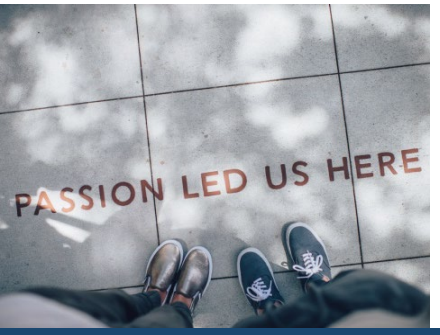


Lab 5: Evaluating Education Policy with Regression Discontinuity Design

Gregory Bruich, Ph.D.

HARVARD UNIVERSITY

OPPORTUNITY INSIGHTS

Lab 5: Academic Probation Policy at a Large Canadian University

- In today's lab, we will discuss *regression discontinuity design* in the context of **evaluating** education policies and programs

- How do we **validate** a regression discontinuity research design?

- **Coding exercise:** estimate the causal effect of academic probation on on-time college graduation using this method *(Lindo, Sanders, and Oreopoulos 2010)*

Key Lessons from Lab 5

- Substantive question: does academic probation help students graduate on time?
- The method is regression discontinuity design (RDD)
- Key methodological tools:
  1. Binned scatter plots to assess smoothness of predetermined characteristics
  2. Using histograms to check for manipulation of the running variable
  3. Using binned scatter plots to visualize regression discontinuity design
  4. Multivariable regression with interaction terms for treatment effect estimation for a regression discontinuity research design

Excerpt of Letter Set by University to Students Being Placed on Probation

Dear < first name >:

Your academic record indicates that you are experiencing challenges with your studies at xxxxxxxxxxx. As a result, you have been placed "On Probation" at the end of the xxxxxxx session. "On Probation" is an academic status applied to a student if he or she:

1. Is having difficulty achieving a term average of at least 1.60 GPA
2. Is having difficulty meeting performance expectations and/or deadlines as outlined by the course instructor.
3. Is having difficulty achieving the minimum grades required for graduation.

A student who at the end of any session during which they are on probation has a sessional GPA of less than 1.6 shall be suspended. Therefore, it is imperative that you seek assistance to improve your academic standing to avoid further sanction.

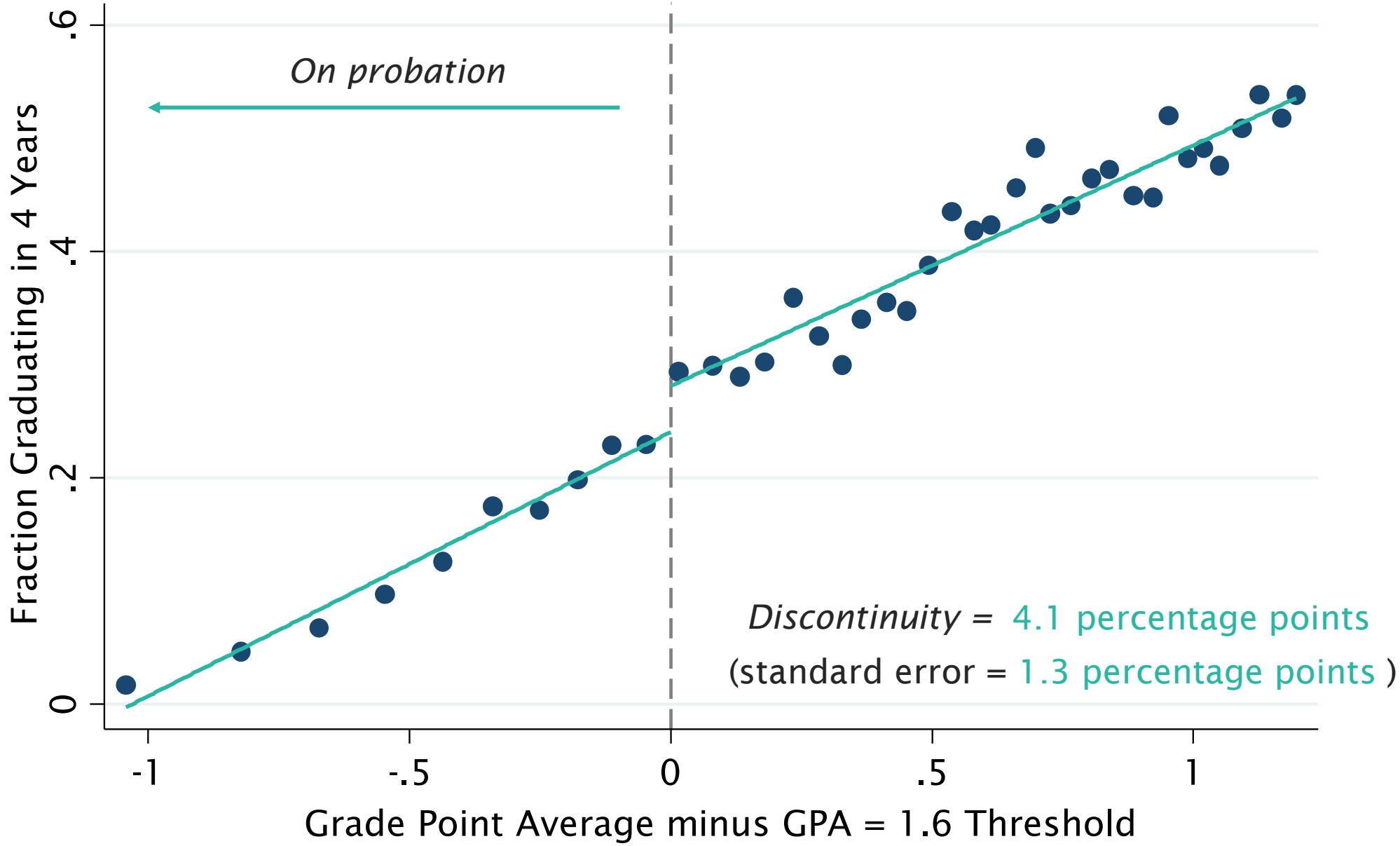
We know that you are capable of academic success, based on your academic record at admission. Let us review your goals and help you develop a plan to achieve them. You have the opportunity and available support to be successful. Please utilize our services to ensure your future success.

Evaluating an Academic Probation Policy at a Large Canadian University

- Download the data set **probation.dta**
- These data contain student-level data from a Canadian University that places students on academic probation if their GPA falls below 1.60
- What is the goal of an academic probation policy like this?
- Is this an effective program?
- Why *don't* we want to compare all students who are on probation with all students who are not on probation to evaluate this program?

Using RDD to Evaluate the Academic Probation Program

Fraction Graduating in Four Years



RD estimating equation with linear control function

- Consider estimating the following regression:

$$Y_i = \beta_0 + \beta_{RD} Z_i + \beta_2 (W_i - w) + \beta_3 Z_i \times (W_i - w) + u_i$$

where $Z_i = 1\{W_i \geq w\}$ is variable equal to 1 if $W_i > w$ and 0 if $W_i < w$

RD estimating equation with linear control function

- Consider estimating the following regression:

$$Y_i = \beta_0 + \beta_{RD} Z_i + \beta_2 (W_i - w) + \beta_3 Z_i \times (W_i - w) + u_i$$

where $Z_i = 1\{W_i \geq w\}$ is variable equal to 1 if $W_i > w$ and 0 if $W_i < w$

- Fitted value for observations with W_i less than w, which have $Z_i = 0$:

$$\hat{Y}_i = \beta_0 + \beta_2 (W_i - w)$$

RD estimating equation with linear control function

- Consider estimating the following regression:

$$Y_i = \beta_0 + \beta_{RD} Z_i + \beta_2 (W_i - w) + \beta_3 Z_i \times (W_i - w) + u_i$$

where $Z_i = 1\{W_i \geq w\}$ is variable equal to 1 if $W_i > w$ and 0 if $W_i < w$

- Fitted value for observations with W_i less than w, which have $Z_i = 0$:

$$\hat{Y}_i = \beta_0 + \beta_2 (W_i - w)$$

- Fitted value for observations with W_i greater than w, which have $Z_i = 1$:

$$\hat{Y}_i = \beta_0 + \beta_{RD} + (\beta_2 + \beta_3)(W_i - w)$$

RD estimating equation with linear control function

- Consider estimating the following regression:

$$Y_i = \beta_0 + \beta_{RD} Z_i + \beta_2 (W_i - w) + \beta_3 Z_i \times (W_i - w) + u_i$$

  where $Z_i = 1\{W_i \geq w\}$ is variable equal to 1 if $W_i > w$ and 0 if $W_i < w$

- Fitted value for observations with W_i less than w, which have $Z_i = 0$:

$$\hat{Y}_i = \beta_0 + \beta_2 (W_i - w)$$

- Fitted value for observations with W_i greater than w, which have $Z_i = 1$:

$$\hat{Y}_i = \beta_0 + \beta_{RD} + (\beta_2 + \beta_3)(W_i - w)$$

- Jump at threshold $W_i = w$ is given by $\Delta \hat{Y}_i = \beta_{RD}$

Trying it on your own

- Use the variables in the **probation.dta** file to validate the research design
- Visual evidence is *always* the best evidence:
  - Binned scatter plots were invented for RDD
  - Histograms (as you drew in Empirical Project Part 1)
- What do you conclude about the validity of the RDD? Is manipulation of the running variable plausible in this setting?
- What do you conclude about the effectiveness of this college's academic probation program? What caveats would you put on your conclusions?