CSP · INGENIERÍA INDUSTRIAL UNIVERSIDAD DE CHILE
CENTRO DE SISTEMAS PÚBLICOS

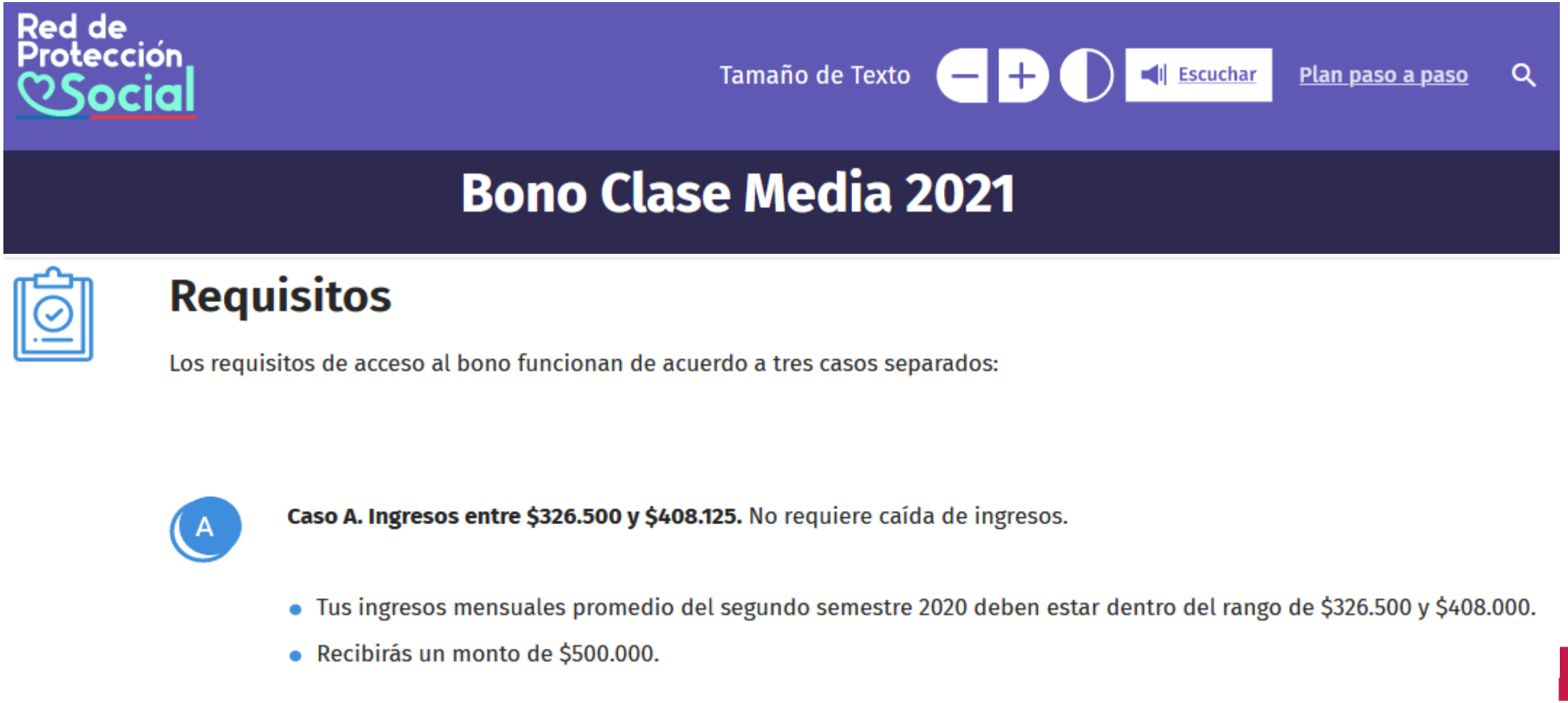
IN4402: Applied statistics for management and economics

Regression Discontinuity – Model estimation

ANDRÉS FERNÁNDEZ

www.sistemaspublicos.cl

INTRODUCTION I
CUASI AND NON EXPERIMENTAL METHODS



INTRODUCTION II
CUASI AND NON EXPERIMENTAL METHODS

- In other words:
  - If we call $Z^* = 408.000$ it's a *cut-off* value for participation

$$D_i = RecibeBono_i = \begin{cases} 1 \; if \; Ptje_i \leq 408.000 \\ 0 \; if \; Ptje_i > 408.000 \end{cases}$$

- The selection point is somewhat *arbitrary* (why not 407.000, 409.000?)
- Intuition says that subjects close to the cutoff *would be similar*

ASSUMPTIONS: DISCONTINUITY ON PARTICIPATION

CUASI AND NON EXPERIMENTAL METHODS

- (Sharp definition):
  - The discontinity comes from the probability to receive treatment (D_i):

ASSUMPTIONS: POTENTIAL OUTCOMES

CUASI AND NON EXPERIMENTAL METHODS

- We *expect potential outcomes* to be **continuous** in Z
  - $\lim_{z \uparrow Z^*} \mathbf{E}[Y_i|\mathbf{Z} = \mathbf{z}] = \lim_{z \downarrow Z^*} \mathbf{E}[Y_i|\mathbf{Z} = \mathbf{z}]$

ASSUMPTIONS: POTENTIAL OUTCOMES

CUASI AND NON EXPERIMENTAL METHODS

- Everything **but** the "running variable" is the same between groups.
- After controlling for the *running variable*, potential results should be **independent** of assignment*:*

$$Y_0, Y_1 \perp D_i \mid Z_i$$

- We expect that:
  - Potential results variable are ***continuous*** in the cut-off point.
  - Observables are ***continuous*** in the cut-off point

ESTIMATING ATE
CUASI AND NON EXPERIMENTAL METHODS

- So, given the assumptions the ATE would be:

$$ATE = \lim_{z \uparrow Z^*} \mathrm{E}[Y_i | Z_i = z] - \lim_{z \downarrow Z^*} \mathrm{E}[Y_i | Z_i = z]$$

- It is estimated by a regression:

- This effect
  - Is a local average for subjects close to the cut-off
  - **Internal validity**: closer to the cut-off the groups are more similar
  - **External validity**: closer to the cut-off we lose sample

CSP · INGENIERÍA INDUSTRIAL UNIVERSIDAD DE CHILE
CENTRO DE SISTEMAS PÚBLICOS

IN4402: Applied statistics for management and economics

Regression Discontinuity – Assumptions and Checks

ANDRÉS FERNÁNDEZ

www.sistemaspublicos.cl

ASSUMPTIONS: POTENTIAL OUTCOMES

CUASI AND NON EXPERIMENTAL METHODS

- We assume the cut-off point is somewhat *arbitrary*
- Everything **but** the "running variable" is the same between groups (near the cut-off).
- After controlling for the *running variable*, potential results should be **independent** of assignment*:*

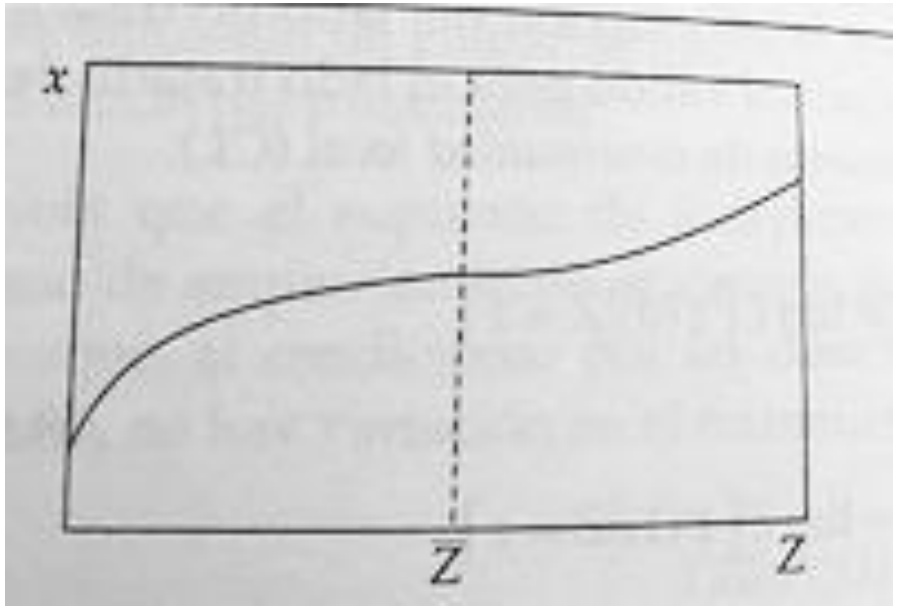
$$Y_0, Y_1 \perp D_i \mid Z_i$$

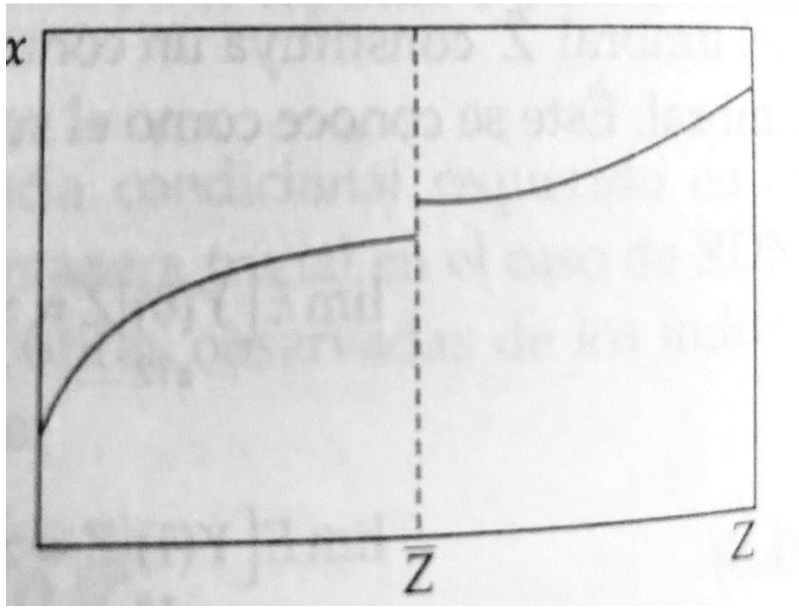
- And we expect that:
  - Potential results variable are ***continuous*** in the cut-off point.
  - Observables are ***continuous*** in the cut-off point

ASSUMPTIONS: CONTINUITY ON OBSERVABLES

CUASI AND NON EXPERIMENTAL METHODS

- In order to compare groups, we expect that $\boldsymbol{X}$ are **continuous** (balance)
- We can observe those variables and *check* their continuity





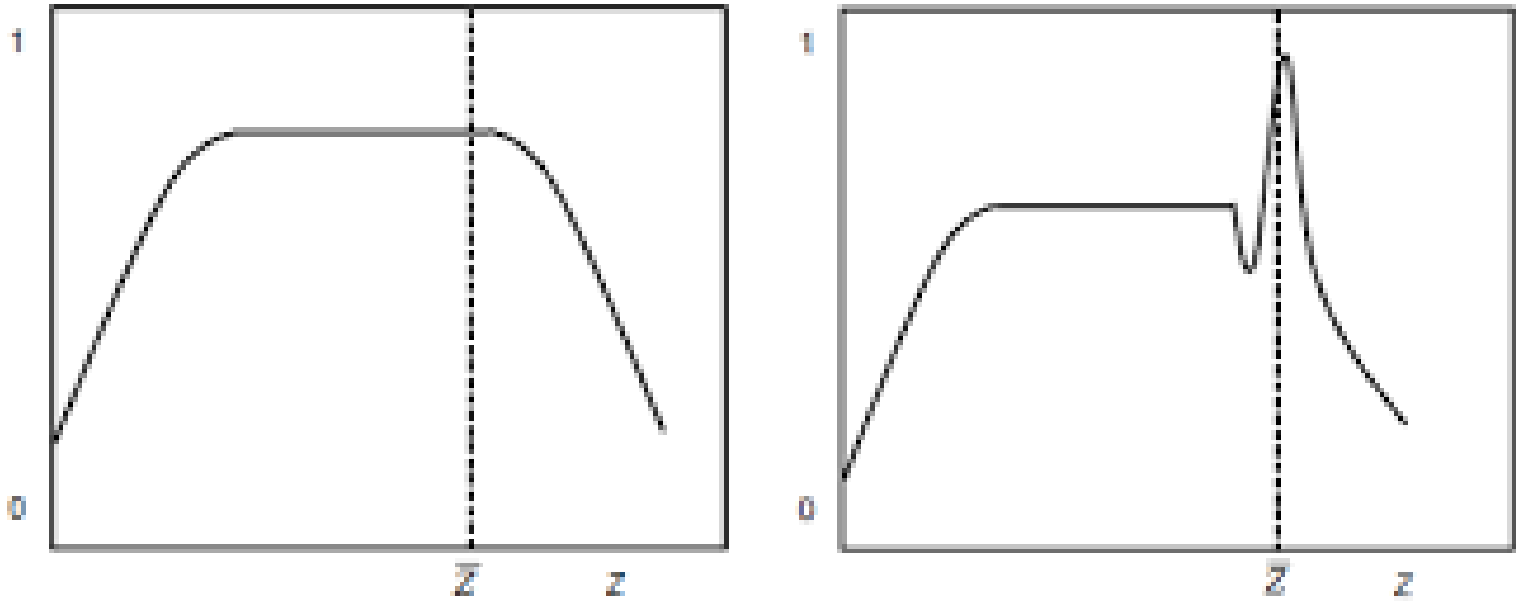
- Not continuous means cut-off is not "*random*" and groups won't be comparable in those variables

ASSUMPTIONS: NO MANIPULATION

CUASI AND NON EXPERIMENTAL METHODS

- Is important that the running variable cannot be **manipulated** by subjects
  - Otherwise a selection problem could arise.
- We check **density** of subjects by the running variable:
  - *Left* means continuous and without manipulation
  - *Right* means a “self-selection” to try to get the scores.

Density



ROBUSTNESS CHECKS
CUASI AND NON EXPERIMENTAL METHODS

- It is estimated by the following regression ($\rho = ATE$):

$$Y_i = \alpha + \beta Z_i + \boldsymbol{\rho} D_i + u_i$$

- Robustness Checks: local effects and functional forms
  - Use different values of $\delta > 0$ for **neighborhood local** effects ($Z^* - \delta < z < Z^* + \delta$)
  - Add **quadratic** (or other polynomial) terms
  - Add **interactions** for linear and quadratic terms

CHECKING ASSUMPTIONS
CUASI AND NON EXPERIMENTAL METHODS

- Because the assumptions is *only* the continuity, we can check for *flexible* relations

$$Y_i = f(Z_i) + \rho D_i + u_i$$

ROBUSTNESS CHECKS
CUASI AND NON EXPERIMENTAL METHODS

- Summary:
  - Regression Discontinuity can be used in situations where treatment is assigned by a *cut-off point in a running variable*
  - Some assumptions must hold:
    - Continuity in potential outcomes and conditional independence
    - Continuity in covariates
  - Some robustness checks should be made
    - Check polinomial relationships
    - Check different neighborhood sizes

CSP · INGENIERÍA INDUSTRIAL UNIVERSIDAD DE CHILE
CENTRO DE SISTEMAS PÚBLICOS

IN4402: Applied statistics for management and economics

Regression Discontinuity – Application

ANDRÉS FERNÁNDEZ

www.sistemaspublicos.cl

RD APPLICATION
CUASI AND NON EXPERIMENTAL METHODS

CSP

24 HORAS

HOME / NACIONAL

"Brigadas de menores": La controvertida idea del alcalde Lavín para fiscalizar venta de alcohol en Las Condes

20 años elmostrador

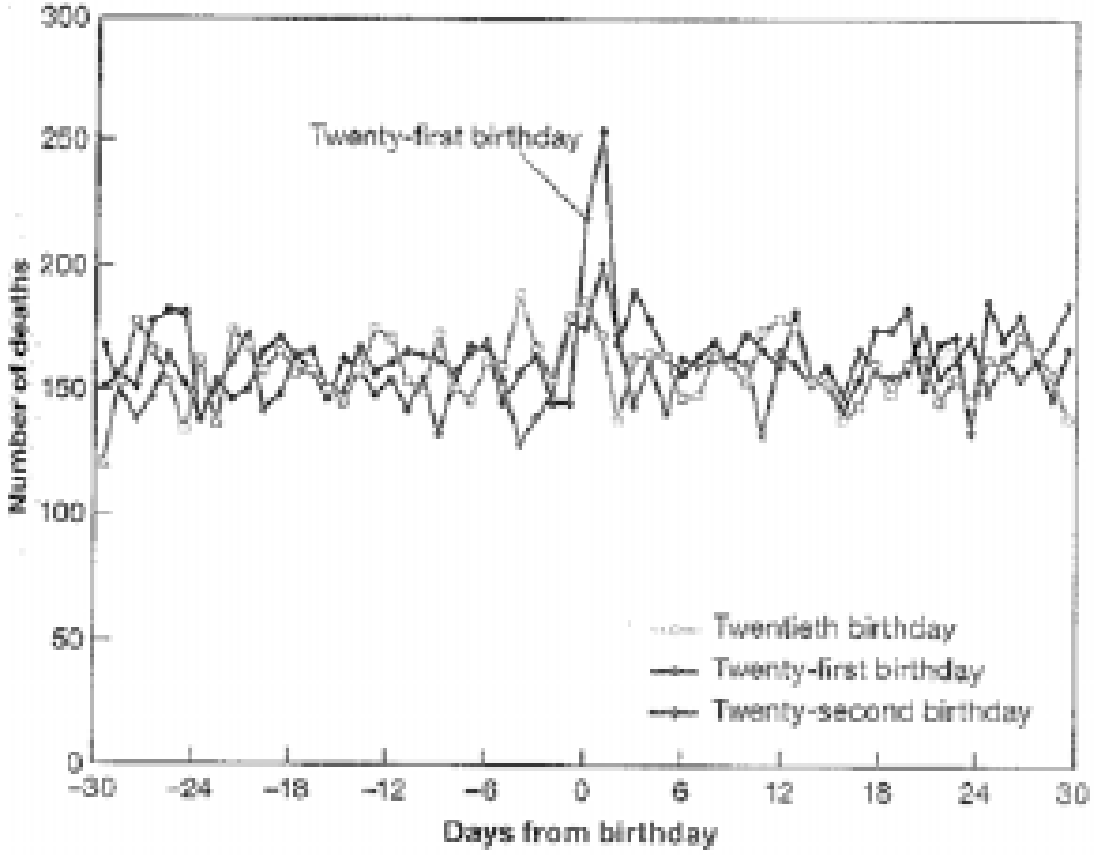
Noticias Mercados TV Cultura Generación M Agenda País Braga Avisos Legales Buscar Ingresar

PAÍS

Multan en Las Condes a Uber Eats y Pedidos Ya por no pedir carné de identidad a compradores de alcohol

RD APPLICATION
CUASI AND NON EXPERIMENTAL METHODS

- Minimum Legal Drinking Age (MLDA) – 21 yo in US
- Small change in age (days or months) has a big impact in access to alcohol



Source: Angrist & Pischke, 2014
in Schwartz, 2019

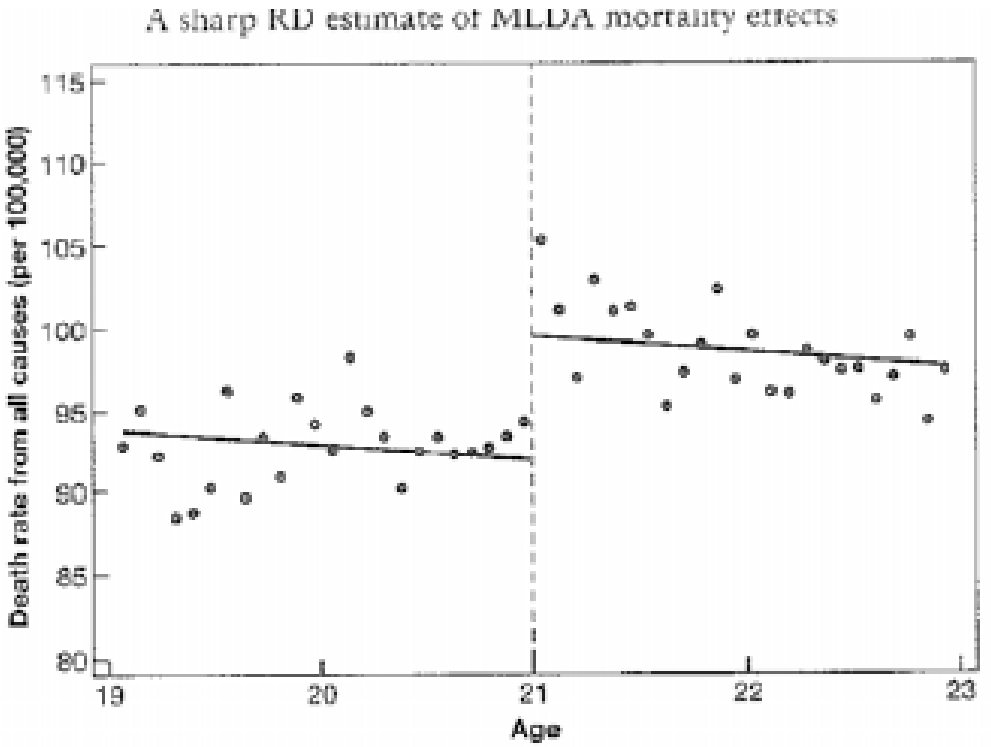
- RD approach could help examine whether spike in deaths is due to the MDLA

RD APPLICATION
CUASI AND NON EXPERIMENTAL METHODS

- Model estimated:

$$NumDeaths_i = \beta_0 + \beta_1 Age_i + \rho Over21_i + u_i$$

- Assumptions:
  - Law says NO alcohol under 21 (with compliance)
  - Are potential outcomes and covariates continuos within Age?



Notes: This figure plots death rates from all causes against age in months. The lines in the figure show fitted values from a regression of death rates on an over-21 dummy and age in months (the vertical dashed line indicates the minimum legal drinking age (MLDA) cutoff).

$\rho = 7.7$

RD APPLICATION
CUASI AND NON EXPERIMENTAL METHODS

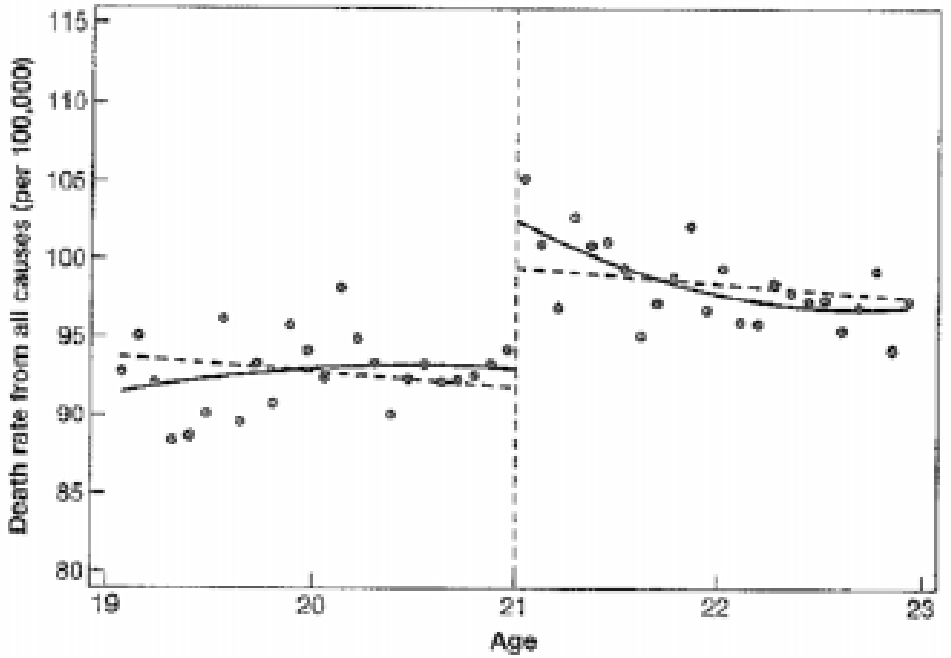
- How can be sure? -> Robustness checks
  - Quadratic terms
  - Check other variables

- Model estimated:

$$NumDeaths_i = \beta_0 + \beta_1 Age_i + \beta_2 Age_i^2 + \rho Over21_i + u_i$$

- This looks more discontinuos than quadratic
- The gap might be caused by the effect

Quadratic control in an RD design



Notes: This figure plots death rates from all causes against age in months. Dashed lines in the figure show fitted values from a regression of death rates on an over-21 dummy and age in months. The solid lines plot fitted values from a regression of mortality on an over-21 dummy and a quadratic in age, interacted with the over-21 dummy (the vertical dashed line indicates the minimum legal drinking age [MLDA] cutoff).

RD APPLICATION
CUASI AND NON EXPERIMENTAL METHODS

- Results of RD in other outcomes:
  - Which ones are alcohol related?
  - Which ones aren't?

Sharp RD estimates of MLDA effects on mortality

Dependent variable	Ages 19–22 (1)	Ages 19–22 (2)
All deaths	7.66	9.55
	(1.51)	(1.83)
Motor vehicle accidents	4.53	4.66
	(.72)	(1.09)
Suicide	1.79	1.81
	(.50)	(.78)
Homicide	.10	.20
	(.45)	(.50)
Other external causes	.84	1.80
	(.42)	(.56)
All internal causes	.39	1.07
	(.54)	(.80)
Alcohol-related causes	.44	.80
	(.21)	(.32)
Controls	age	age, age^2, interacted with over-21
Sample size	48	48

Notes: This table reports coefficients on an over-21 dummy from regressions of month-of-age-specific death rates by cause on an over-21 dummy and linear or interacted quadratic age controls. Standard errors are reported in parentheses.