API-202B Empirical Methods II

Session #19:

Quasi-experimental methods: Regression discontinuity

miguel_santos@hks.harvard.edu
@miguelsantos12

Where we are and what lies ahead in API202B?

API202B

Introduction
- Evidence and policy (1) Jan 22nd
- Causality and validity (2) Jan 24th

Linear regression
- Multivariate regression (3-4) Jan 29th - 31st
- Dummy variables (5) Feb 5th
- Omitted variables (5-6) Feb 7th

Non-linear regression
- Binary dependent (7) Feb 12th
- Quadratic and logarithms (8) Feb 14th
- Interactions (9) Feb 21st

Midterm Exam

Experiments
- Experiments RTC Debate (10-11) Feb 28th - Mar 5th

Quasi-experiments
- Instrumental variables (12-13-14) Mach 7th 19th 21st
- Fixed effects (15) March 26st – April 2nd
- Differences in differences (16-17) Mar 28th-Apr 2nd
- Final Project Work Session (18) April 4th
- **Regression discontinuity (19) April 9th**

Review
- Review - Final exercise discussion (20) – April 11th

FP

Final Exam (May 4th)

Today's class

- Introduction to regression discontinuity
- Example: Does merit aid affect college enrollment?
- Takeaways
- Survey on topics to be reviewed

Regression discontinuity

- Today we will discuss our final quasi-experimental technique, called **regression discontinuity** (RD; regression discontinuity design)

- RD can be used when treatment is based on a continuous variable called the **running variable** (or assignment variable, or forcing variable)

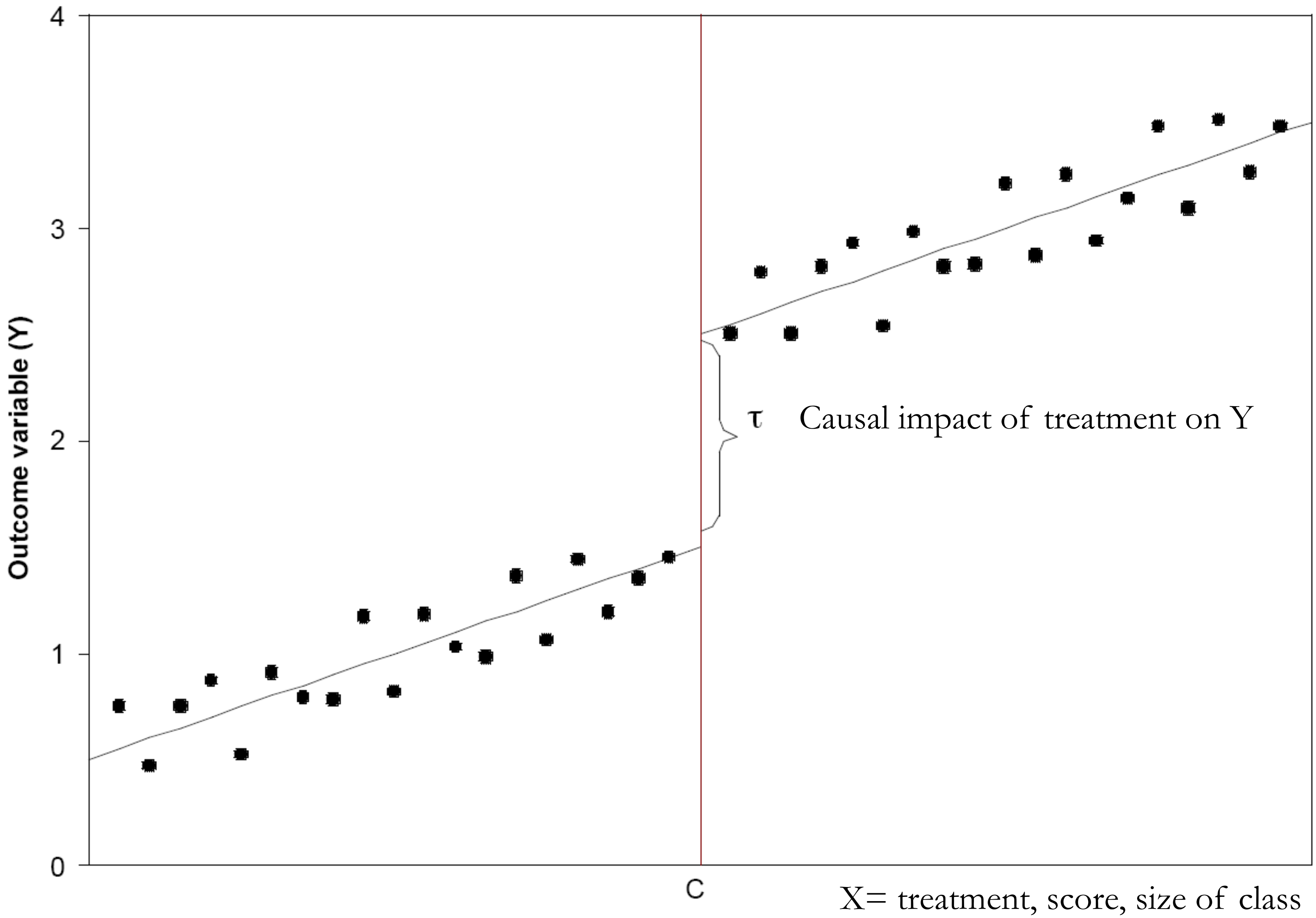
  - If the value of the assignment variable falls on one side of the cutoff, the person is assigned to the "treatment" group

  - If the value of the assignment variable falls on the other side of the cutoff, the person is assigned to the "control" group

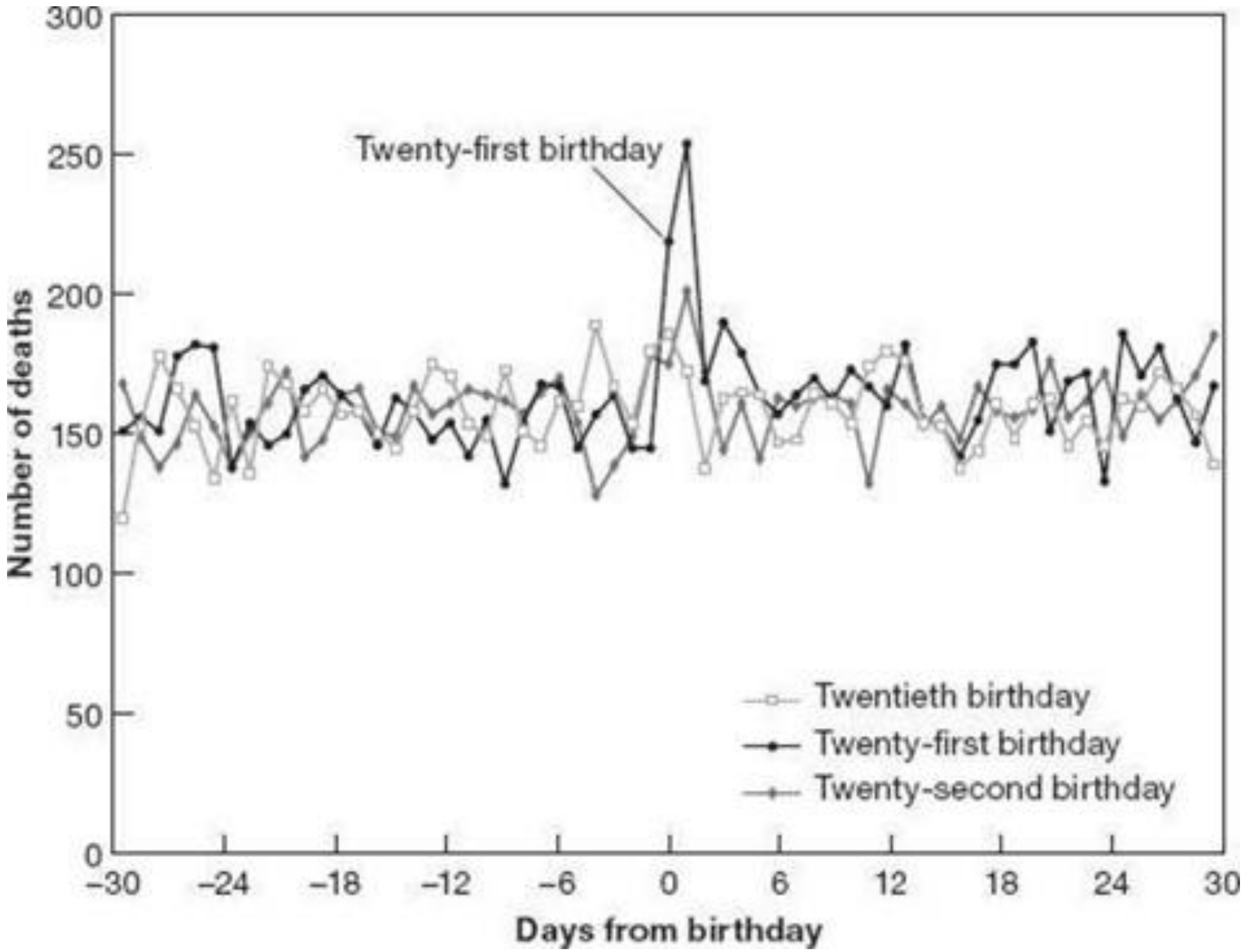
Regression discontinuity

- Assignment variable can be anything defined prior to treatment
- Examples include:
  - Assigning social welfare benefits on the basis of income
  - Opening new classrooms in classes where certain size is reached
- In many settings the assignment variable is a measure of need or merit
- Basic idea: Check for a noticeable difference (discontinuity) in the outcome between nearly identical people who fall on either side of arbitrary threshold
- Big advantage over diff-in-diff: No need for multiple time periods!

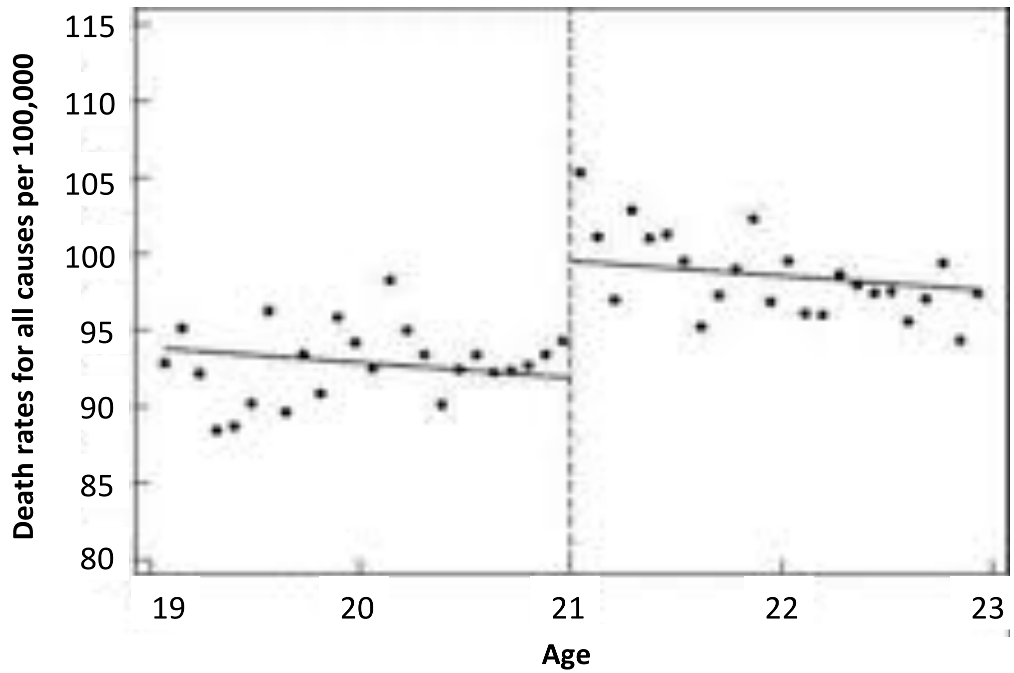
Regression discontinuity



Regression discontinuity: Minimum legal drinking age



Regression discontinuity: Minimum legal drinking age



- Treatment is a deterministic function of the **running variable age** (sharp RD):

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

- A simple RD analysis of the MLDA estimates causal effects using a regression:

$$\overline{M}_a = \beta_0 + \beta_1 D_a + \beta_2 age + \varepsilon$$

Regression discontinuity: Minimum legal drinking age

- To the left of the cut-off the regression line becomes:

$$\bar{M}_a = \beta_0 + \beta_2 age + \varepsilon$$

- To the right of the cut-off the regression line becomes:

$$\bar{M}_a = (\beta_0 + \beta_1) + \beta_2 age + \varepsilon$$

- The coefficient of interest in this case is: β_1

- How would that equation look if we were to control for a possible non-linear relationship between age and death? Add a quadratic term we would need to add a square term, or even a higher order polynomial (seems more appropriate)

$$\bar{M}_a = \beta_0 + \beta_1 D_a + \beta_2 age + \beta_3 age^2 + \varepsilon, \quad \text{or even}$$

$$\bar{M}_a = \beta_0 + \beta_1 D_a + \beta_2 age + \beta_3 age^2 + \beta_3 age^3 + \varepsilon$$

Example: Merit aid and college choice

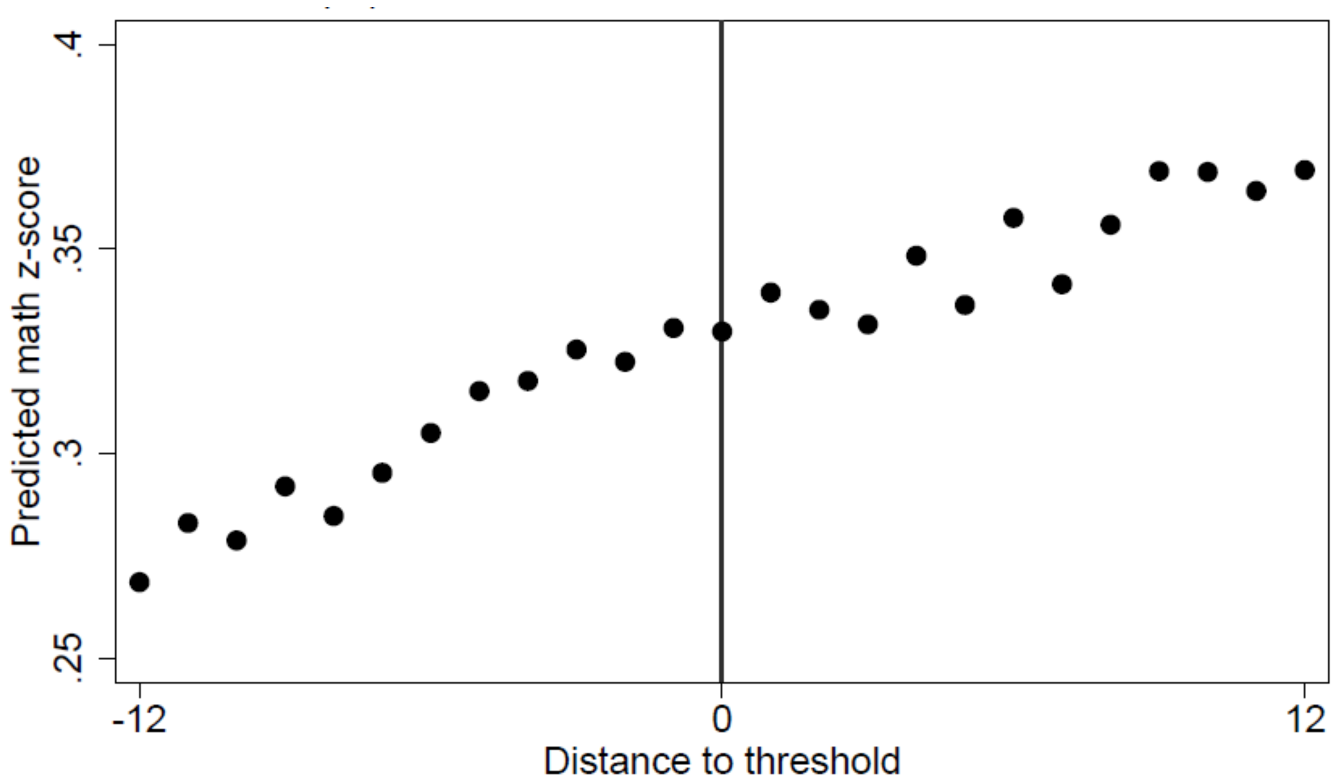
- Policy question:
  - How does financial aid based on academic merit affect college choice?
- The Adams Scholarship is a Massachusetts merit aid program that:
  - Waives tuition at in-state public colleges ("Adams colleges")
  - If a student's 10th grade test score is at or above the 75th percentile
- We have data on students' test scores and college choices
- Why might simple comparisons of eligible and ineligible students' college choices be biased?
- Comparing outcomes of those eligible and ineligible for the Adams Scholarship would confound the impact of the scholarship with the fact that eligible students have higher academic skills, that in turn might be the mirror of a whole set of more favorable characteristics such as family environment, etc.

Example: Merit aid and college choice

- To get rid of OVB: Exploit the eligibility threshold
- Students just above and just below the threshold should be nearly identical in all ways except for scholarship eligibility
  - Whether student gets one question right feels like coin toss (i.e. random)
- RD is sometimes called "local randomization" (i.e. random near threshold)
- Students at the 75th percentile have nearly the same academic skills as those at the 74th percentile
  - Should also be similar in terms of gender, race, family background, etc.
- As long as they can't control the side of threshold they fall on

Covariate balance test

- Just like RCTs, we can check whether the threshold generates two nearly identical groups!
- Do the treatment and control groups 8th grade math scores look balanced?



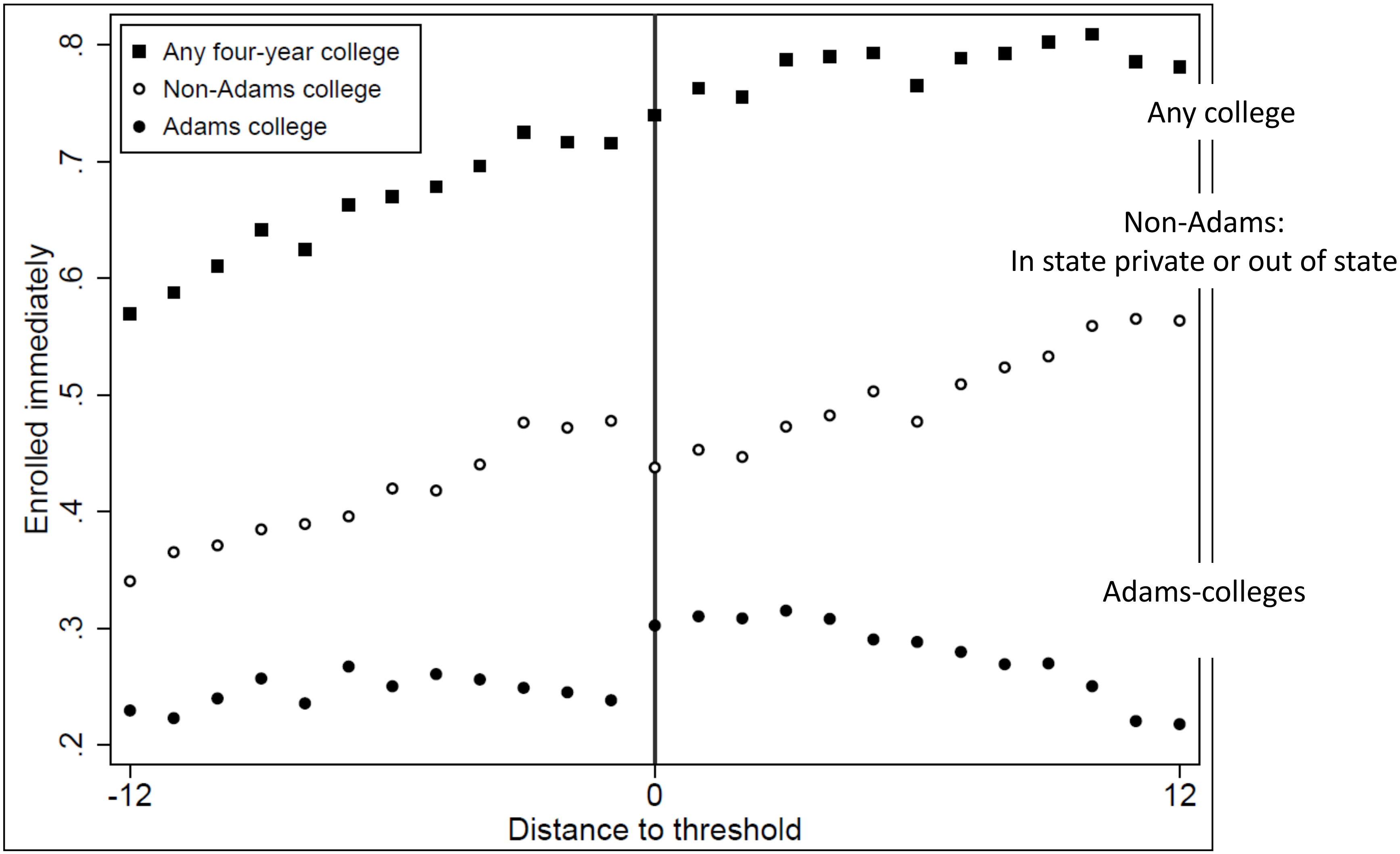
- Graph would look similarly "smooth" if Y= income, race, gender, etc.

Covariate balance test

- Why is it important that individuals not be able to control the assignment variable? (Assignment variable is not susceptible to being manipulated)
  - We do not want people to choose their treatment status, because T/C groups would no longer be balanced
    - In this case, the earliest cohort of students of the Adams Scholarship took their MCAS test prior to the announcement of the existence of the Adams Scholarship (at the time of test administration, student have no idea in which percentile they are)
    - MCAS exams are centrally scored and raw scores transformed into scaled scores via an algorithm unknown to students, families or even teachers.
- If eligibility were based on having a high school GPA of 3.5 or higher, would regression discontinuity yield internally valid estimates?
  - No. Again, students differ in many characteristics that are driving their GPA scores, other than treatment.

Virtual estimation of impacts

- Once we are satisfied that students across threshold look similar, we can examine mean outcomes by graphing them by distance to the threshold



Virtual estimation of impacts

- Students just above threshold are 7% likely to enroll in in-state public colleges than students just below threshold.

- Students just above threshold are -6% likely to enroll in other four-year colleges than students just below threshold.

- Students just above threshold are not more (impact visually hard to distinguidh form zero) likely to enroll in any four-year college than students just below threshold.

Regression estimation of impacts

- Regression version of eyeball test is OLS of the form:

$$College = \beta_0 + \beta_1 Above + \beta_2 Distance + \beta_3 Above * Distance + \ \varepsilon$$

College = 1 if student enrolls in a particular type of college.

Eligible = 1 if a student's test score qualified for the scholarship.

Distance measures student's distance from the threshold (test score points)

- β_1 is difference in outcome between those just above and below threshold
- Because that regression draws two lines, one on either side of the threshold:
  - To the left: $College = \beta_0 + \beta_2 * Distance + \ \varepsilon$
  - To the right: $College = (\beta_0 + \beta_1) + (\beta_2 + \beta_3) * Distance + \ \varepsilon$

Regression estimation of impacts

- Here are the regression results:

	Adams college, four-year (1)	Non-Adams college, four-year (2)	Any four-year college (3)
Enrolled immediately	0.069***	-0.060***	0.009
	(0.010)	(0.010)	(0.008)
$\bar{Y}$	0.238	0.478	0.716

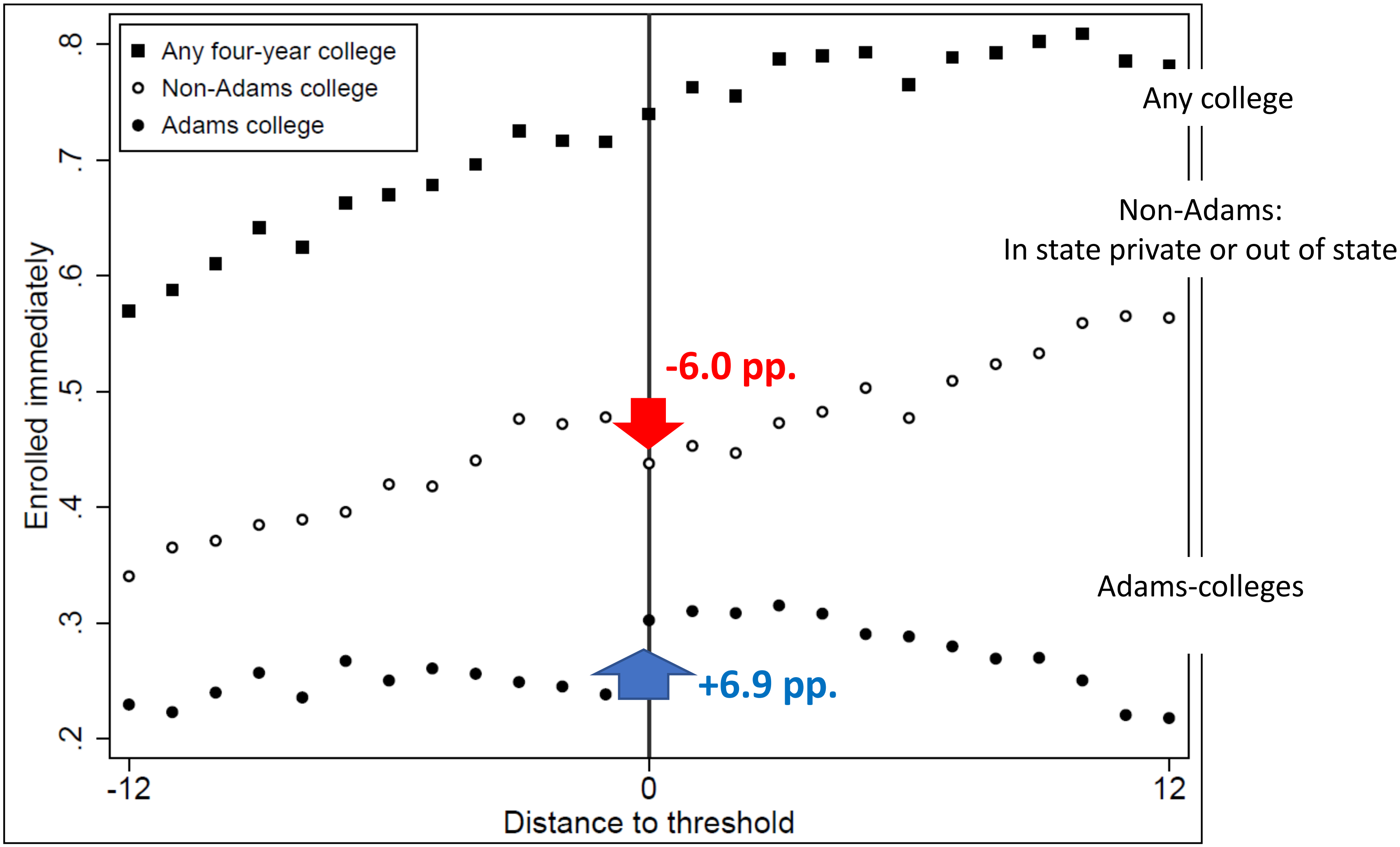
- Do these match our eyeball test?

  Roughly yes! (credit to Mark and Juan Felipe)

- Does this scholarship seem like a good use of public funds?
  - Good question. It does achieve the goal of increasing the share of good students that remain in state, but at the cost of getting them in schools of lower quality, that have less resources available, where students take more time to graduate, and completion rates are lower (the latter part, however, belongs to the last paper and was not discussed in class). As a Mass policy maker, you might consider that a success – as in response to the program 6%-7% of students switched to stat colleges, but your are doing that at a cost to them.

Virtual estimation of impacts

- Once we are satisfied that students across threshold look similar, we can examine mean outcomes by graphing them by distance to the threshold



Local Average Treatment Effects

- Instrumental variables and regression discontinuity have a common property:
  - The estimates produced have a marginal interpretation, i.e., they are driven by a subset of the sample studied (LATE)
  - The estimate of β_1 has a marginal interpretation because it is generated by students near the eligibility threshold
- Important to ask whether causal impacts would be different for individuals far from the threshold
- Would merit aid affect lower scoring students differently?
- Merit aid does not seem to have stimulated lower scoring students in particular, as the slope of the line to the left of the threshold seems pretty flat.

Takeaways

- Regression discontinuity is a powerful tool for causal inference
  - It allows comparison of two groups who are nearly identical except for the fact they fall on different sides of an arbitrary threshold
  - It's as close to an RCT as a non-RCT can get
- Many public policies involve eligibility thresholds!
- If you are convinced that those on both sides of the policy threshold are similar but for the policy treatment itself, RD can provide causal estimates
  - RD estimates are LATEs driven by those near the threshold
- Today's example was a **sharp RD**, where the threshold determines 100% of treatment status (i.e. below = no treatment, above = treatment)
- There are also cases of **fuzzy RD**, where the threshold changes the probability of treatment, but not from 0 to 100% (can be addressed by IV!)

Vocabulary

- Regression discontinuity
- Running variable
- Covariate balance test