Program Evaluation (Causal Inference) 2: Regression Discontinuity Design

Instructor: Yuta Toyama

Last updated: 2020-03-30

Section 1

Introduction

Introduction

- Regression Discontinuity Design
  - Exploit the discontinuous change in treatment status to estimate the causal effect.
  - Example:
    - Threshold of test score for college admission
    - Eligibility of policy due to age.
    - Geographic boundary of two regions.
- Pros: Strong internal validity
  - Assumption for identification is weak.
- Cons: Very little external validity
  - What we estimate is the effect on people at the boundary.

Idea in Figure

Reference

- Angrist and Pischke “Mostly harmless econometrics” Chapter 6
- R packages: https://sites.google.com/site/rdpackages/rdrobust

Section 2

Framework

Framework

- Y_i: observed outcome for person i
- Define *potential outcomes*
  - Y_{1i}: outcome for i when she is treated (treatment group)
  - Y_{0i}: outcome for i when she is not treated (control group)
- D_i: treatment status is deterministically determined (sharp RD design)

$$D_i = \mathbf{1}\{W_i \geq \bar{W}\}$$

  - W_i: running variable (forcing variable).
  - Probabilistic assignment is allowed (fuzzy RD design)

Example: Incumbent Advantage

- Consider the two-candidate elections
  - D_i: dummy for incumbent in the election
  - Y_i: whether the candidate win in the election
  - W_i : the vote share in the previous election.
- The incumbent status is defined as

$$D_i = \mathbf{1}\{W_i \geq 0.5\}$$

- Idea of RD:
  - Suppose that you won with 51%.
  - You are similar to the guy who lose at 49% (main assumption of RD).
  - If you focus on these people, D_i is as if it were randomly assigned.

Framework cont.d

- Note that $D_i = \mathbf{1}\{W_i \geq \bar{W}\}$ implies the unconfoundedness

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

- But the overlap assumption does not hold

$$P(D_i = 1 | W_i = w) = \begin{cases} 1 & if \;\; w \geq \bar{W} \\ 0 & if \;\; w < \bar{W} \end{cases}$$

- To compare people with and without treatment, we need to rely on some sort of extrapolation around the threshold.

Linear approach

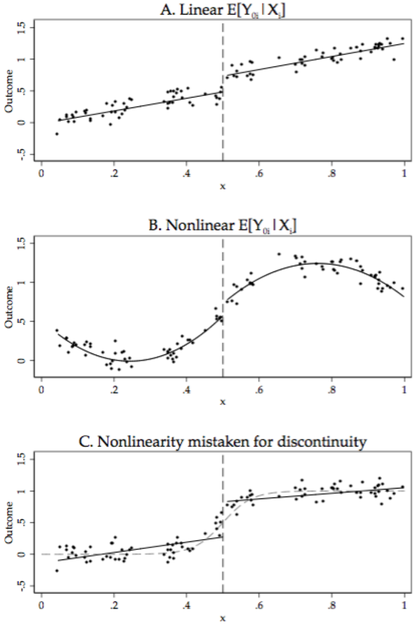
- Suppose for a moment that

$$Y_{1i} = \rho + Y_{0i}$$
$$E[Y_{0i}|W_i = w] = \alpha_0 + \beta_0 w$$

- This leads to a regression

$$Y_i = \alpha + \beta W_i + \rho D_i + \eta_i$$

  - ρ is the causal effect.
- This approach relies on linear extrapolation. May not be good.
  - What if $E[Y_{0i}|W_i = w]$ is nonlinear?


A. Linear $E[Y_{0i} | X_i]$
B. Nonlinear $E[Y_{0i} | X_i]$
C. Nonlinearity mistaken for discontinuity
Outcome
x

A more general approach

- Allowing for nonlinear effect of the running variable W_i

$$Y_i = f(W_i) + \rho \mathbf{1}\{W_i \geq \bar{W}\} + \eta_i$$

- A function $f(\cdot)$ might be a pth order polynomial.

$$f(W_i) = \beta_1 W_i + \beta_2 W_i^2 + \cdots + \beta_p W_i^p$$

  - nonparametric approach later.

Implementation in Regression

- Consider

$$E[Y_{0i}|W_i = w] = f_0(W_i - \bar{W})$$
$$E[Y_{1i}|W_i = w] = \rho + f_1(W_i - \bar{W})$$

  - $\tilde{W}_i = W_i - \bar{W}$ is a normalization.
- Then the regression equation is (See page 255 in Angrist and Pischke)

$$\begin{aligned} Y_i = {} & \alpha + \beta_{01}\tilde{W}_i + \cdots + \beta_{0p}\tilde{W}_i^p \\ & + \rho D_i + \beta_1^* D_i \tilde{W}_i + \cdots + \beta_p^* D_i \tilde{W}_i^p + \eta_i \end{aligned}$$

  - ρ is the causal effect.
- When running regression, need to focus on the sample around threshold.
  - How close the sample should be to the threshold can be taken care by statistical procedure.

Section 3

Example

From Mastering Metrics Sec 4.1: Effects of the minimum age drinking law

FIGURE 4.1
Birthdays and funerals

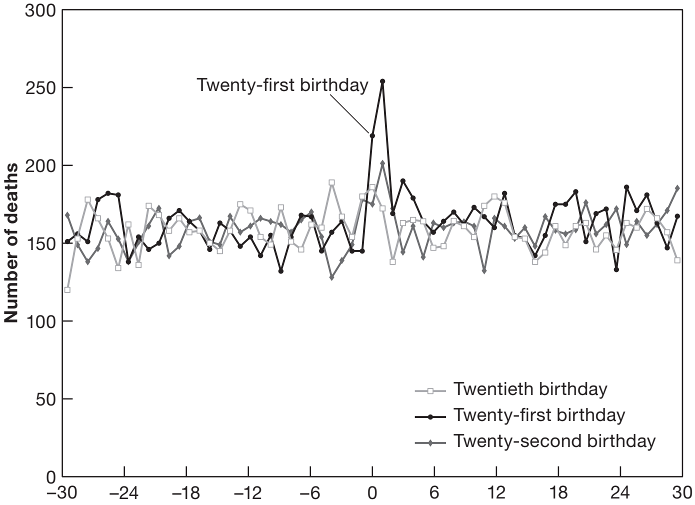
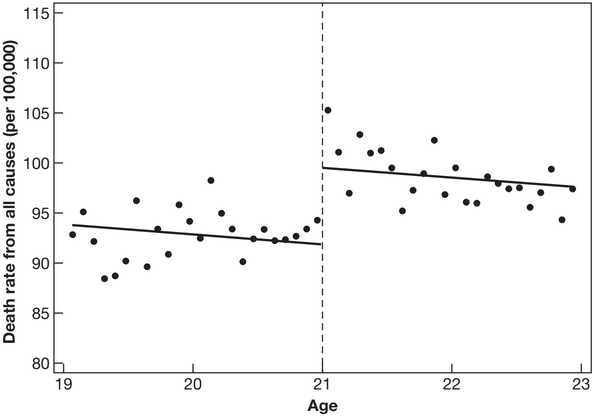
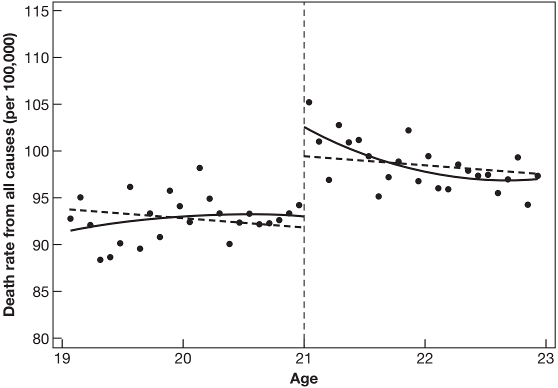


FIGURE 4.2
A sharp RD estimate of MLDA mortality effects



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).

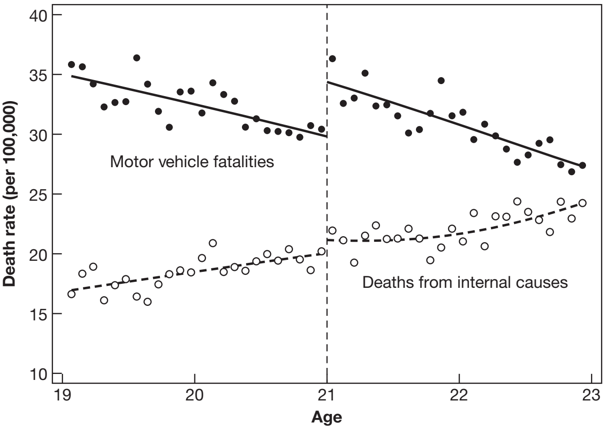
FIGURE 4.4
Quadratic control in an RD design



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).

FIGURE 4.5

RD estimates of MLDA effects on mortality by cause of death



Notes: This figure plots death rates from motor vehicle accidents and internal causes against age in months. Lines in the figure plot fitted values from regressions of mortality by cause on an over-21 dummy and a quadratic function of age in months, interacted with the dummy (the vertical dashed line indicates the minimum legal drinking age [MLDA] cutoff).

Section 4

Formal Analysis

Formal Identification Analysis

- Key: continuity assumptions: Both $E[Y_{1i}|W_i = w]$ and $E[Y_{0i}|W_i = w]$ are continuous at the threshold $w = \bar{W}$.
  - This is not directly testable assumption (because we cannot observe Y_{1i} below the threshold).
  - Will discuss several validating approaches.
- To see how this works, notice that

$$
\begin{aligned}
E[Y_i|W_i = w] =& E[Y_{0i}|W_i = w] \\
&+ \mathbf{1}\{w \geq \bar{W}\} \left(E[Y_{1i}|W_i = w] - E[Y_{0i}|W_i = w] \right)
\end{aligned}
$$

- Taking the limit of w to $\bar{W}$ from above and below

$$
\lim_{w \uparrow \bar{W}} E[Y_i|W_i = w] = \lim_{w \uparrow \bar{W}} E[Y_{0i}|W_i = w] = E[Y_{0i}|W_i = \bar{W}]
$$

$$
\lim_{w \downarrow \bar{W}} E[Y_i|W_i = w] = \lim_{w \downarrow \bar{W}} E[Y_{1i}|W_i = w] = E[Y_{1i}|W_i = \bar{W}]
$$

  - Notice that we use continuity in the second equalities!

- Remember that

$$\lim_{w \uparrow \bar{W}} E[Y_i | W_i = w] = \lim_{w \uparrow \bar{W}} E[Y_{0i} | W_i = w] = E[Y_{0i} | W_i = \bar{W}]$$

$$\lim_{w \downarrow \bar{W}} E[Y_i | W_i = w] = \lim_{w \downarrow \bar{W}} E[Y_{1i} | W_i = w] = E[Y_{1i} | W_i = \bar{W}]$$

- So, we have

$$E[Y_{1i} - Y_{0i} | W_i = \bar{W}] = \lim_{w \downarrow \bar{W}} E[Y_i | W_i = w] - \lim_{w \uparrow \bar{W}} E[Y_i | W_i = w]$$

  - LHS: Average treatment effect at the threshold
  - RHS: We can observe from the data.
    - Conditional expectation near the threshold.

Nonparametric Implementation

- Too much details for this course. I will skip.
- The package 'rddrobust" will take care for all the details.

Section 5

Validation of Assumptions

Validation of Assumptions

- The key assumptions : Both $E[Y_{1i}|W_i = w]$ and $E[Y_{0i}|W_i = w]$ are continuous at the threshold $w = \bar{W}$.
- This is not directly testable because we cannot observe Y_{1i} below the threshold.
- There are two common approaches that support this assumption:
  1. Covariate test
  2. Density test (no bunching in the running variable).

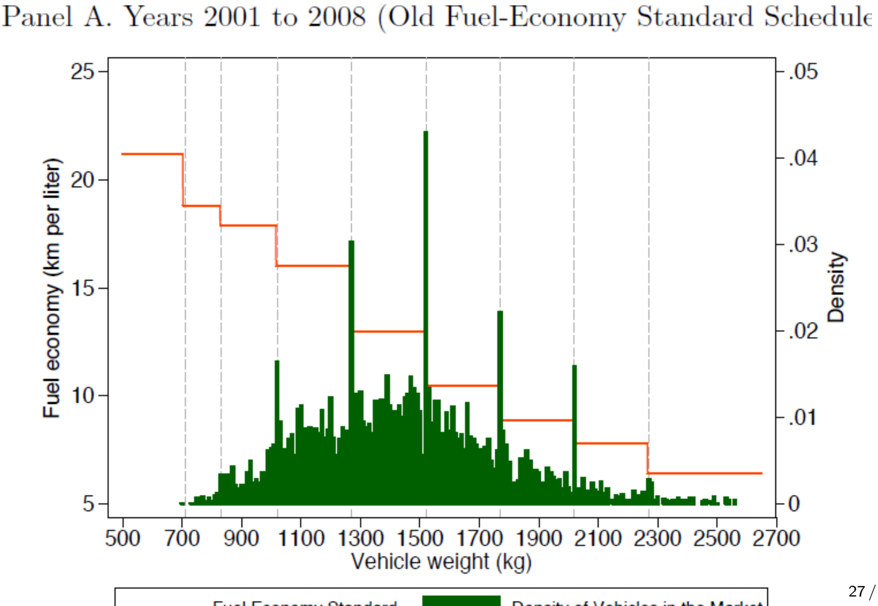
Covariate Test

- The underlying idea of RDD: Comparing outcomes right above and right below $\bar{W}$ provides a comparison of treated and control agents who are similar due to the assumed continuity in conditional distributions
- If this is a valid comparison, then we would expect that covariates X also change smoothly as we pass through the threshold.
  - Run the RDD on the covariate X.
- If we found the discontinuity, it suggests that the conditional expectation of Y on W may not be continuous either.
- If X has a direct effect on Y, the discontinuity in $E[Y_i|W]$ at $\bar{W}$ will confound the treatment effect.
- Example:
  - Y hours worked,
  - D: older-than-65 discounts,
  - W: age, X: social security benefit (non-work income)

Density Test, or No Bunching

- There is also a concern about manipulation if agents know about the institutional details
    - For example, if schools scoring lower than w = 50 on standardized tests get labeled as dysfunctional, we might expect a lot of schools to be right above 50
- In this case, we observe a bunching around the threshold.
    - Agents are "manipulating" treatment assignment around the threshold.
    - Density of W_i is discontinuous at $\bar{W}$
- We would expect that $E[Y_{1i}|W_i = w]$ would be also discontinuous.
- McCrary (2008) suggests a test of the null hypothesis that the density of W_i is continuous at $\bar{W}$.
- Note: Bunching itself is an interesting economic phenomenon. It can be used to analyze a different question.

Ito and Sallee (2018, REStat)



Bunching Estimator