Regression Discontinuity Design

Jörn-Steffen Pischke

LSE

October 26, 2018

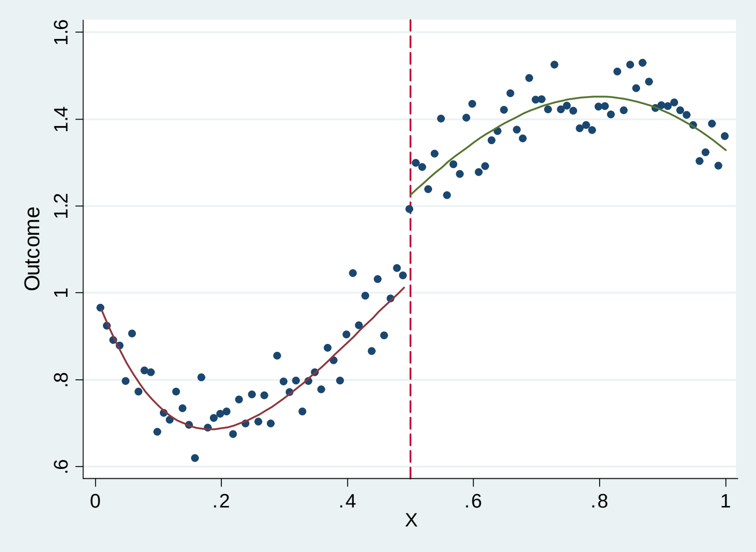
Sharp RD

- Regression: conditional independence assumption $E[Y_{0i}|X_i, D_i] = E[Y_{0i}|X_i]$. Once we control for a confounder X_i, treatment assignment is as good as random.
- The key to the RD design is that we have a deep understanding of the mechanism which underlies the assignment of treatment D_i. In this case, assignment to treatment depends on a single variable X_i. In the sharp RD design this variable fully determines the treatment according to the cutoff rule:

$$D_i = \begin{cases} 1 & \text{if } X_i \geq X_0 \\ 0 & \text{if } X_i < X_0 \end{cases}.$$

- X_i is called the *running variable*.

Sharp RD



Estimating the Sharp RD

$E[Y_{0i}|X_i]$ is a function of X_i, so write

$$\begin{aligned} Y_{0i} &= f(X_i) + e_i \\ Y_{1i} &= Y_{0i} + \beta \\ Y_i &= f(X_i) + \beta D_i + e_i \\ &= f(X_i) + \beta 1(X_i \geq X_0) + e_i. \end{aligned}$$

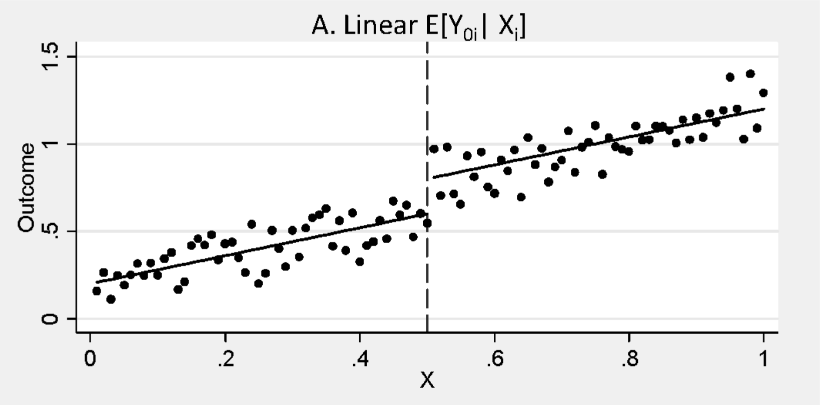
The function $f(\cdot)$ must be continuous at X_0.
In practice, we will have to assume some flexible functional form for $f(\cdot)$, for example a polynomial.
A. Gelman and G. Imbens "Why high-order polynomials should not be used in regression discontinuity designs," *JBES* (2017)

Interpreting RD

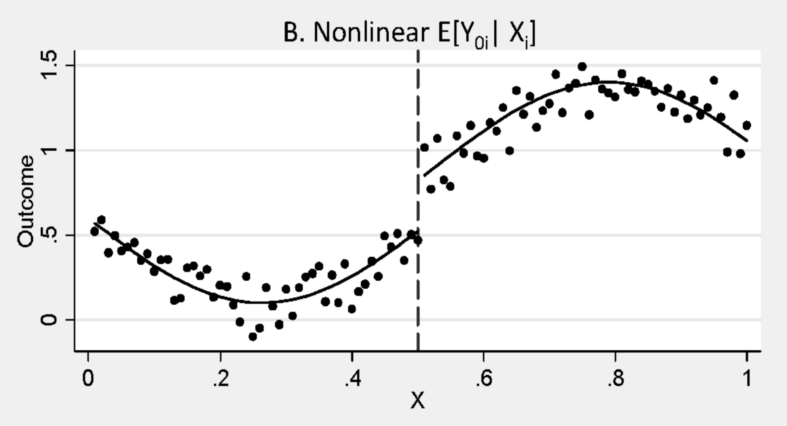
RD relies on regression, yet RD identification is distinct.

- In regression (matching) we hope that treatment is as good as randomly assigned after conditioning on controls.
  - There will be units with the same values of the controls (matches) but with different treatment status.
- RD extrapolates at the discontinuity
  - There is no value of the running variable where we observe both treatment and control observations. We need to be willing to extrapolate across the discontinuity.
  - But for observations very close to the discontinuity we effectively have an experiment.
  - RD estimates are local to the cutoff.

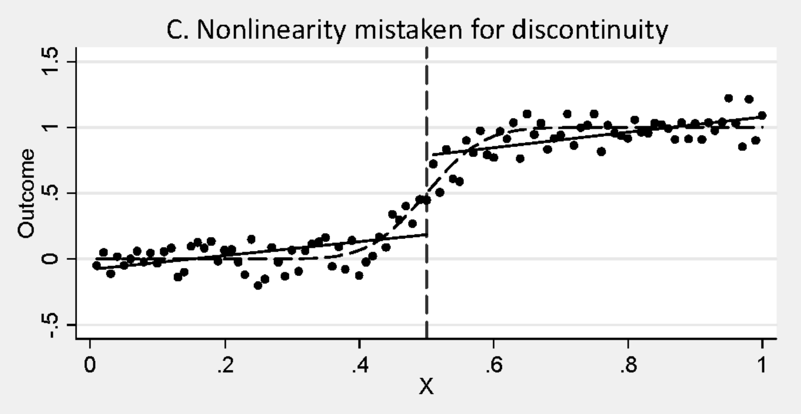
How to get the functional form of f(X) right?



How to get the functional form of f(X) right?



How to get the functional form of f(X) right?



How to get the functional form of f(X) right?

- Flexible enough to get the functional form of $E[Y_{0i}|X_i]$ right away from X_0.
  - Look at the picture to see how much flexibility you need.
- Make the window of your data around X_0 smaller and stick with linear.
  - RD only provides local information on treatment at X_0.
  - You need a lot of data near X_0 for this.
  - Try a few different windows.
- Data dependent window choice: G. Imbens and K. Kalyanaraman "Optimal bandwidth choice for the regression discontinuity estimator," *Review of Economic Studies* (2012)

Generalizations: Different Functions either Side of the Cutoff

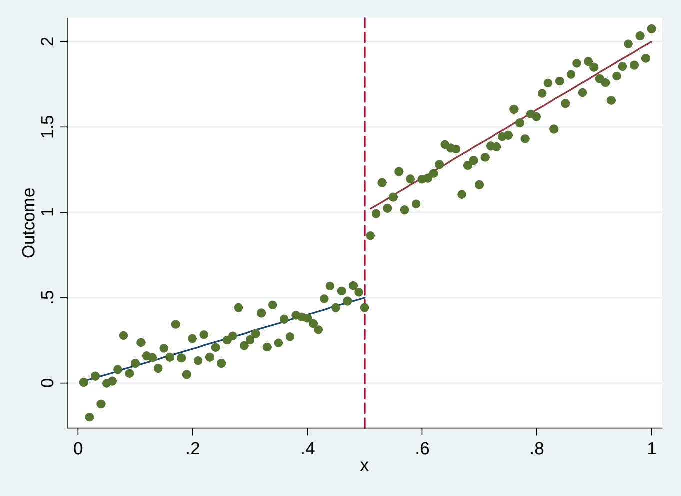
We can allow for different functions to the left and the right of the cutoff:

$$Y_i = f_l(X_i)1(X_i < X_0) + f_r(X_i)1(X_i \geq X_0) + \beta D_i + e_i$$

Need to impose $f_l(X_0) = f_r(X_0)$. For a linear function for $f(X_i)$ this amounts to

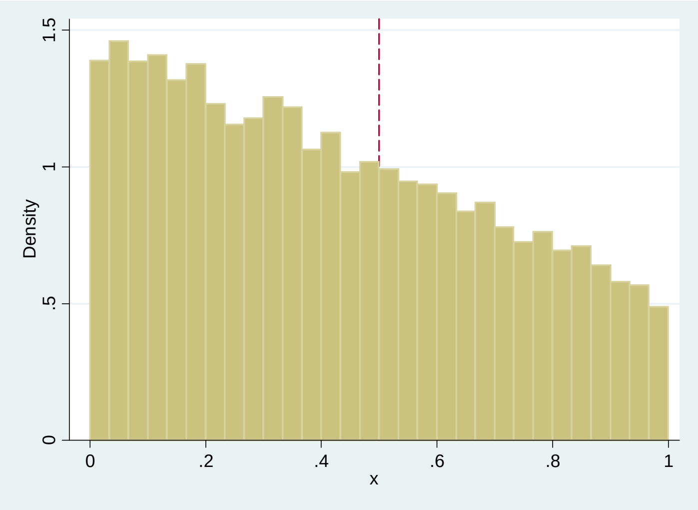
$$Y_i = \alpha + \gamma_1 1(X_i < X_0)(X_i - X_0) + \gamma_2 1(X_i \geq X_0)(X_i - X_0) + \beta D_i + e_i.$$

Unequal Slopes



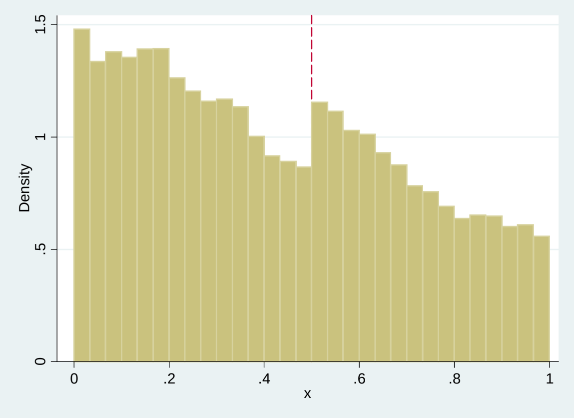
Pitfalls of RD

Is the Distribution of the Running Variable Smooth?



Pitfalls of RD

Running Variable Looks Manipulated



Justin McCrary "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of Econometrics,* 142(2), February 2008, Pages 698–714.

Fuzzy RD

- Instead of a deterministic assignment rule there may only be a change in the probability of treatment at the cutoff.

$$\begin{aligned}\Pr(D_i &= 1) = p(X_i) \\ \lim_{X_i \uparrow X_0} p(X_i) &\neq \lim_{X_i \downarrow X_0} p(X_i).\end{aligned}$$

- The probability of treatment $p(X_i)$ is also a continuous function, except at X_0.
- Sharp RD is the special case

$$p(X_i) = \begin{cases} 1 & \text{if } X_i \geq X_0 \\ 0 & \text{if } X_i < X_0 \end{cases}.$$

Fuzzy RD is IV

There are various regressions now

$$
\begin{aligned}
Y_i &= f_1(X_i) + \beta D_i + e_i && \text{structural equation} \\
Y_i &= f_2(X_i) + \pi_2 1(X_i \geq X_0) + \xi_{2i} && \text{reduced form} \\
D_i &= g(X_i) + \pi_1 1(X_i \geq X_0) + \xi_{1i} && \text{first stage.}
\end{aligned}
$$

The cutoff induces a change in the probability of treatment. If treatment matters, this induces a change in the outcome. Since the treatment doesn't affect all units, the jump at the cutoff in the outcome needs to be rescaled by the jump at the cutoff in the probability of treatment
$=>$ standard IV

$$\beta = \frac{\pi_2}{\pi_1}.$$