 25
- Nomenklatura* position, 311–316
- “Obsession with Small Differences” account of inequality, 26
- Occupational Changes in a Generation (OCG) survey, 188
- Occupational closure, and income inequality
 - closure mechanisms/strategies, 177–179
 - closure theory, 176–177
 - closure/earnings, 182–183
 - data/variables, 179–182
 - methods/models, 179
- Occupational mobility, 425–435
- Occupations
 - achievement as prestige or socioeconomic, 202–204
 - attainment process and education, 516–526
 - as contextual variable, 209–211
 - Duncan’s SEI measurement scale, 205–209
 - and family background/income in adulthood, 514
 - industrialism/industrial man
 - large scale of society, 957–959
 - pluralistic industrialism, 960–965
 - society’s consensus, 959–960
 - work force, 954–957
 - and Kohn on job complexity/adult personality, 893–902
 - and meaning of prestige, 195–196
 - measuring status of, 187–190
 - postindustrial economies and, 981
 - prestige of, 191–194
 - prestige ratings and, 196–201
 - prestige/socioeconomic indicators, 211–212
 - and Reskin on occupational sex composition, 802–810
 - socioeconomic indexes for, 213–217
 - See also* Industrial society stratification
- Oliver, Melvin L., on racial inequality, 709–715
- Open/closed relationships, Weber on
 - economic, 130–131
 - social, 128–130
- Oppression, class structure and, 87
- Opting-out. *See* Labor
- Orchestras and gender discrimination, 747–758
- Organization for Economic Cooperation and Development (OECD), 1037
- Ownership and decomposition of capitalism, 91–92
- Pager, Devah, “Race, Crime, and Finding Work in an Era of Mass Incarceration,” 683–689
- Pakulski, Jan, on death of class, 1022–1029
- Palestinian Center for Policy and Survey Research (PCPSR), 1058–1060
- Panel Study of Income Dynamics (PSID), 58, 483
- Party, Weber on, 123–124
- Pay equity, 834
- Pebly, Anne R., “Neighborhoods, Poverty, and Children’s Well-Being,” 360–368
- Pecuniary canons of taste, leisure class and, 869–870
- Pecuniary emulation, leisure class and, 862–864
- Persistence account, poverty/inequality and, 2
- Personal Responsibility and Work Opportunity Reconciliation Act, 408
- Personality, and Kohn on job complexity and, 893–902
- Personnel
 - stratification
 - differential functional importance, 31–32n3
 - differential scarcity/variations, 32
- Petersen, Trond
 - on discrimination, 780–785
 - on gender wage gap, 826–832
- Physical assets, and inequality, 6
- Piketty, Thomas, 67–72
- Piore, Michael J., on dual labor market, 550–553
- Pluralistic industrialism, 960–965
- Poland, socialism and, 33–39

- Political environment
 Domhoff on power and, 290–295
 and Giddens on power elite, 285–289
 and Gouldner on rise of a New Class, 296
 inequality of opportunity and, 20
 and Mills on power elite, 282
 and Mosca on ruling class, 268–274
 power elite and, 276–278
 socialist countries, and Marxism, 34
- Portes, Alejandro
 on educational/early occupational attainment process, 516–526
 on immigrant enclave, 646–656
 on segmented assimilation, 658–667
- Positional rank and inequality, 31–32
- Possessing class, self-alienation and, 78
- Post-communism, and capitalism as built by non-capitalists, 1000–1005
- Post-communist managerialism, 311–316
- Post-Fordism, sociotechnical change and, 169–171
- Postindustrial economies' social foundations
 cost-disease/service expansion and, 984–986
 and Giddens on social justice, 1016–1017
 globalization/technological change and, 979–980
 microfoundations of employment and, 986–988
 pessimism and, 981–984
 problems of new service economy and, 980–981
 service occupations and, 981
 social democracy and, 990–992
 welfare regimes and, 988–990
- Post-industrialist society
 components of, 967–969
 meritocracy and, 975–977
 power of, 969–975
 and social structure changes, 966–967
- Postmodernist Phase of social class measurement, 11–12
- Postsocialist stratification, 994–998
- Poverty
 adjusted disposable income (ADPI)/unadjusted household income (DPI), 337–338n4
 and Beck and Lau on meta-change of modern society, 1033–1035
 Chicago neighborhoods
 Community Survey of the Project on Human Development in Chicago Neighborhoods (PHDCN), 380–384
 consequences of inequality, 379–380
 geographical/racial change, 371–376, 385n2, 386n6
 predictors of change, 377–379
 dual labor market and, 551–553
 and effect of family/schooling in America, 498–504
 Ehrenreich on, 317–326
 fields of inquiry for, 3–4
 neighborhoods/children's well-being and, 360–368
 Smeeding on, 327–338
 sources of concern for, 2
 stratification process and, 495–497
 terrorism and, 1056–1062
 Weber on, 95
 Wilson on jobless, 340–349
See also Income inequality; Inequality
- Poverty hypothesis, 904
- Power
 class conflicts and, 94–97
 Domhoff on politics and, 290–295
 economically determined, Weber on, 114
 and Giddens on power elite, 285–289
 and Gouldner on rise of a New Class, 298
 Mills on power elite, 275–285
nomenklatura position and, 311–316
 of old over young, 46–47
 in post-industrialist society, 969–975
 power elite and, 284–285n3
 and prestige of occupations, 192–193
 as reason to reduce inequality, 56
 ruling class and, 268–274
- Power assets, and inequality, 6
- Presidential voting realignment, 945–951
- Prestige ratings
 occupational achievement and, 202–204
 occupations and, 195–201
 and socioeconomic measures, 211–212
- Private property
 alienated labor and, 78
 capitalism and, 17–18
 post-communist managerialism and, 315–316
 ruling class and, 270–271
 socialist countries, and Marxism, 34, 36–37
 stratification systems and, 15
- Production vs. act of production, and alienation of worker, 75–78
- Proletarian class
 and class conflict in world economy, 113
 labor and development of, 84–87
 self-alienation and, 78

- Race/ethnicity
- American apartheid, 349–359
 - and Bonacich on ethnic antagonism, 632–644
 - in Chicago, 716–726
 - and family background/income in adulthood, 514
 - and Feagin on antiblack discrimination, 703–709
 - and Fernandez on networks/hiring and, 587–595
 - incarceration and, 391–392
 - jobless poverty and, 343–348
 - labor market discrimination, 673–677
 - Lareau on class/race/family life
 - children's time use, 928–930
 - language use, 930–932
 - parental intervention institutions, 932–935
 - study participants/observations, 927–928
 - and Oliver on racial inequality, 709–715
 - Pager on finding work after incarceration, 683–689
 - and Portes on segmented assimilation, 658–667
 - poverty in Chicago neighborhoods
 - Community Survey of the Project on Human Development in Chicago Neighborhoods (PHDCN), 380–384
 - consequences of inequality, 379–380
 - geographical/racial change, 371–376, 385n2, 386n6
 - predictors of change, 377–379
 - racial identities in 2000
 - governmental (OMB) decision on measuring race, 625–627
 - maximum/minimum counts of racial categories, 629–630
 - multiple-race response option, 624–625, 627–629
 - and Reskin on employment discrimination, 770–777
 - and Reskin on occupational sex composition, 802–810
 - stereotype threat and African-American student achievement, 678–682
 - stratification of, 4
 - Wilson on declining significance of, 691–702
 - See also* Immigrant enclave
- Rank and equality, 31–32
- Rational action theory, educational differentials
- empirical generalizations, 617–620
 - generation of class differentials, 614–617
 - history/generalities, 610–612
 - model, 612–614
 - model implications, 620–621
- Realignment in U.S. presidential voting, 945–951
- Really existing socialism, 34–39, 39n2
- Really existing socialist societies, 34–39, 39n2
- Reductionism, Structuralist Phase of social class measurement and, 9
- Relationships (open/closed), Weber on
- economic, 130–131
 - social, 128–130
- Relative deprivation hypothesis, 905
- Relative income hypothesis, 904–905
- Religion
- as reason to reduce inequality, 54–55
 - ruling class and, 271–272
- Rent. *See* Class analysis, rent-based
- Residential segregation
- American apartheid and, 351–352
 - Cultural Phase of social class measurement and, 9–10
 - MacLeod on aspirations in low-income neighborhood, 528–545
 - poverty/children's well-being and, 360–368
- Reskin, Barbara F.
- on employment discrimination, 770–777
 - on occupational sex composition, 802–810
- Rewards in society
- analysis of stratification systems
 - institutionalized social inequality, 46–47
 - scarce/desired goods and power/esteem as positively functional, 47–48, 49n2
 - political environment in America and, 50, 52–53n6
 - socialism and, 33–34
- Rigidity, 13
- stratification and, 43
- Rituals, 119
- Robert, Stephanie, on health/income/inequality, 904–912
- Roemer on class/exploitation
- assets/exploitation/classes, 105–108
 - exploitation concept, 100–101
 - game-theory approach, 102–103
 - general class analysis, 103–105
 - labor-transfer approach, 101–102
 - middle class in contemporary society, 108–109
 - in terms of domination within production, 103
- Rosenbaum, James E., and Chicago Gautreaux Housing Program, 393–399

- Ruling class, 268–274
 Domhoff on power and, 290–295
- Saez, Emmanuel, 67–72
- Sampson, Robert J., on social processes in Chicago, 371–385
- Sastry, Narayan, “Neighborhoods, Poverty, and Children’s Well-Being,” 360–368
- Secularization, and Gouldner on future of intellectuals, 296
- Segmented assimilation, 658–667
- Segregation, status, 119–120. *See also* Gender inequality; Race/ethnicity
- Self-alienation, 75–76
- Sen, Amartya K., “From Income Inequality to Economic Inequality,” 235–247
- Sewell, William H., on educational/early occupational attainment process, 516–526
- Siblings, social ascription and, 596–608
- Slavery stratification, 13, 16
- Smeeding, Timothy M., on poverty in America comparatively, 327–338
- Smith, Marshall, on effect of family/schooling in America, 498–504
- Social ascription, sibling correlations and, 596–608
- Social assets, and inequality, 6
- Social capital
 and Burt on structural holes, 583–585
 and Fernandez on networks/race/hiring and, 587–595
 and Granovetter on strength of weak ties, 576–579
- Social classes
 and aspirations in low-income neighborhood, 528–545
 Beck and Lau on meta-change of modern society, 1030–1035
- disaggregation
 class identification, 166–167
 class outcomes, 168
 collective action, 167–168
 social closure, 167
- and Giddens on social justice, 1016–1022
- inequality trend and, 18–19
- issues of trend
 associational change, 171–172
 site of production, 168–169
 sociotechnical change, 169–171
- Marx on alienation and, 74–78
- post-industrialist society
 components, 967–969
 meritocracy, 975–977
 power, 969–975
 social structure changes, 966–967
 social clubs, 293–294
- Social control and sex segregation, 797–801
- Social fluidity, 473–476, 479
- Social inclusion account, poverty/inequality and, 2–3
- Social justice
 Giddens on, 1016–1022
 globalization and, 1043
 as reason to reduce inequality, 54
- Social networks
 and Burt on structural holes, 583–585
 and Fernandez on race/hiring and, 587–595
 and Granovetter on strength of weak ties, 576–579
 Lin on status attainment and, 580–582
- Socialism
 and capitalism as built by non-capitalists, 1000–1005
 Marxism, countries governed by, 33–39
 worker motivation and, 35–37
- Society, public preference as reason to reduce inequality, 57–58
- Socioeconomic order, and rise of a New Class, 295–303
- Socioeconomic ratings
 indexes for occupations
 conceptual issues, 214–215
 occupational prestige and, 215–216
 structural models, 216–217
 occupational achievement and, 202–204
- Sociological realism, 1007
- Sociotechnical change and social classes, 169–171
- Solon, Gary, on intergenerational income mobility, 481–484
- “Some Principles of Stratification” (Davis and Moore), 41
- Somoza, Anastasio, 34
- Sørensen, Aage B., on matching persons to jobs, 553–561
- Sørensen, Jesper B.
 “Are There Big Social Classes?,” 165–175
 on rent-based class analysis, 219–234
- Soviet republics, socialism and, 33–39
- Spatial dynamics of neighborhood poverty, 377–379
- Standard Occupational Classification (SOC), 179–180
- Staniszki, Jadwiga, 311
- State socialism stratification, 13

- Status groups/classes
 occupations and, 187–190
 sibling correlations and, 596–608
 social networks and, 580–582
 Weber on
 acquisition, 126–127
 class concepts/situation, 124–125
 property classes, 125–126
 social stratification, 127–128
- Status/class/party, Weber on
 communal action from class interest, 116–117
 determination of class by market, 114–116
 economically determined, 114
 effects of status stratification on economic conditions, 121–123
 ethnic segregation/caste, 119–121
 parties/power, 123–124
 status honor, 118–119
 status privileges, 121
 status stratification guarantees, 119
 types of class struggle, 117–118
- Steele, Claude, on stereotype threat and African-American student achievement, 678–682
- Stereotype threat and African-American student achievement, 678–682. *See also* Race/ethnicity
- Stereotyping
 employment discrimination and, 773–776
 sex segregation and, 791–792
See also Race/ethnicity
- Stiglitz, Joseph E., on globalism's discontents, 1036–1043
- Stratification/stratification systems, 26n3, 26n4
 analysis
 certain societal positions are more functionally important, 42–43
 classification of prerequisites of positions, 46
 conversion of talents into skills, 44–45
 few individuals possess talents trainable into important skills, 43–44
 institutionalized social inequality, 46–47
 necessity of differential inducement value, 45–46
 scarce/desired goods and power/esteem as positively functional, 47–48
 and Chan and Goldthorpe on arts, 916–925
 and changing Black American experiences, 700–701
 and death of class, 1022–1029
 and decomposition of labor, 92–93
 dysfunctions, 47–48
 examples of ideal-typical systems, 13–14
 functional importance of, 31–32n3
 inequality and, 3–5
 as inherently antagonistic, 43
 Marx on, 88
 and measuring status of occupations, 187–190
 Meyer on evolution of modern, 1006–1015
 positional rank, and differential functional importance, 31–32
 postsocialist, 994–998
 power elite and, 278–280
 principles
 functional necessity of stratification, 30–31
 positional rank, 31–32
 process
 basic model, 487–491
 conceptual issues, 486–487
 path coefficients, 491–495
 self-perpetuation, 495–497
 term origin, 26n1
 theatre/dance/cinema attendance
 analysis of data, 916–917
 conclusions, 924–925
 data/concepts, 917–919
 latent class analysis of attendance, 919–921
 types of musical consumer, 921–924
See also Class structure
- Structural holes, 583–585
- Structuralist Phase of social class measurement, 7–9
- Style of life, 119
- Suicide bombers, 1056–1062
- Supply side accounts, income inequality and, 63
- Swidler, Ann, 49–53
- Szalai, Erzsébet, 311
- Szelényi, Iván
 on capitalism, 1000–1005
 on post-communist managerialism, 311–316
- Takeoff account, 2
- Tam, Tony, on gender wage gap, 838–842
- Terrorism and poverty, 1056–1062
- Theories of class. *See* Class theories
- Townsley, Eleanor
 on capitalism, 1000–1005
 on post-communist managerialism, 311–316
- Trade agreements, as reason to reduce inequality, 56
- Treiman, Donald J., on occupational prestige, 191–194
- Tribalism stratification, 13–14, 26n2
- Trickle-down-economics, 1041–1043. *See also* Economic inequalities

- Truly Disadvantaged, The* (Wilson), 353
- Tumin, Melvin M., 41–49
- Turner, Ralph H., on class mobility, 420–425
- Tyree, Andrea
 on status of occupations, 187–190
 on stratification process, 486–497
- Underclass, American apartheid and, 349–359
- Unemployment, Western on, 388–392
- Unions
 and future of income inequality, 65
 income inequality and, 24
See also Labor
- Universality, class and, 90
- Upper class, Brooks on new, 304–310
- Veblen, Thorstein, on leisure class, 862–870
- Visibility account, poverty/inequality and, 2
- Voss, Kim, 49–53
- Wage competition model, 557
- Wallerstein, Immanuel, “Class Conflict in the Capitalist World Economy,” 111–113
- War, effect on top income, 69–71, 72n1, 72n3
- Warren, John Robert, “Socioeconomic Indexes for Occupations,” 213–217
- Waters, Malcolm, on death of class, 1022–1029
- Waters, Mary C., on West Indian immigrants, 669–672
- Wealth
 and Beck and Lau on meta-change of modern society, 1033–1035
 class analysis and, 219–223
 and class as life conditions, 223–224
 and exploitation classes/collective action in modern capitalism, 229–232
 power elite and, 278
 rent/exploitation and, 224–229
- Weber, Max
 “Class, Status, Party,” 114–124
 open/closed relationships, 555–556
 on power, 95
 “Status, Groups and Classes,” 128–132
- Weeden, Kim A.
 on labor market gradation, 249–264
 “Occupational Closure and Earnings Inequality,” 176–186
- Welfare policies
 effect on children/adolescents, 400–406
 Hays on women and, 407–416
 postindustrial economies and, 988–990
- West Indian immigrants, 669–672
- Western, Bruce, 61–66
 “Incarceration, Unemployment, and Inequality,” 388–392
- When Work Disappears: The World of the New Urban Poor* (Wilson), 340–341
- Wilson, William Julius
 on declining significance of race, 691–702
 on jobless poverty, 340–349
 on race/ethnicity in Chicago, 716–726
- Wolfe, Barbara, on health/income/inequality, 904–912
- Working class culture
 downward mobility and, 10
 oppression and, 87
See also Alienation
- Workplace segregation, Cultural Phase of social class measurement and, 9–10
- Wright, Eric Olin, 98–110
- Zhikov, Todor, 34–35

“[This text] reflects a clear vision of the scientific and intellectual history of the field, but it will also shape research and scholarship for years to come. For researchers, teachers, and students, *Social Stratification* will be an indispensable resource.”

—Robert M. Hauser, University of Wisconsin–Madison

“*Social Stratification* is the definitive reader in the field ... a must-read.” —Barbara Reskin, University of Washington

The study of poverty and inequality has been thrust into the foreground as scholars, politicians, and policymakers respond to the spectacular increase in economic inequality and the slowing, stalling out, or even reversal of long-standing downward trends in other forms of inequality. A mainstay of the field for more than a decade, *Social Stratification* has now been fully updated and revised with additional readings and newly commissioned pieces, all from top scholars of poverty and inequality. This collection reflects ongoing changes in the structure of inequality and in the tools and concepts that have been used to understand these changes.

Oriented toward the advanced student, *Social Stratification* provides a demanding, comprehensive, no-holds-barred overview of classic and contemporary scholarship. The history of the field unfolds systematically from the early and modern classics to the cutting-edge scholarship that currently drives the field. The resulting collection, even more comprehensive and diverse than its predecessor, can be used as a stand-alone text for courses on stratification, poverty, and inequality, as well as occupations, labor markets, and social mobility.

DAVID B. GRUSKY is professor of sociology and director of the Center for the Study of Poverty and Inequality at Stanford University. He is the editor, with Szonja Szelényi, of *The Inequality Reader: Contemporary and Foundational Readings in Race, Class, and Gender* (Westview Press, 2006) and *Inequality: Classic Readings in Race, Class, and Gender* (Westview Press, 2006).



A MEMBER OF THE PERSEUS BOOKS GROUP
www.westviewpress.com
www.perseusacademic.com

ISBN-13: 978-0-8133-4373-0
ISBN-10: 0-8133-4373-9

