Regression Discontinuity Designs

Kristiina Huttunen
Aalto University

Semptember 17, 2021

`Mini-course on Causal Inference`

Motivation: Selection problem

- In economics, we often want to ask causal "if-then questions", i.e. How does university education increase individual's earnings?
- We cannot simply compare those who went to university and those who did not. Those who went to university may be more motivated and talended...
- We cannot randomly allocate students to universities.

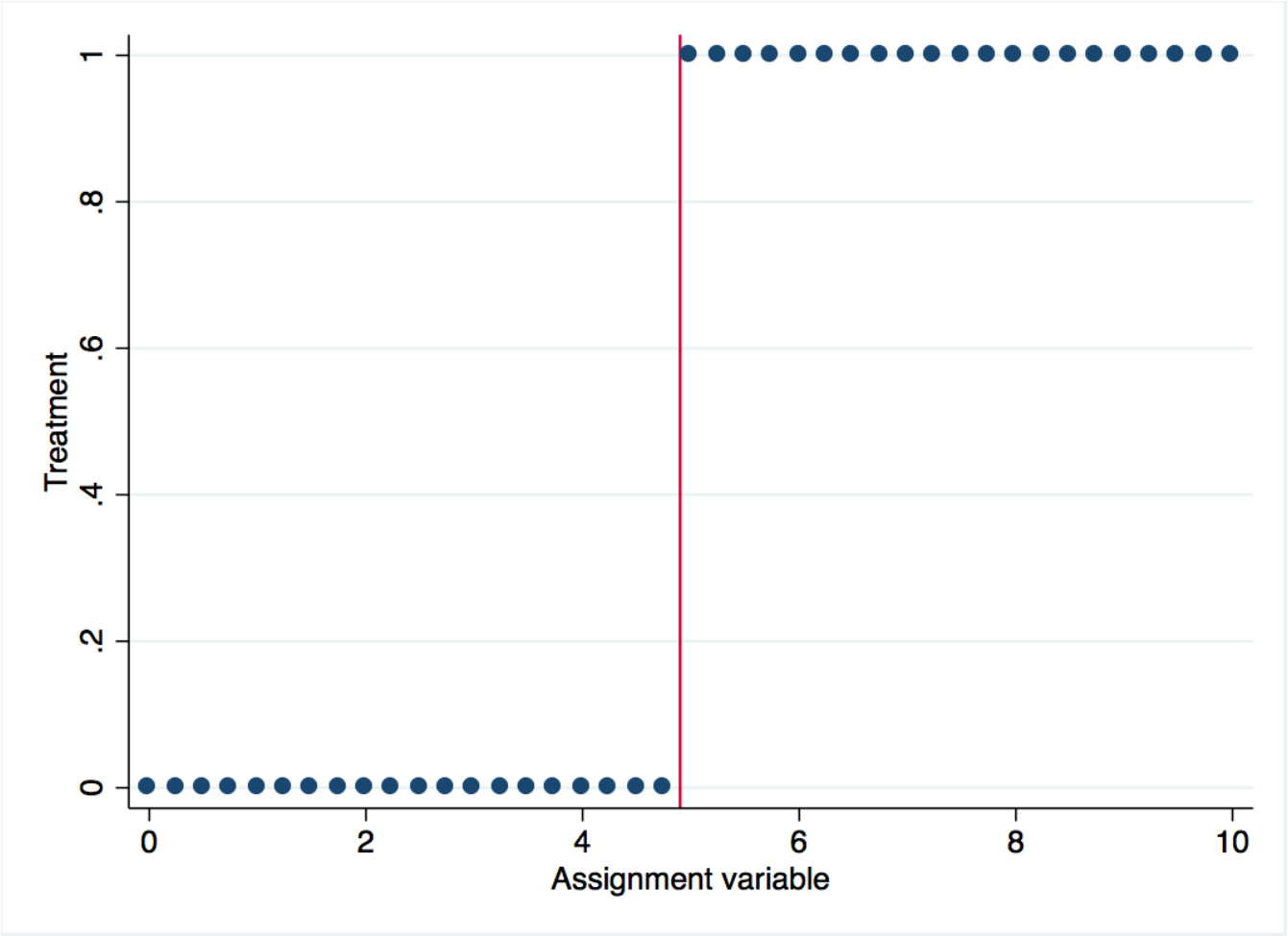
Regression discontinuity design idea

"Rules create experiments"

- Institutional rules often assign individuals to treatments
- We can exploit these rule to estimate the causal effects of the treatment
- The most typical case are threshold rules that are based on some ex-ante variable
  - Test score in school entry exam
  - Income for support eligibility
  - Growth hormone deficit test result for growth hormone treatment
- This ex-ante variable is called the assignment, running (forcing) variable

(Sharp) Regression Discontinuity Design

Those whose value of the assigment variable is above threshold level receive the treatment, others not



Treatment is received if the value of the test is above 5.

Regression Discontinuity Design

- The main idea in the RDD is to compare the outcomes below (control) and above (treated) the threshold
  - Near the treshold the treatment is as good as randomly assigned
- We assume that
  - Treatment status is a deterministic function of the assignment (running) variable
  - Treatment staus is discontinuous function of the assignment (running) variable
- Sharp RD design: Treatment jumps at threshold from 0 to 1
- Fuzzy RD design:The probability of treatment jumps at the threshold

Sharp Regression Discontinuity Design

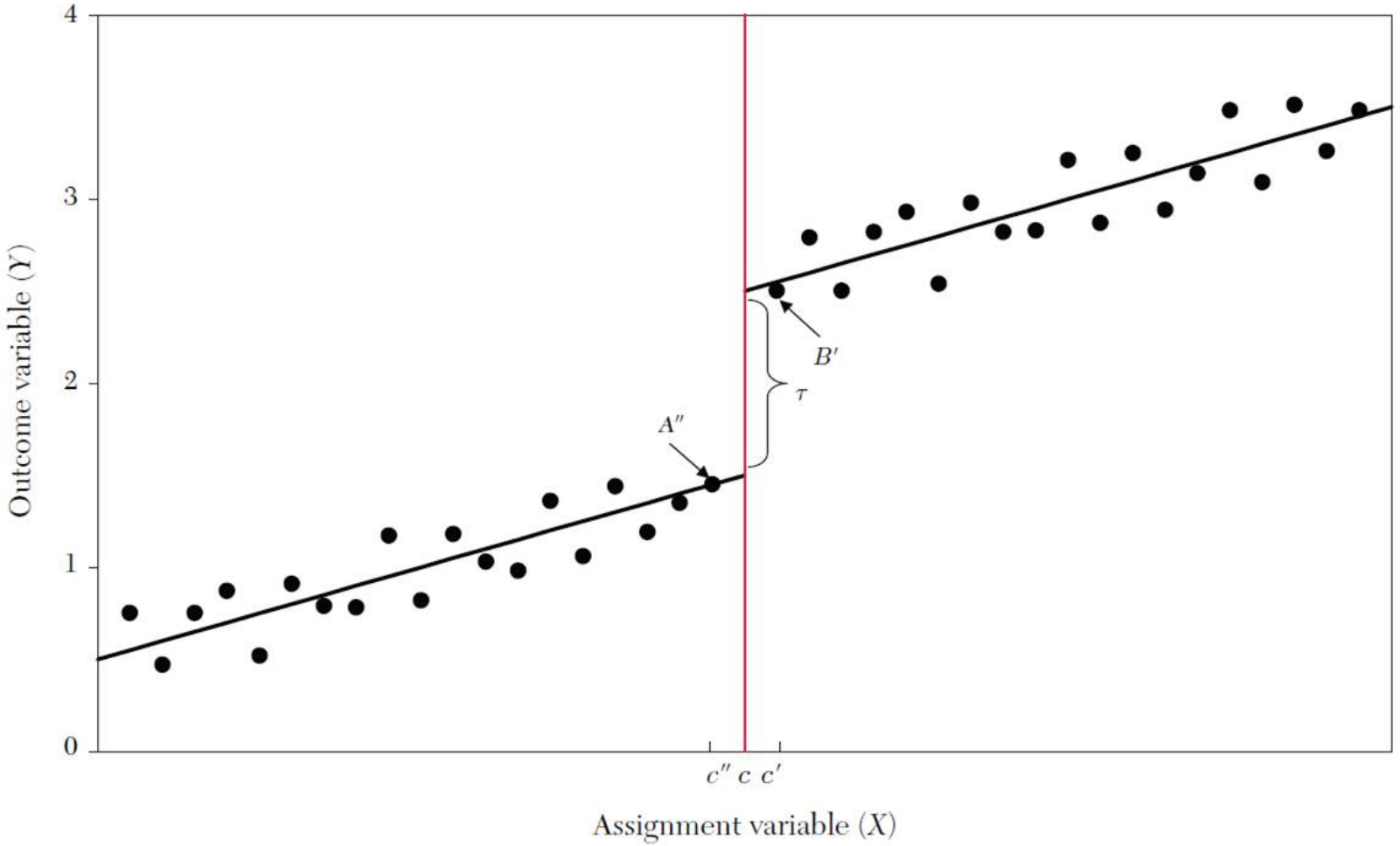
- Treatment status (D_i) is deterministic and discontinuous function of running (assignment, forcing) variable x_i

$$D_i = \begin{cases} 1 & \text{if } x_i > c \\ 0 & \text{if } x_i < c \end{cases}$$

- All individuals to the right of the cut off (c) are exposed to treatment, and all those to the left are denied the treatment.

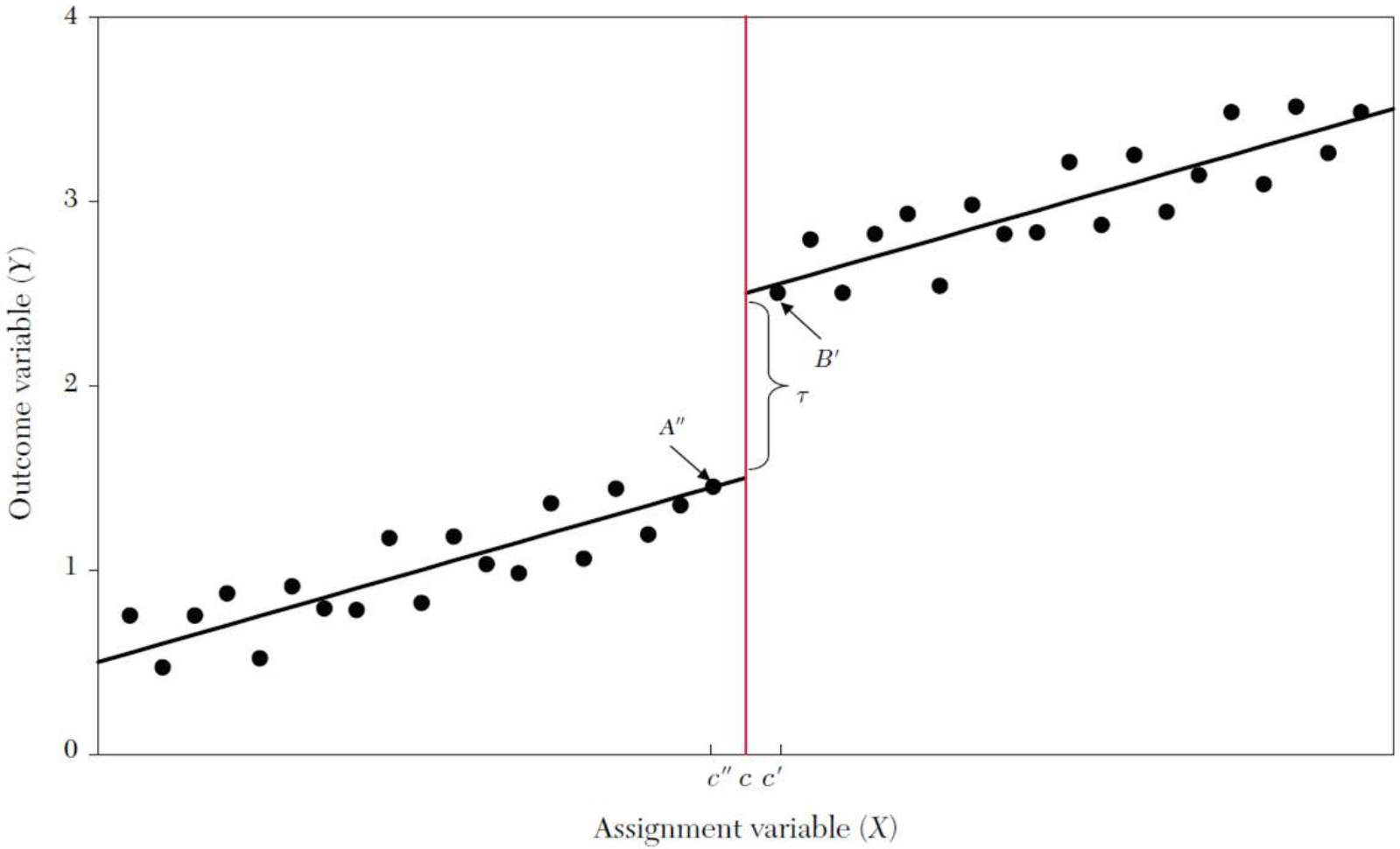
Sharp Regression Discontinuity Design

Outcome Variable and Assigment Variable



Sharp Regression Discontinuity Design

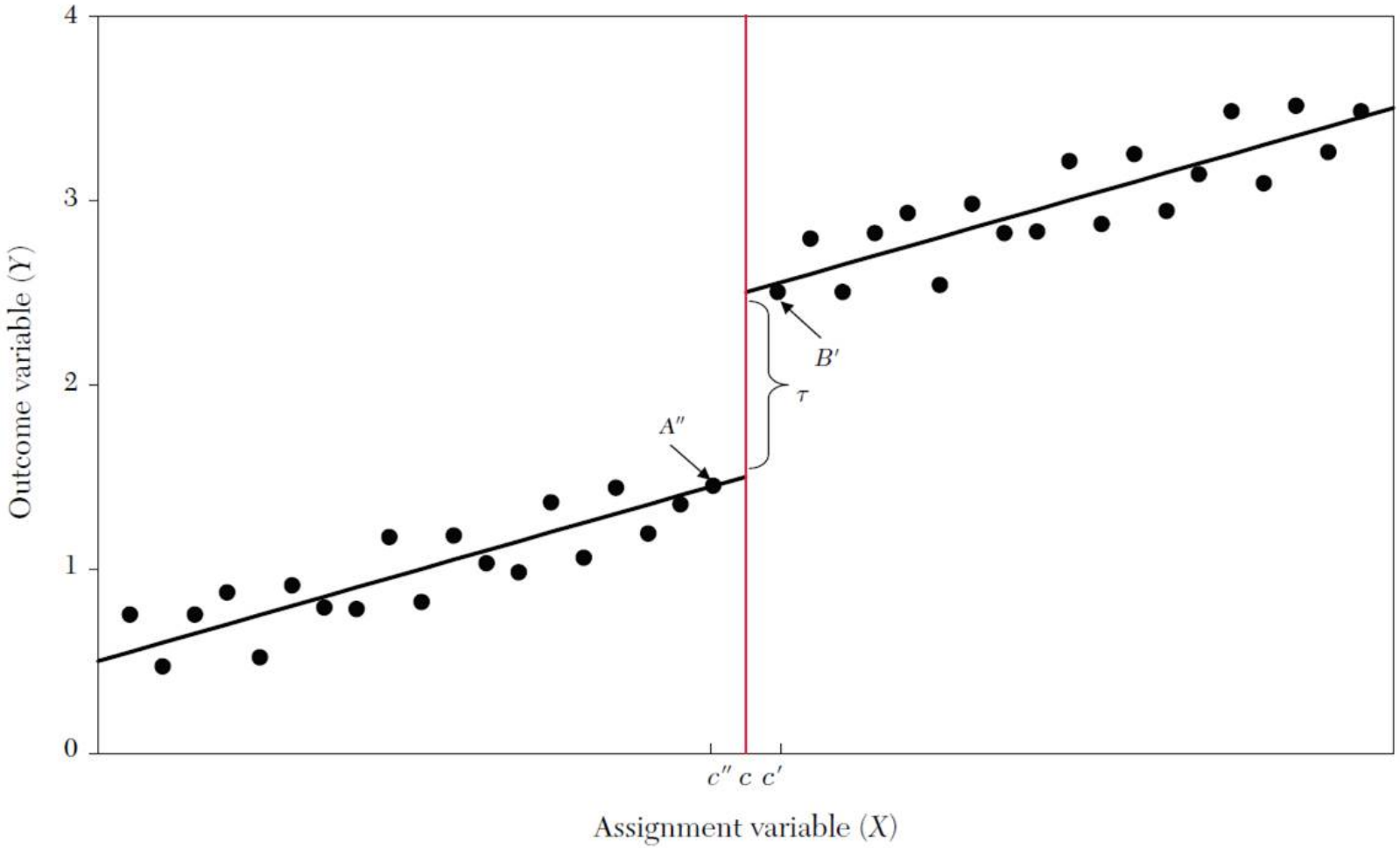
Outcome Variable and Assigment Variable



B' value of Y for individual receiving treatment

Sharp Regression Discontinuity Design

Outcome Variable and Assigment Variable

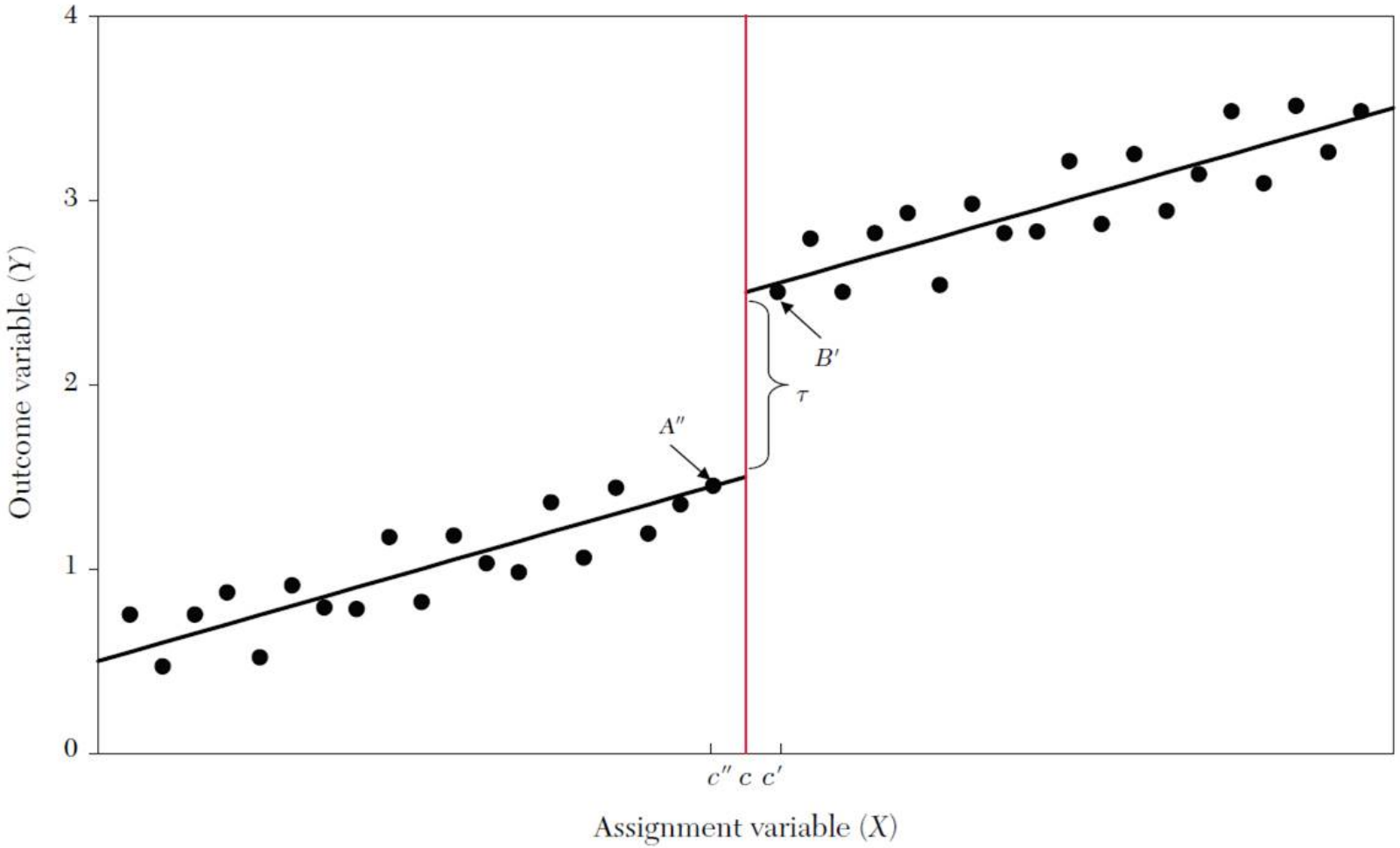


B' value of Y for individual receiving treatment

A" reasonable guess for Y for the same individual in no-treatment state

Sharp Regression Discontinuity Design

Outcome Variable and Assigment Variable



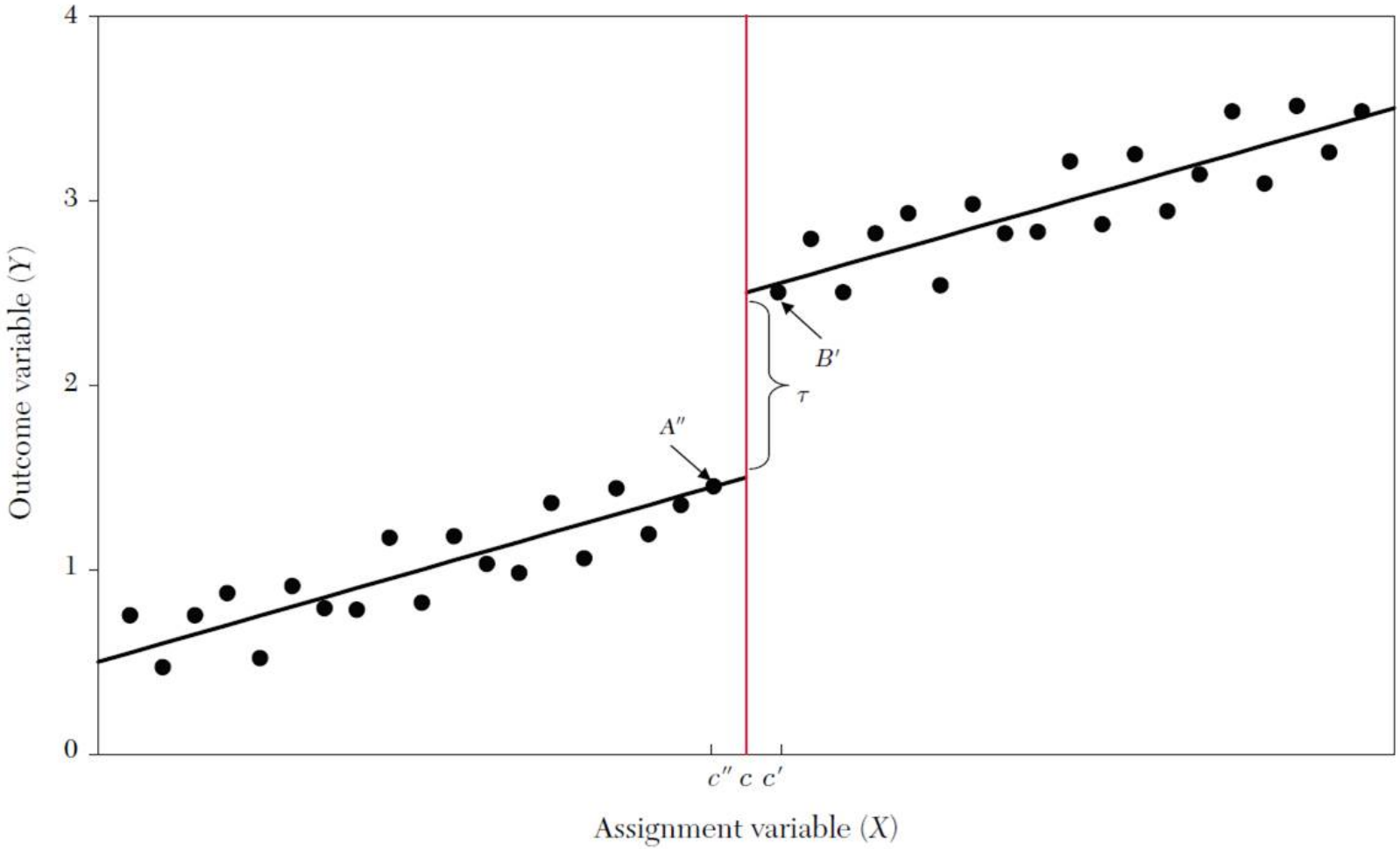
B' value of Y for individual receiving treatment
A" reasonable guess for Y for the same individual in no-treatment state
B'-A" would be the impact of treatment

How to estimate the effect of treatment in RDD?

- Inherent tradeoff in RDD:
  - The estimates are more accurate the closer we are to threshold
  - The closer to the threshold we are the less data we have
- We need to use data further away and assume a functional form for the relationship between Y and X.

Sharp Regression Discontinuity Design

Simple Linear RD Set Up



How to estimate the effect of treatment in RDD?

Simple Linear RD Set Up

- Assuming linear relationship between X (assignment variable) and Y (outcome)
  $Y_i = \alpha + \tau D_i + \beta X_i + u_i$
- All other factors determining Y need to evaluate smoothly with respect to X
- But the relationship between X and Y may be more complex...

What if the relationship between X and Y is nonlinear?

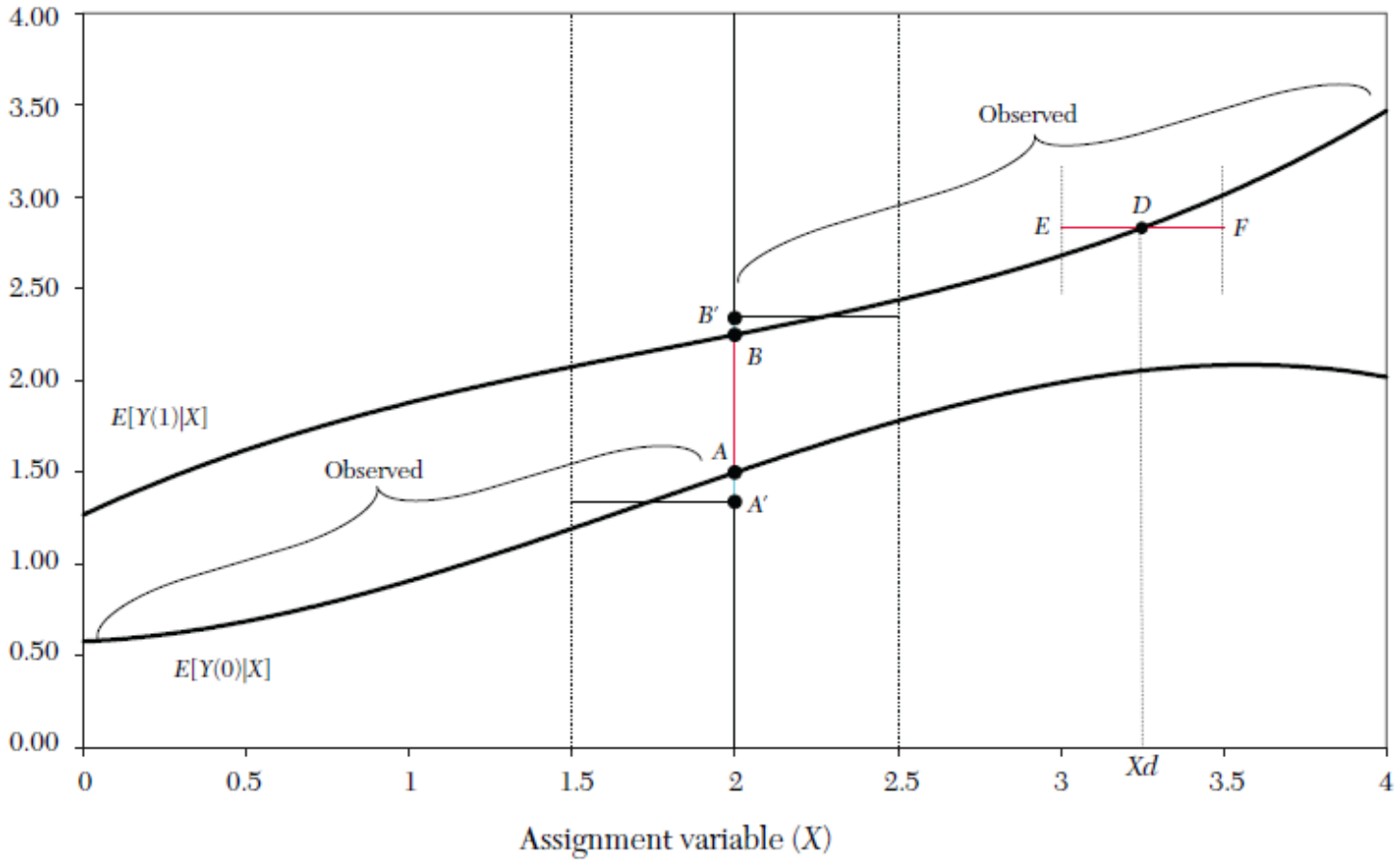


Figure 2. Nonlinear RD

How to estimate the effect of treatment in RDD?

1. Specify correct functional form

- More general functional form:
  $Y_i = \alpha + \tau D_i + f(X_i) + u_i$
- For example ρth order polynomial
  $f(X_i) = \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3 ... + \beta_\rho x_i^\rho$
- Validity of RD Estimates depends whether the nonlinear model provide an adequate description of $E[Y_i(0)|X_i]$
- There is critisisim toward use of global higher-order polynomials: Gelman and Imbens (2018).
- Alternative way: estimate the effect using data closer to cutoff (within a bandwidth)

Sharp Regression Discontinuity Design

2. Estimation closer to cutoff

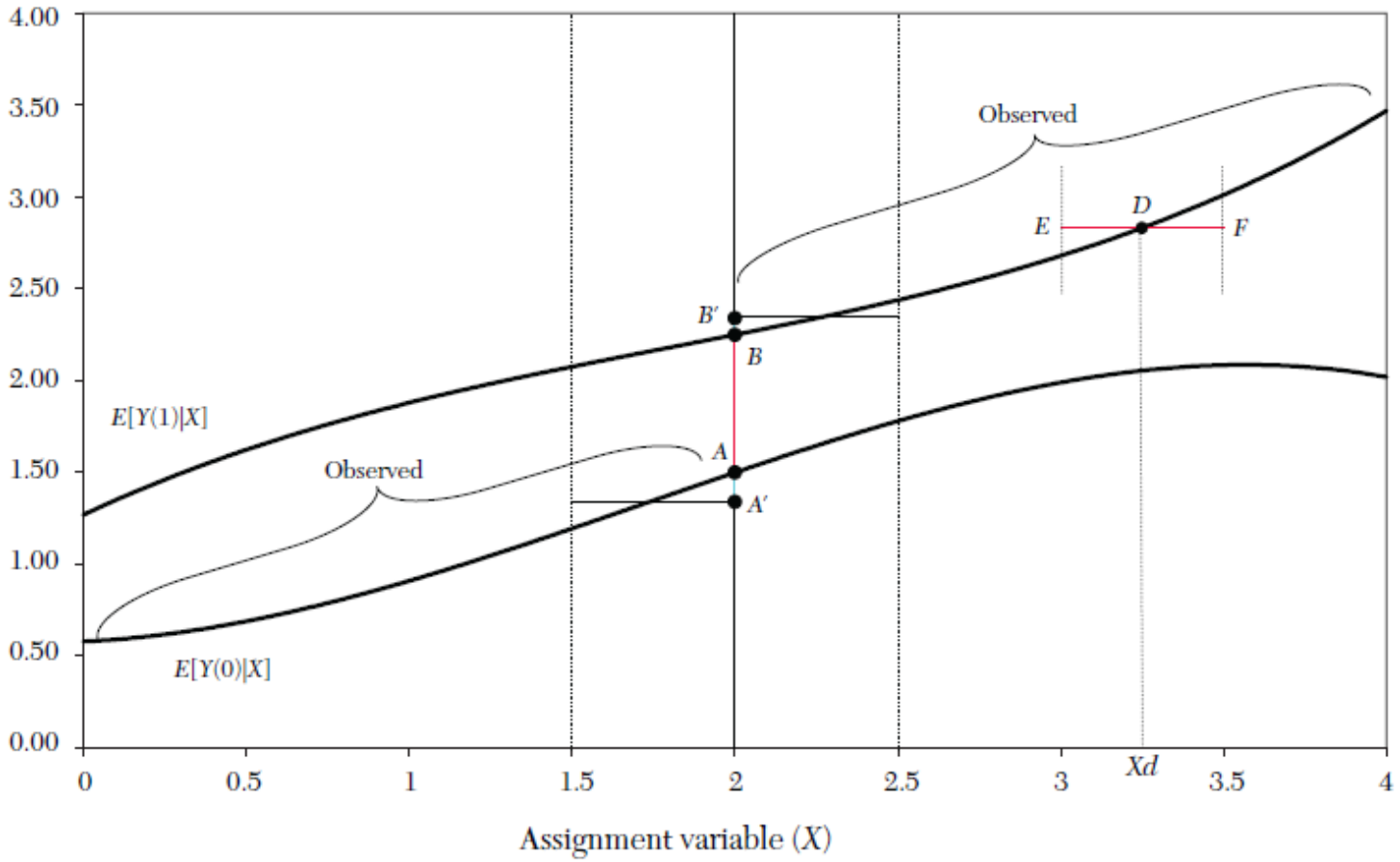


Figure 2. Nonlinear RD

- In the previous graph

$$B - A = \lim_{\varepsilon \to 0} E[Y_i | X_i = c + \varepsilon] - \lim_{\varepsilon \to 0} E[Y_i | X_i = c - \varepsilon]$$

- Which at the limit is equal to

$$E[Y_i(1) - Y_i(0) | X_i = c]$$

- This is the average effect of treatment at the cutoff c

Sharp Regression Discontinuity Design

2. Estimation closer to cutoff

- How to estimate
  $E[Y_i|X_i = c + \varepsilon]$ and $E[Y_i|X_i = c - \varepsilon]$?
- Non-parametric methods: local linear regression within a given window (bandwidth) of width h around the cutoff point
- How to choose a bandwidth h? Tradeoff in precision and bias (bias gets smaller with smaller h, but we lose precision).
- Data driven badwidth selection procedures (Calonico, Cattaneo, and Titiunik, 2014, Imbens, and Kalyanaraman, 2012)

Testing the validity of RD set up

- Few important test for the validity of RD design (no manipulation of cutoff)
  1. Density of running variable should be continuous (McCrary test)
  2. Baseline characteristics should evolve smoothly at cutoff (use these as outcome and controls)
  3. Explore the sensitivity of the results to bandwidth choice: plot the local linear discontinuity estimates agains continumm of bandwidths

Example of Sharp Design: Causal Effect of Incumbency, Lee (2008)

- Does a Democratic candidate for a seat in the U.S. House of Representatives has an advantage if his party won the seat last time?
- Treatment: Party won the seat in previous election
- Outcome: Winning the seat in next election.
- Assigment variable: x_i is the vote share margin of victory (the difference between Demogratic and Republican votes shares).

Example of Sharp Design:

Causal Effect of Incumbency, Lee (2008)

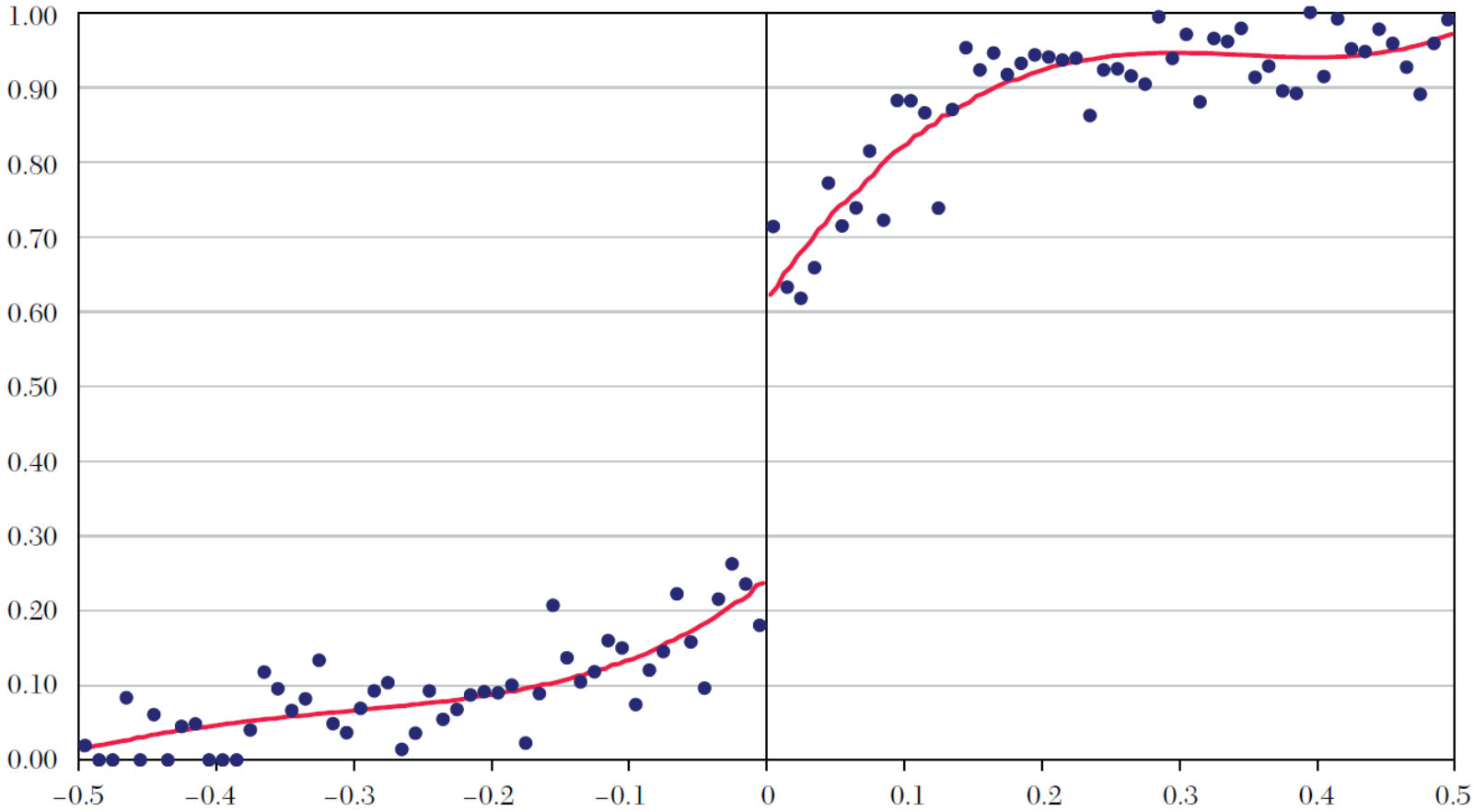


Figure 10. Winning the Next Election, Bandwidth of 0.01 (100 bins)

Assigment variable: Difference between Demogratic and Republican votes shares in previous election.

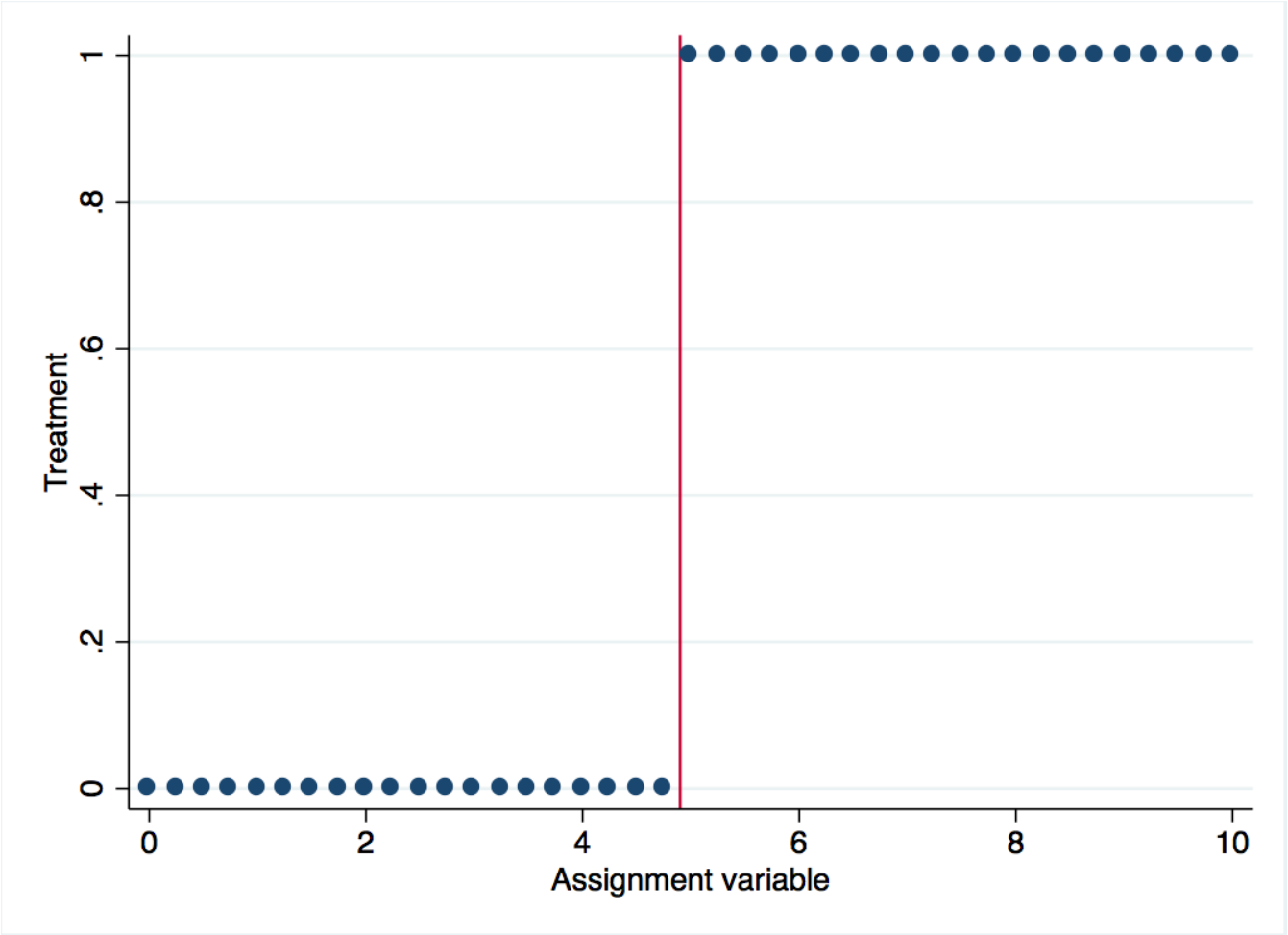
Example of Sharp Design:

Causal Effect of Incumbency, Lee (2008)

- Imcumbency seems to raise party re-election probabilities by 40 %
- Checks
  - How sensitive the results are to specification (polynomial) and choice of bandwidth
  - Is there bunching in distribution of x near the cutoff
  - Discontinuities in pretreatment covariates

Sharp Design

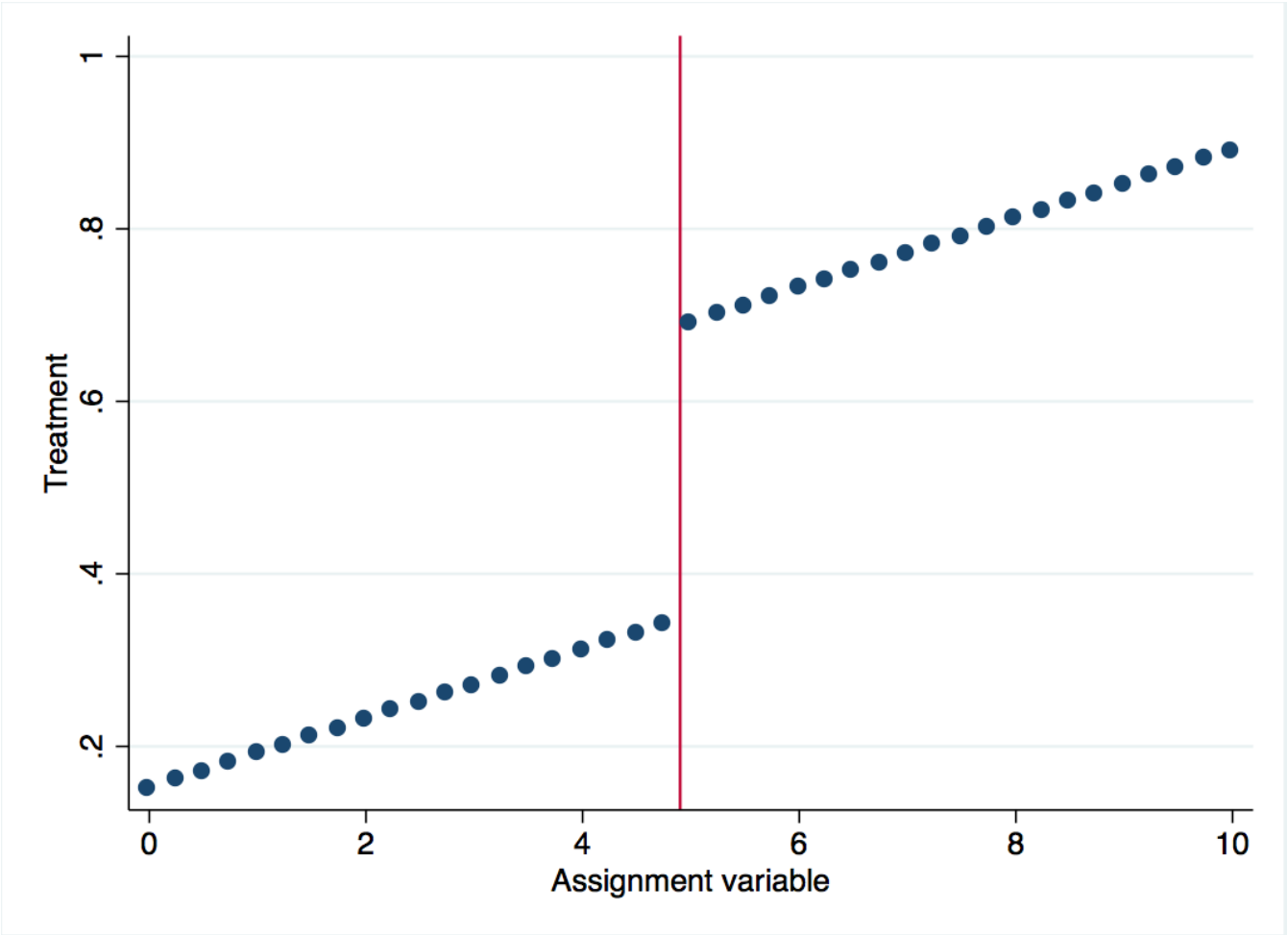
Those whose value of the assigment variable is above threshold level receive the treatment, others not



Treatment is received if the value of the test is above 5.

Fuzzy Regression Discontinuity Design

The probability of treatment jumps at the threshold



Probability of treatment jumps discontinuously at value 5.

Fuzzy Regression Discontinuity Design

- In "Sharp" RD Design probability of treatment jumps from 0 to 1 when X crosses the threshold c
- In "Fuzzy" RD Design design there is a smaller jump in the probability of assignment to the treatment

$$Pr(D_i = 1|x_i) = \begin{cases} g_1(x_i) & \text{if } x_i \geq c \\ g_0(x_i) & \text{if } x_i < c \end{cases}$$

Wald Estimator in Fuzzy RD

- A treatment effect can be recovered by dividing the jump in the relationship between Y and X at threshold (reduced form) by the jump in the treatment propensity at threshold (first stage)

$$\tau = \frac{\lim_{\varepsilon \to 0} E[Y_i | X_i = c+\varepsilon] - \lim_{\varepsilon \to 0} E[Y_i | X_i = c-\varepsilon]}{\lim_{\varepsilon \to 0} E[D_i | X_i = c+\varepsilon] - \lim_{\varepsilon \to 0} E[D_i | X_i = c-\varepsilon]}$$

- Similar to IV set up: The rule is like an instrument that is used to created exogenous variation in treatment propensity...
- We identify the effect for individual at threshold.

Example of "Fuzzy Design":

Abdulkadiroglu, Angrist, and Pathak, Econometrica 2014

- What is the effect of attending an elite high school on student achivement?
- Competitive elite schools in Boston and New York that select their students based on admissions tests
- Autors use these entry thresholds to estimate the effect of attending an elite school on test scores
- Parallel situation in many other locations (Finland)

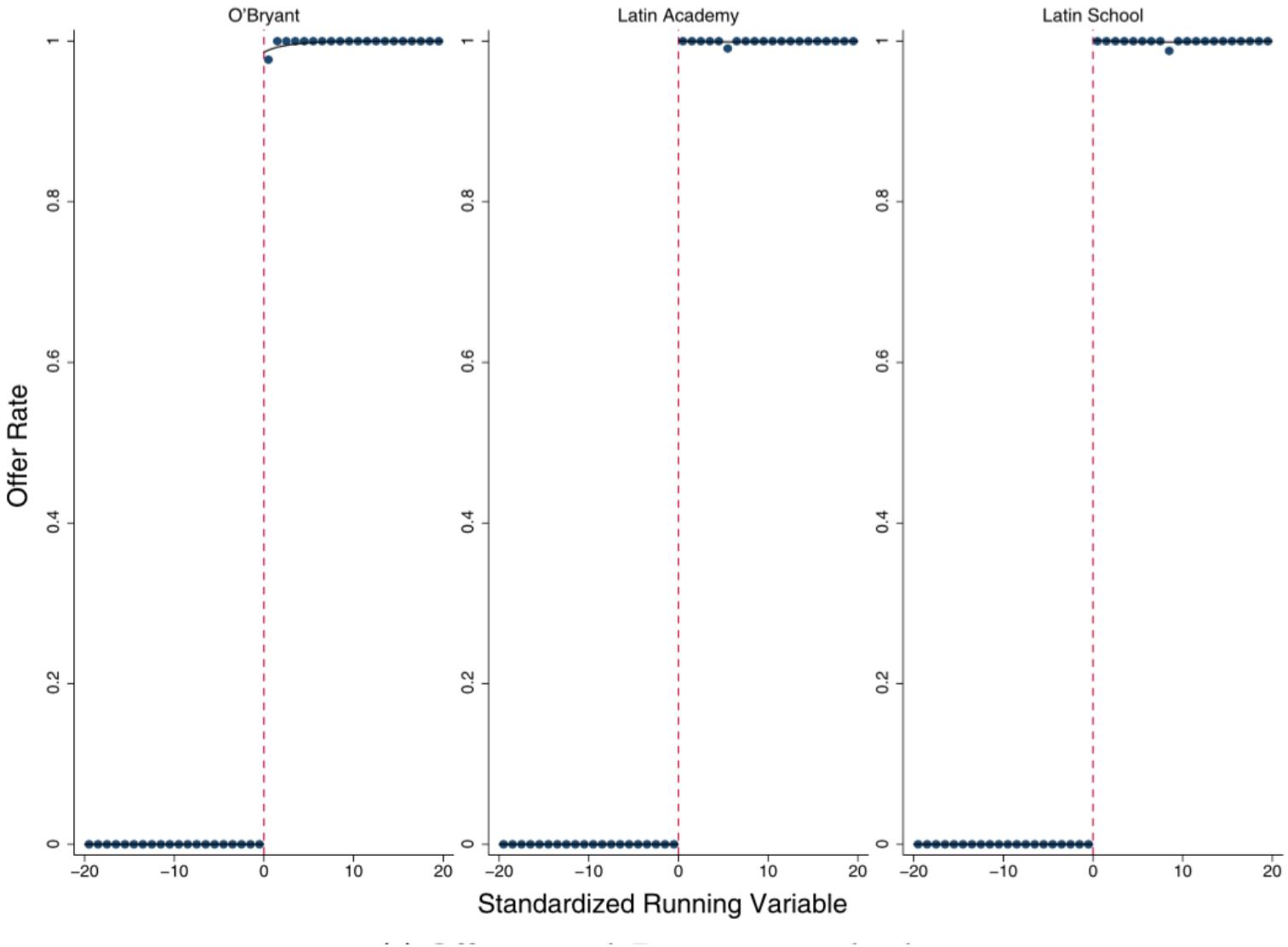
Example of "Fuzzy Design":

Abdulkadiroglu, Angrist, and Pathak, Econometrica 2014

- Probability of receicing an offer to elite school jumps from 0 to one 1 at threshold (cut off)
- However, the probability of enrollment may not jump from 0 to 1
  - Some applicants receive multiple offers and only choose to enroll in the preferred school
  - Rejected slots will be filled from the waiting list below the threshold

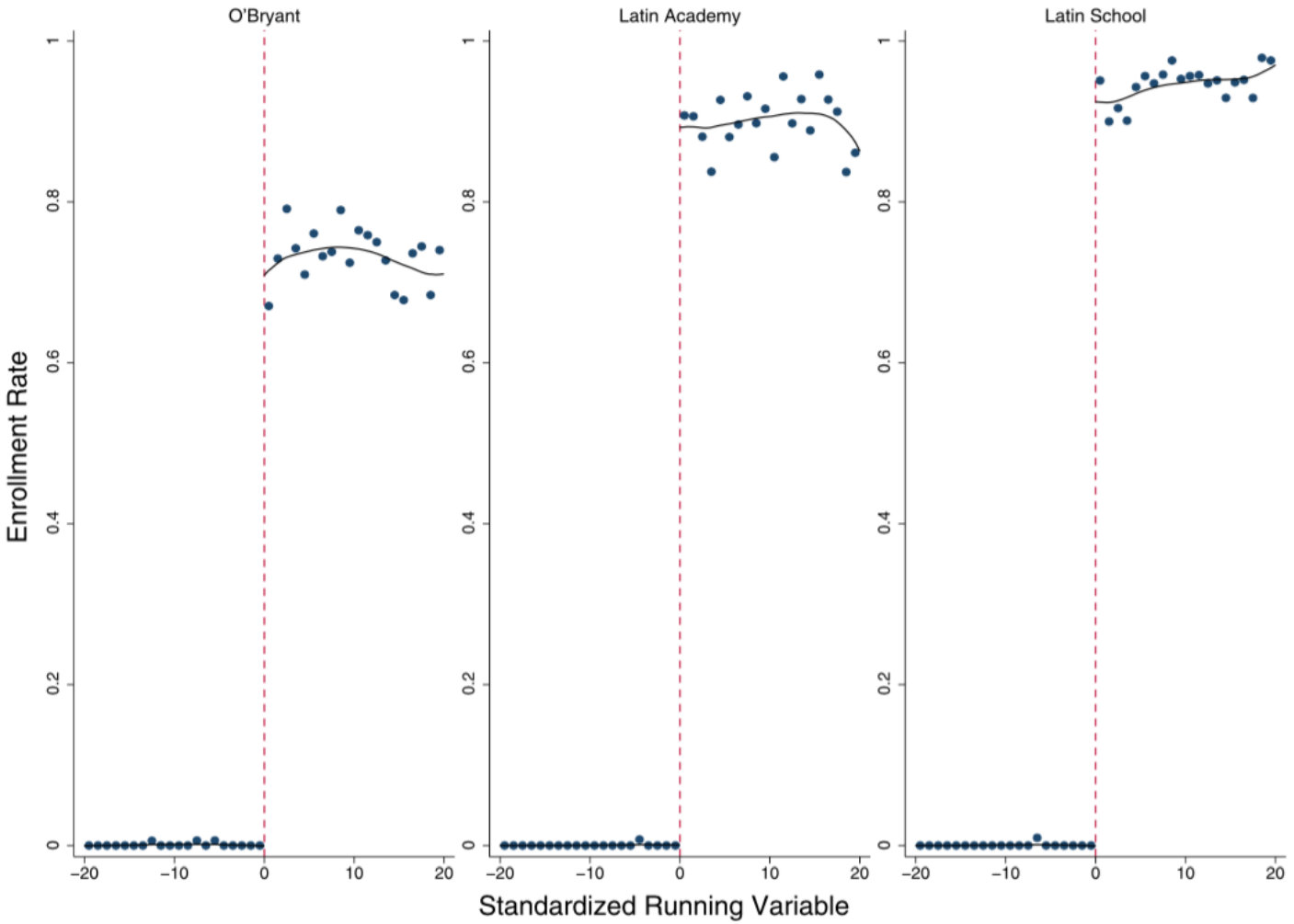
Offers at each Boston exam school

Abdulkadiroglu, Angrist, and Pathak, Econometrica 2014



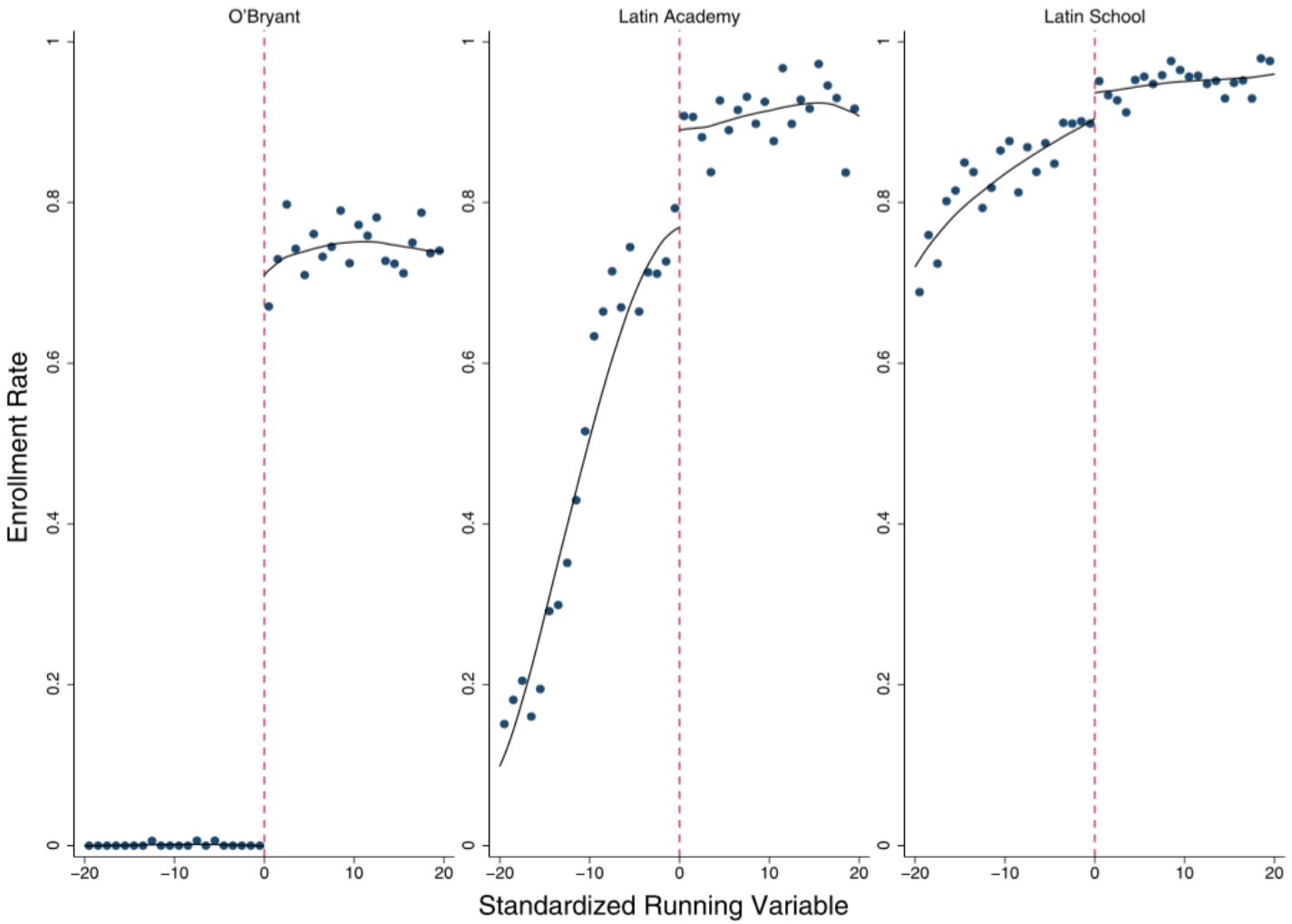
Enrollment at each Boston exam school

Abdulkadiroglu, Angrist, and Pathak, Econometrica 2014



Enrollment at any Boston exam school

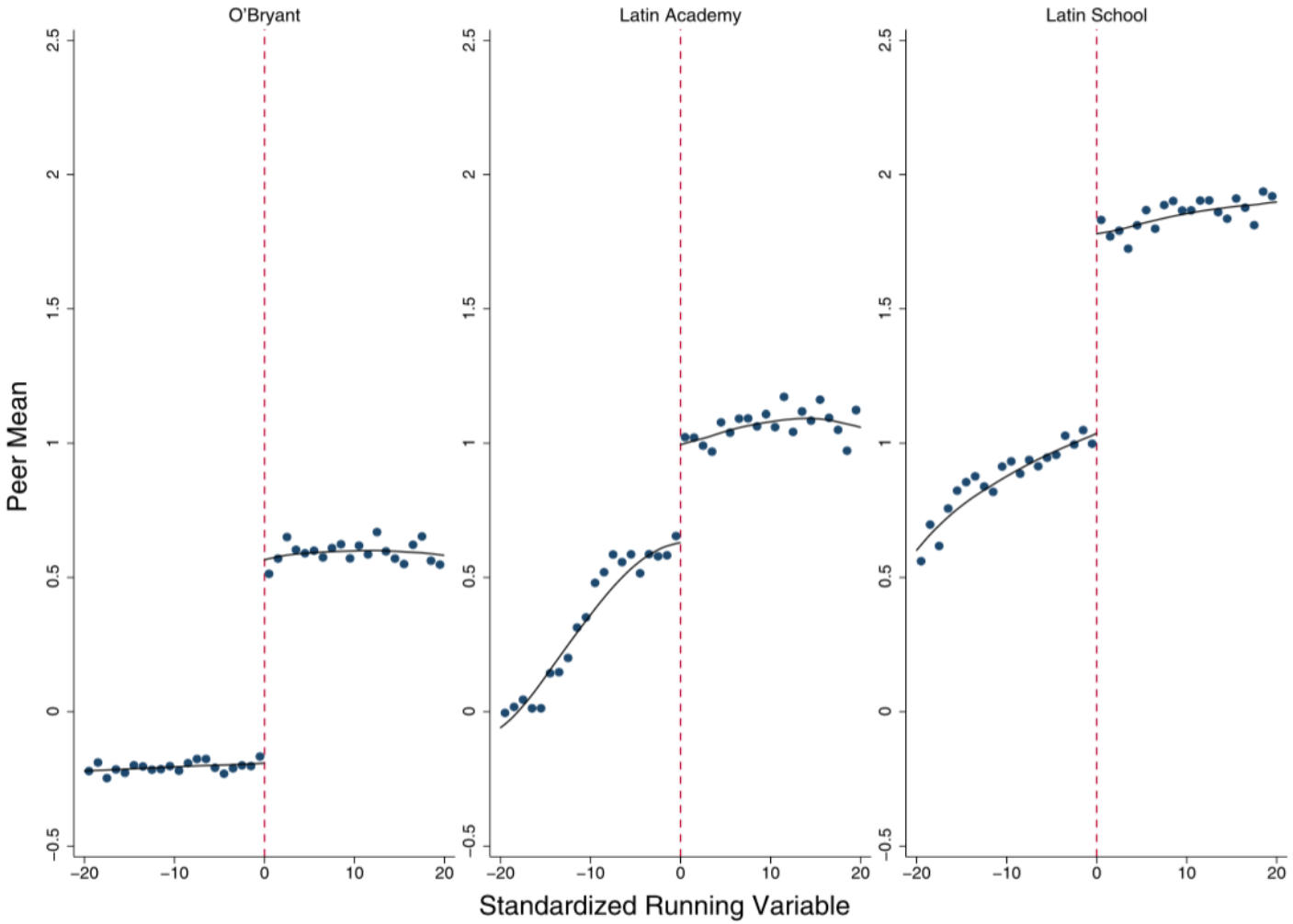
Abdulkadiroglu, Angrist, and Pathak, Econometrica 2014



(c) Enrollment at any Boston exam school

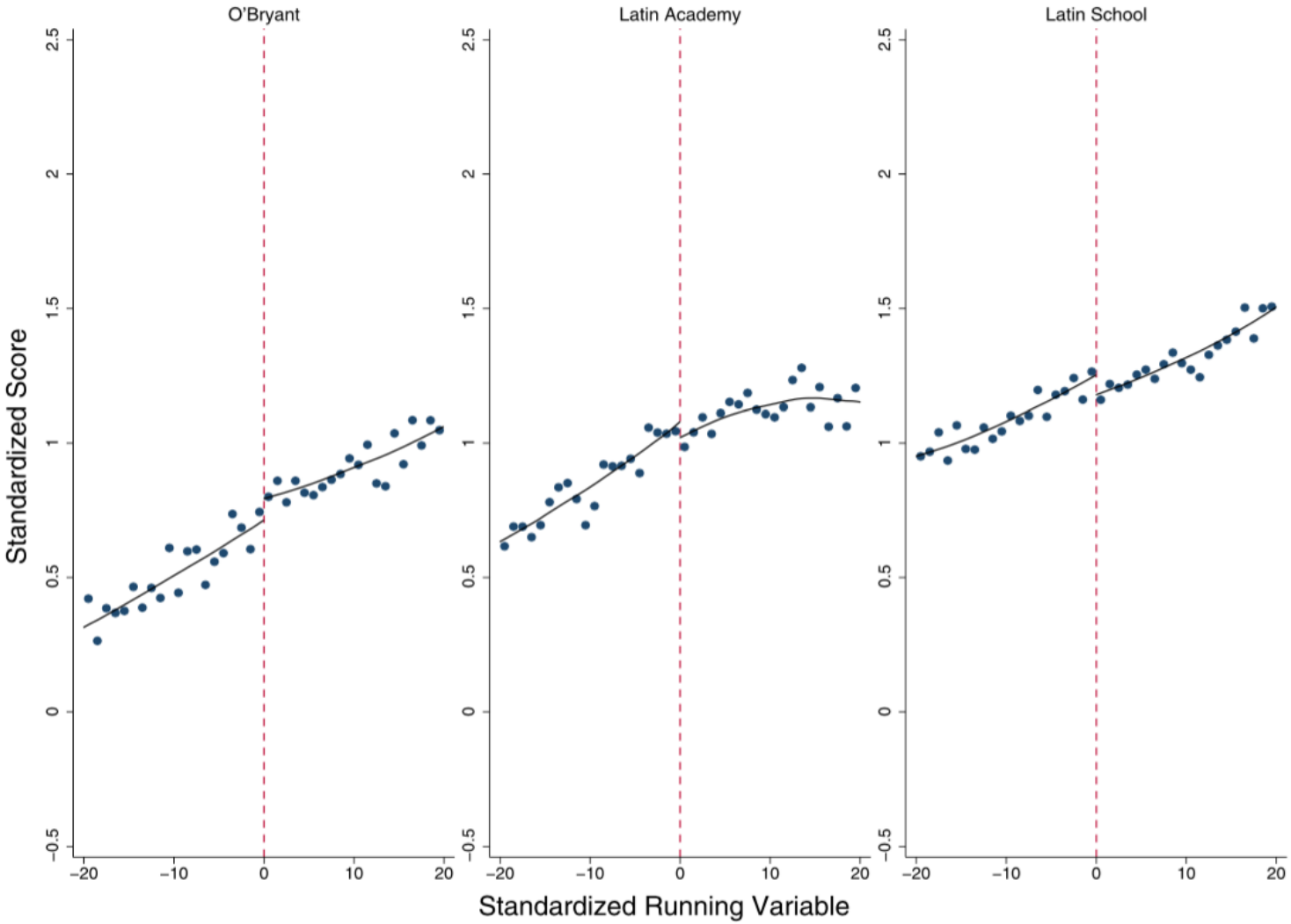
Baseline peer math score at Boston exam schools

Abdulkadiroglu, Angrist, and Pathak, Econometrica 2014



10th grade math at Boston exam schools

Abdulkadiroglu, Angrist, and Pathak, Econometrica 2014



Example of "Fuzzy Design":

Abdulkadiroglu, Angrist, and Pathak, Econometrica 2014

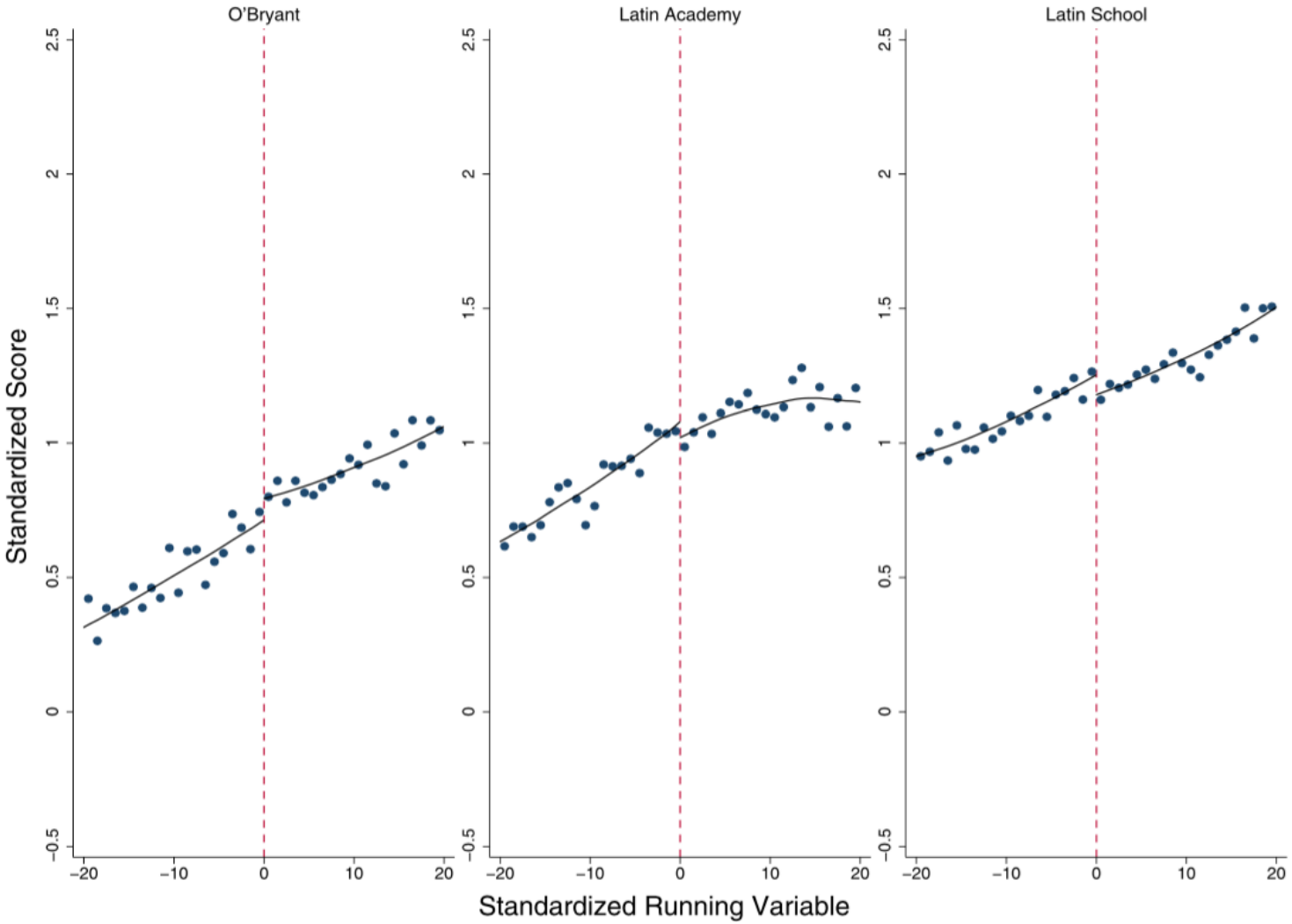
- We want to estimate the effect of enrolling to elite school on test scores
- The reduced form can be written as

$$Y_i = \alpha + \rho D_i + \beta R_i + u_i$$

- Where D_i is the indicator for being above threshold (have an offer) and R_i is the running variable

10th grade math at Boston exam schools for 7th and 9th grade applicants

Abdulkadiroglu, Angrist, and Pathak, Econometrica 2014



Example of "Fuzzy Design":

Abdulkadiroglu, Angrist, and Pathak, Econometrica 2014

- No visible reduced form effect
- Given this, it is not surprising that 2SLS estimates are approximately zero for all outcomes
- Elite schools have no effect on achievement
- RDD is local treatment effect: What this imply for the interpretation