Introduction to RD Designs

Northwestern Causal Inference Workshop

Gonzalo Vazquez-Bare

Department of Economics, UC Santa Barbara

August 1, 2024

Motivation: the Head Start program

- Head Start:
  - Established in 1965 as part of the War on Poverty
  - Provided health, nutritional and social services
  - Targeted children ages three to five
  - Served over 900,000 children at a cost of $7 billion in 2006*
- Understanding the impact of HS is crucial:
  - Is the program having its intended effect?
  - Is the program cost-effective?

*See Ludwig and Miller (2007, Quarterly Journal of Economics) for details

Motivation: the Head Start program

- We have county-level data for period 1973-1983
- Let Y_i denote the child mortality rate for county i in that period
- In 1965, some counties received assistance to develop HS proposals
  - HS participation 50-100% higher in counties that received assistance
- Let D_i be a treatment indicator:

$$D_i = \begin{cases} 1 & \text{if } i \text{ received assistance in 1965} \\ 0 & \text{otherwise} \end{cases}$$

- How do we measure the effect of HS on child mortality rates?

Motivation: the Head Start program

- We can compare counties with and without HS assistance
- Naive estimator: difference in means

$$\hat{\Delta} = \bar{Y}_1 - \bar{Y}_0$$

where $\bar{Y}_d$ = average outcome among counties with $D_i = d$

Motivation: the Head Start program

- We can compare counties with and without HS assistance
- Naive estimator: difference in means

$$\hat{\Delta} = \bar{Y}_1 - \bar{Y}_0$$

where $\bar{Y}_d$ = average outcome among counties with $D_i = d$

	$D_i = 1$	$D_i = 0$	Diff	p-val
Y_i	2.42	2.23	0.19	0.52

Motivation: the Head Start program

- We can compare counties with and without HS assistance
- Naive estimator: difference in means

$$\hat{\Delta} = \bar{Y}_1 - \bar{Y}_0$$

where $\bar{Y}_d$ = average outcome among counties with $D_i = d$

	$D_i = 1$	$D_i = 0$	Diff	p-val
Y_i	2.42	2.23	0.19	0.52
population	17.60	41.50	-38.96	0.00
% urban	13.39	30.97	-17.58	0.00
% black	33.91	7.89	26.02	0.00
income PC	8,334.79	11,031.27	-10,742.67	0.00
mortality pre-HS	10.68	6.84	3.83	0.00

Motivation: the Head Start program

- Small and insignificant difference between treated and untreated
- But treated and untreated counties differ in many other dimensions
- What are we actually measuring with $\hat{\Delta}$?
- What do we want to measure?
  - What do we mean by "the causal effect" of HS?

The potential outcomes framework

- We observe a unit i from a population of interest
- Each unit i can be in one of two scenarios: treated or untreated
- Potential outcomes:
  - $Y_i(1)$: potential outcome when treated
  - $Y_i(0)$: potential outcome when untreated
- The *treatment effect* for unit i is:

$$\tau_i = Y_i(1) - Y_i(0)$$

The potential outcomes framework

- The observed outcome is:

$$Y_i = \begin{cases} Y_i(1) & \text{if } D_i = 1 \\ Y_i(0) & \text{if } D_i = 0 \end{cases}$$

- We can rearrange this as:

$$Y_i = Y_i(1)D_i + Y_i(0)(1 - D_i)$$

or equivalently

$$\boxed{Y_i = Y_i(0) + \tau_i D_i}$$

The Fundamental Problem of Causal Inference

The Fundamental Problem of Causal Inference (Holland, 1986)

It is impossible to observe the value of $Y_i(1)$ and $Y_i(0)$ on the same unit and, therefore, it is impossible to observe the effect of the treatment on unit i.

- There is always a *counterfactual* scenario that is not observed
- The causal effect is unobservable
- How can we infer the counterfactual?

Observed data and selection bias

- Our measure of the treatment effect so far is:

$$\hat{\Delta} = \bar{Y}_1 - \bar{Y}_0$$

which is an estimator of:

$$\Delta = \mathbb{E}[Y_i | D_i = 1] - \mathbb{E}[Y_i | D_i = 0]$$

- What is the relationship between Δ and τ_i?

Observed data and selection bias

$$
\begin{aligned}
\Delta &= \mathbb{E}[Y_i | D_i = 1] - \mathbb{E}[Y_i | D_i = 0] \\
&= \mathbb{E}[Y_i(1) | D_i = 1] - \mathbb{E}[Y_i(0) | D_i = 0] \\
&= \mathbb{E}[Y_i(1) | D_i = 1] - \mathbb{E}[Y_i(0) | D_i = 1] \\
&\quad + \mathbb{E}[Y_i(0) | D_i = 1] - \mathbb{E}[Y_i(0) | D_i = 0] \\
&= ATT + SB
\end{aligned}
$$

- $ATT = \mathbb{E}[\tau_i | D_i = 1]$: average treatment effect on the treated
- $SB = \mathbb{E}[Y_i(0) | D_i = 1] - \mathbb{E}[Y_i(0) | D_i = 0]$: selection bias

Observed data and selection bias

$$\mathbb{E}[Y_i|D_i = 1] - \mathbb{E}[Y_i|D_i = 0] = ATT + SB$$

- $SB = \mathbb{E}[Y_i(0)|D_i = 1] - \mathbb{E}[Y_i(0)|D_i = 0]$ is generally $\neq 0$
- Treated and untreated differ in factors other than the treatment
- Key issue to analyze: *treatment assignment mechanism*
  - What determines who is treated and who is not?

Head Start and the RD design

- How was Head Start assistance offered?
- Let X_i denote county i's poverty index in 1960
- County i was offered HS assistance if:

$$X_i \geq 59.1984$$

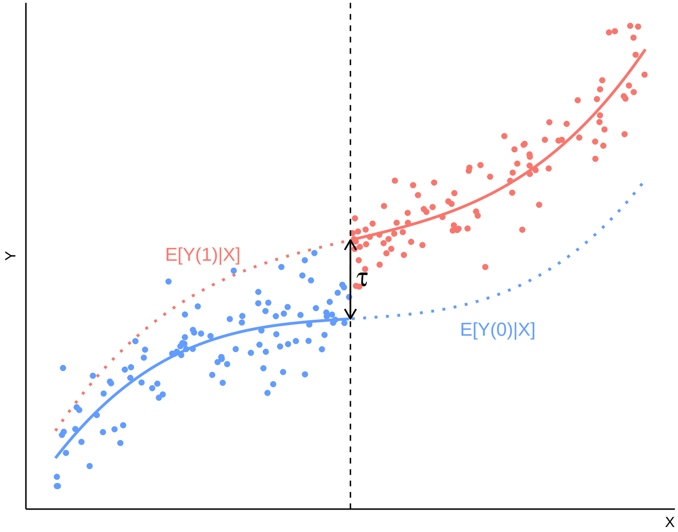
- This introduces a discontinuity in the probability of treatment:

$$\mathbb{P}[D_i = 1 | X_i] = \mathbb{1}(X_i \geq 59.1984)$$

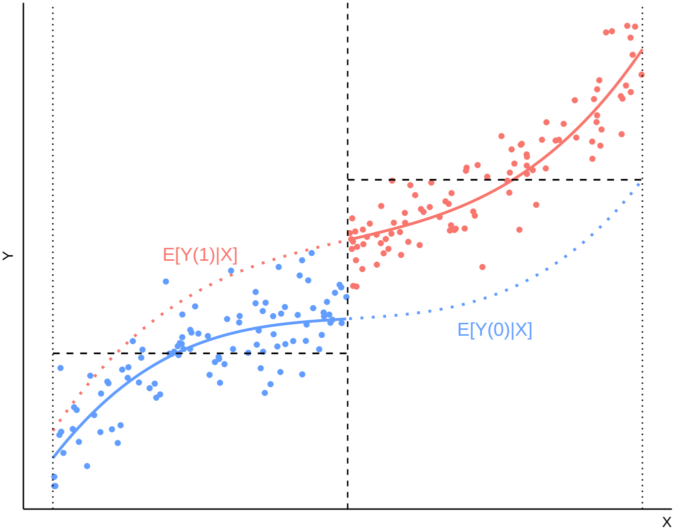
Head Start and the RD design

- How do we exploit this discontinuity?
- Suppose:
  1. Counties don't have precise control on their poverty index
  2. Unobserved confounders do not "jump" at the cutoff
- Compare counties slightly above and slightly below the cutoff
- Conditional averages: $\mathbb{E}[Y(0)|X]$, $\mathbb{E}[Y(1)|X]$

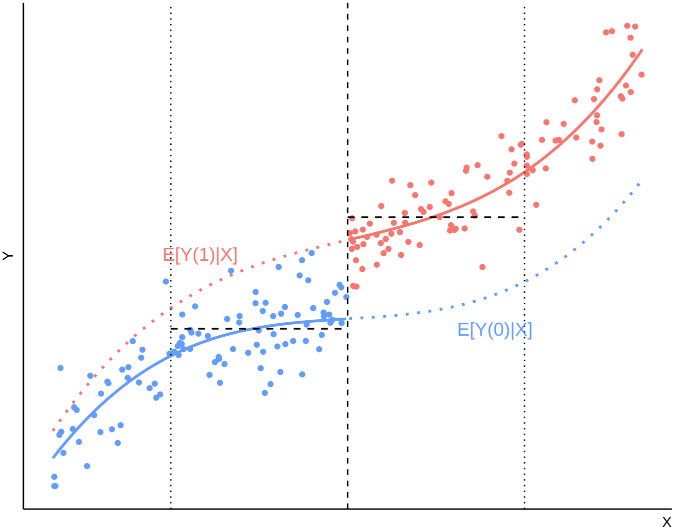
RD: intuition



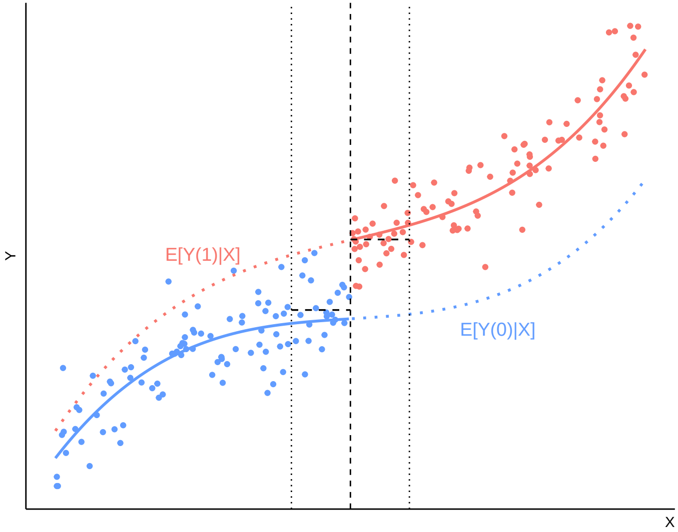
RD: intuition



RD: intuition



RD: intuition



Identification (Hahn, Todd and Van der Klaauw, 2001)

Nonparametric identification in RD

Suppose:

1. (sharp design): $D_i = \mathbb{1}(X_i \geq c)$
2. (smoothness): $\mathbb{E}[Y_i(0)|X_i = x]$, $\mathbb{E}[Y_i(1)|X_i = x]$ are continuous at $x = c$

Then,

$$\mathbb{E}[\tau_i|X_i = c] = \lim_{x \downarrow c} \mathbb{E}[Y_i|X_i = x] - \lim_{x \uparrow c} \mathbb{E}[Y_i|X_i = x]$$

Intuition

In a sharp design, if the underlying functions are continuous, the difference in the right and left limits of the average observed outcomes identifies the average treatment effect at the cutoff.

Components of an RD analysis

- Graphical presentation
- Estimation and inference
- Model validation / falsification
- Sensitivity and robustness

RD software packages

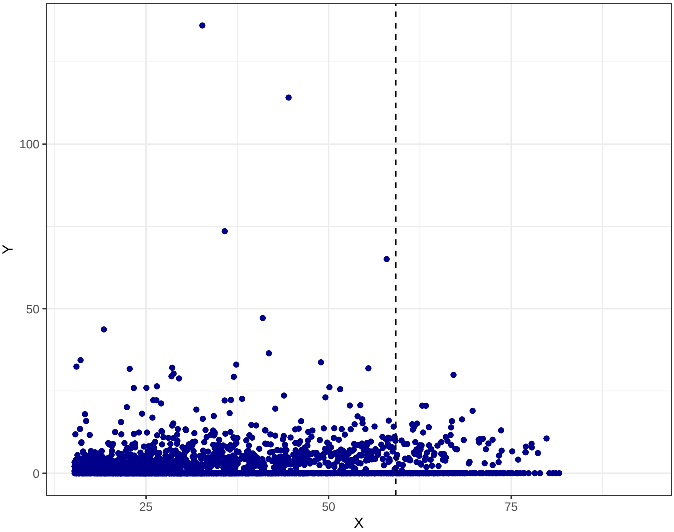
- `rdrobust`: graphical presentation and local polynomial methods
- `rddensity`: density discontinuity tests (manipulation testing)
- `rdlocrand`: local randomization methods
- `rdmulti`: analysis of RD with multiple cutoffs or scores
- `rdpower`: power and sample size calculation

https://rdpackages.github.io/

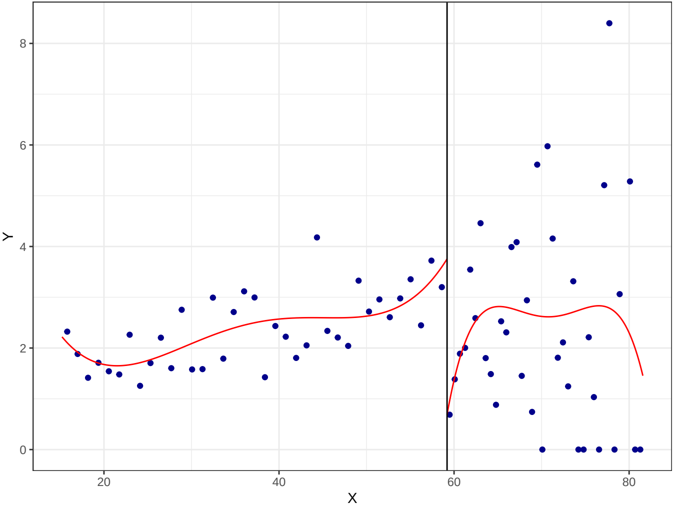
Graphical presentation of RD designs

- RD designs are easy to visualize graphically
- Plots add transparency to the analysis
- Goal: estimate regression functions above and below cutoff
- Scatter plots may not depict information clearly
- Typical RD plot:
  - Global polynomial fit
  - Local means / binning

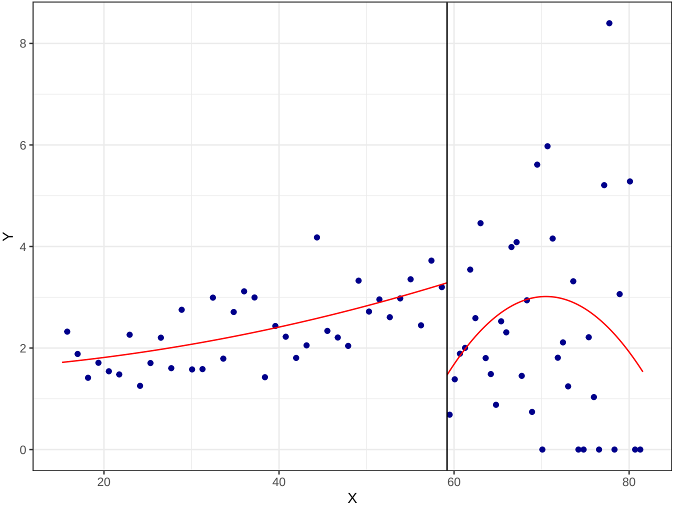
Graphical presentation: scatter plot



Graphical presentation: `rdplot`



Graphical presentation: `rdplot`



Graphical presentation: `rdplot`