

# Econometrics 2

Arieda Muço

Central European University

Spring 2020

# Regression Discontinuity

Consider a potential outcome model

$$Y_i = (1 - D_i)Y_{0i} + D_iY_{1i}$$

where in addition to  $(Y_i, D_i)$  we observe a variable  $z_i$  which defines treatment assignment by a cutoff  $z_0$ .

- Cutoff in vote share decides election outcome
- Minimum test score to pass a test
- Minimum number of days employed to be eligible for unemployment benefits

Idea: use the discontinuity created by this rule to identify the treatment effect

## 2 types of designs

### 1. Sharp Regression Discontinuity

$D$  is a deterministic function of  $z$

$$D_i = \begin{cases} 1 & z_i \geq z_0 \\ 0 & z_i < z_0 \end{cases}$$

### 2. Fuzzy Regression Discontinuity

Probability of receiving treatment does not change from 0 to one at the cutoff, but

$$\lim_{z \uparrow z_0} P(D_i = 1 | z_i = z) \neq \lim_{z \downarrow z_0} P(D_i = 1 | z_i = z)$$

other variables also determine treatment assignment, but incentives to participate change discontinuously at the threshold.

# Sharp regression discontinuity

Fig 1: Assignment Probabilities (Sharp RD)

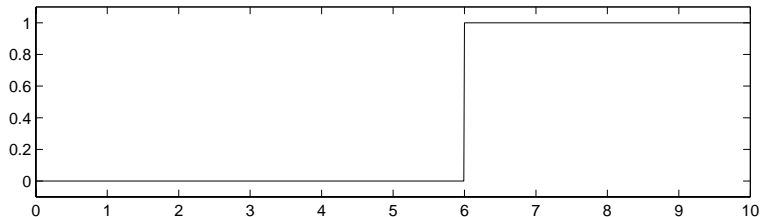


Fig 2: Potential and Observed Outcome Regression Functions

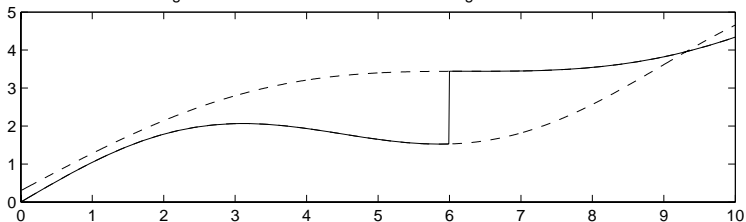


Fig 3: Assignment Probabilities (Fuzzy RD)

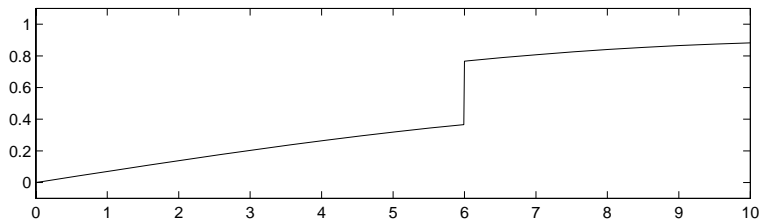
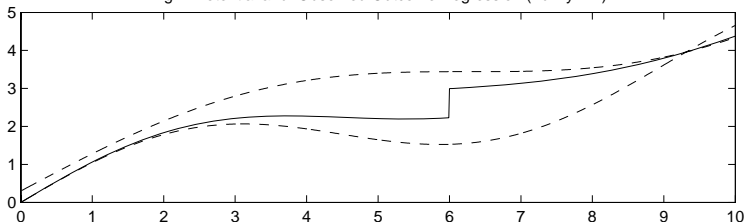


Fig 4: Potential and Observed Outcome Regression (Fuzzy RD)



Selection on observables:

$$(Y_{1i}, Y_{0i}) \perp D_i | z_i$$

holds trivially. In the sharp design there is no variation in  $D_i$  conditional on  $z_i$ .

But there is **zero overlap** with respect to  $z$ .

# Identification Assumptions

- RD.1** No difference in the outcome for those close to the cutoff in the absence of treatment.  $E(Y_{0i}|z_i = z)$  is continuous at  $z_0$ ,  $E(Y_{1i}|z_i = z)$  is continuous at  $z_0$ .
- RD.2** No sorting based on anticipated gains from treatment near the cutoff  $E(x_i|z_i = z)$  is continuous at  $z_0$ .

# Identification of the treatment effect

sharp design

$$\hat{\rho}_s = \lim_{z \downarrow z_0} E(Y_i | z_i = z) - \lim_{z \uparrow z_0} E(Y_i | z_i = z)$$

fuzzy design

$$\hat{\rho}_f = \frac{\lim_{z \downarrow z_0} E(Y_i | z_i = z) - \lim_{z \uparrow z_0} E(Y_i | z_i = z)}{\lim_{z \downarrow z_0} E(D_i | z_i = z) - \lim_{z \uparrow z_0} E(D_i | z_i = z)}$$

Notes:

- We identify a *local* treatment effect around the discontinuity
- Limited external validity, but strong internal validity of the result.
- The fuzzy RD estimator can be interpreted as an IV estimator (compare to Wald estimator)



# Estimation of the treatment effect

## 1. Graphical analysis: Important part of any RD analysis!

- Group the data in bins with similar values of  $z_i$  to the left and right of  $z = z_0$ .
- Plot mean *outcomes*  $Y_i$  in each bin against  $z$ .
- Is there a jump in  $Y$  at  $z = z_0$ ?
- Are there jumps in mean outcomes at other values of  $z \neq z_0$ ?
- First stage: Plot mean values of  $D_i$  in each bin against  $z$ , sharp or fuzzy design?
- Testable assumptions: Plot mean values of *covariates*  $X_i$  in each bin.
- Distribution of  $z$ : Plot the number of observations in each bin.

# Estimation of the treatment effect

**2. Regression analysis** Local linear regression at the boundary, only use observations with values of  $z$  close to the discontinuity

$$Y_i = \alpha + \rho D_i + \beta \mathbf{z}_i + \varepsilon_i$$

Include more observations and control for flexible functions in  $z$  to the left and right of  $z = Z_0$

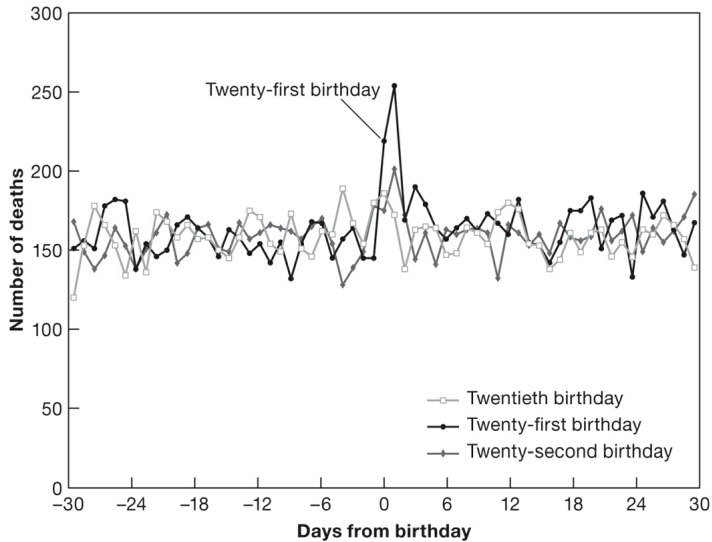
$$Y_i = \alpha + \rho D_i + f_1(z_i - z_0) + D_i * f_2(z_i - z_0) + \beta \mathbf{x}_i + \varepsilon_i$$

- $f_1$  and  $f_2$  are polynomials in  $(z - z_0)$
- $\delta$  measures the jump in  $y$  at  $z = z_0$

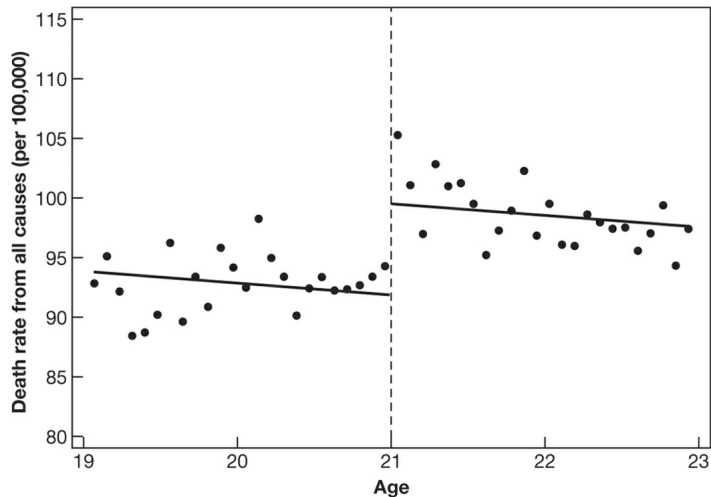
# Applications: Sharp Regression Discontinuity

- Americans over 21 can drink legally
- American college presidents have lobbied to return the Minimum Legal Drinking Age (MLDA) at 18
- Theory (Amethyst Initiative): legal drinking at 18 avoids binge drinking contrasting the view that the age-21 MLDA reduces youth access to alcohol, preventing some harm

# Birthdays and funerals



# Minimum Drinking Age and Mortality Effects



# Minimum Drinking Age and Mortality Effects

$D_a$  is a deterministic function of  $age$ . Age is the running variable.  
MLDA is a sharp function of age

$$D_a = \begin{cases} 1 & a \geq 21 \\ 0 & a < 21 \end{cases}$$

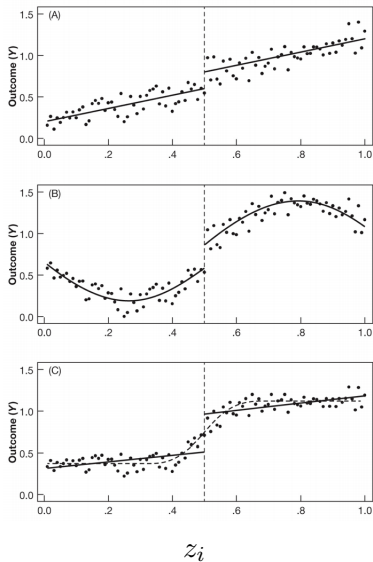
- Treatment status is a deterministic function of  $a$ . Once we know  $a$ , we know  $D_a$
- Treatment status is a discontinuous function of  $a$ , because no matter how close  $a$  gets to the cutoff,  $D_a$  remains unchanged until the cutoff is reached.

# Estimation of the treatment effect

$$M_a = \alpha + \rho D_a + \gamma a + \varepsilon_a$$

- where  $M_a$  is mortality in month  $a$  defined as 30 day interval
- RD tools aren't guaranteed to produce reliable causal estimates.
  - ▶ Get as close to the cutoff as possible
  - ▶ Model nonlinearities directly as polynomial functions of the running variable on both sides of the discontinuity (no need to be the same degree of polynomial on both sides)

# Three cases



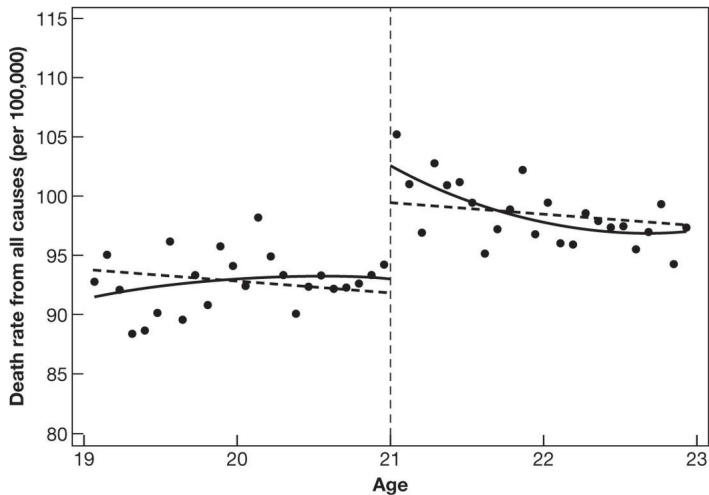


## What polynomial degree?

- Gelman and Imbens (2018) recommend polynomials up to a quadratic to avoid the problem of overfitting. Usually, researchers report estimates with various degrees.
- To make the model with interactions easier to interpret, we center the running variable by subtracting the cutoff

$$M_a = \alpha + \rho D_a + \gamma_1(a - 21) + \gamma_2(a - 21)^2 + \delta_1[D_a(a - 21)] + \delta_2[D_a(a - 21)^2] + \varepsilon_a$$

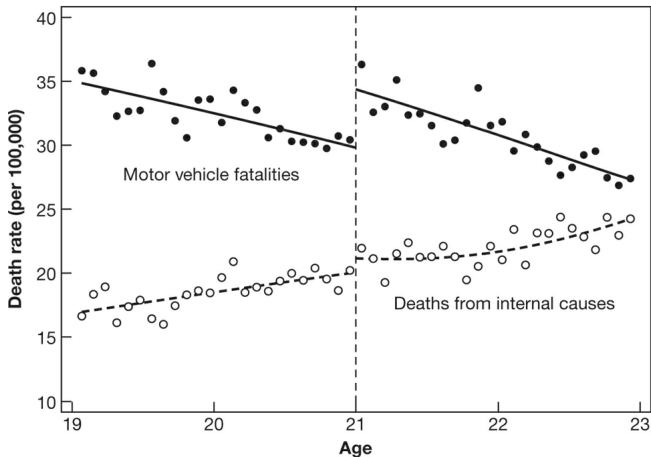
## Quadratic control in an RD design



# Covariate Balance

Dependent variable	Ages 19–22		Ages 20–21	
	(1)	(2)	(3)	(4)
All deaths	7.66 (1.51)	9.55 (1.83)	9.75 (2.06)	9.61 (2.29)
Motor vehicle accidents	4.53 (.72)	4.66 (1.09)	4.76 (1.08)	5.89 (1.33)
Suicide	1.79 (.50)	1.81 (.78)	1.72 (.73)	1.30 (1.14)
Homicide	.10 (.45)	.20 (.50)	.16 (.59)	–.45 (.93)
Other external causes	.84 (.42)	1.80 (.56)	1.41 (.59)	1.63 (.75)
All internal causes	.39 (.54)	1.07 (.80)	1.69 (.74)	1.25 (1.01)
Alcohol-related causes	.44 (.21)	.80 (.32)	.74 (.33)	1.03 (.41)
Controls	age	age, age <sup>2</sup> , interacted with over-21	age	age, age <sup>2</sup> , interacted with over-21
Sample size	48	48	24	24

## RD estimates of MLDA effects on mortality by cause of death



## How close to the boundary?

- As close as possible
- Drawback: few observations left. This causes the resulting estimates are likely to be too imprecise to be useful.
- Solution: Non-parametric RD
- To make the model with interactions easier to interpret, we center the running variable by subtracting the cutoff

$$M_a = \alpha + \rho D_a + \gamma a + \varepsilon_a$$

- in a sample such that:  $a_0 - b \leq a \leq a_0 + b$
- $b$  is the bandwidth which is a function of sampling size
- Theoretical econometricians have proposed strategies for making such bias-variance trade-offs efficiently. The bandwidth selection algorithm is not completely data-dependent and requires researchers to choose certain parameters.

## Local linear nonparametric regression

- Kernel regression as a weighted regression restricted to a window (bandwidth)
- The kernel provides the weights to that regression. A rectangular kernel would give the same result as taking  $E[Y]$  at a given bin on  $z$ . The triangular kernel gives more importance to the observations closest to the center.

$$(\hat{\alpha}, \hat{\rho}) = \underset{\alpha, \rho}{\operatorname{argmin}} \sum_i^n (Y_i - \alpha - \rho(z_i - z_0))^2 K\left(\frac{z_i - z_0}{h}\right) \mathbb{1}(z_i > z_0)$$

- Hahn, Jinyong, Petra Todd, and Wilbert Van der Klaauw. "Identification and estimation of treatment effects with a regression-discontinuity design." *Econometrica* 69.1 (2001): 201-209.
- Imbens, Guido W. and Thomas Lemieux (2008) "Regression Discontinuity Designs: A Guide to Practice" *Journal of Econometrics*, 142, 615-635.
- Gelman, Andrew, and Guido Imbens. "Why high-order polynomials should not be used in regression discontinuity designs." *Journal of Business & Economic Statistics* (2018): 1-10.

# Book References

- Angrist, Joshua D., and Jörn-Steffen Pischke. Mostly harmless econometrics: An empiricist's companion. Princeton university press, 2008.
- Angrist, Joshua D., and Jörn-Steffen Pischke. Mastering' metrics: The path from cause to effect. Princeton University Press, 2014.
- Cunningham, Scott. "Causal Inference: The Mixtape."