Regression Discontinuity: Extensions

INFO/STSCI/ILRST 3900: Causal Inference

26 Oct 2023

Learning goals for today

At the end of class, you will be able to:

1. Explain how fuzzy discontinuities can be used for causal identification
2. Understand how manipulation or other discontinuities may post potential threats to identification

The big idea

The big idea

- Treatment of interest depends only on whether a **running variable** is above or below a threshold c

The big idea

- Treatment of interest depends only on whether a **running variable** is above or below a threshold c
- Assume $\mathrm{E}(Y^{a=1} \mid R = r)$ and $\mathrm{E}(Y^{a=0} \mid R = r)$ varies smoothly

The big idea

- Treatment of interest depends only on whether a **running variable** is above or below a threshold c
- Assume $\mathrm{E}(Y^{a=1} \mid R = r)$ and $\mathrm{E}(Y^{a=0} \mid R = r)$ varies smoothly
- Above the the cut-off $\mathrm{E}(Y^{a=1} \mid R = r) = \mathrm{E}(Y \mid R = r)$
- Below the the cut-off $\mathrm{E}(Y^{a=0} \mid R = r) = \mathrm{E}(Y \mid R = r)$

The big idea

- Treatment of interest depends only on whether a **running variable** is above or below a threshold c
- Assume $\mathrm{E}(Y^{a=1} \mid R = r)$ and $\mathrm{E}(Y^{a=0} \mid R = r)$ varies smoothly
- Above the the cut-off $\mathrm{E}(Y^{a=1} \mid R = r) = \mathrm{E}(Y \mid R = r)$
- Below the the cut-off $\mathrm{E}(Y^{a=0} \mid R = r) = \mathrm{E}(Y \mid R = r)$
- Using observed data, estimate, $\mathrm{E}(Y \mid R = r)$ for r closer and closer to the cut-off
- Estimate local ATE $\mathrm{E}(Y_i^{a=1} \mid R_i = c) - \mathrm{E}(Y_i^{a=0} \mid X_i = c)$ by

$$\underbrace{lim_{x \to c^+} \mathrm{E}(Y \mid X = x)}_{\text{from above the cut-off}} - \underbrace{\lim_{x \to c^-} \mathrm{E}(Y \mid X = x)}_{\text{from below the cut-off}}$$

What can go wrong?

- **Other discontinuity**: Something other than treatment also jumps at the threshold
- **Fuzzy RDD**: Some units are treated on either side of threshold
- **Manipulation**: Units have control over over their running variable

Other discontinuities

- **RDD Assumption**: Treatment has discontinuity at the cut-off, but everything else continues smoothly
- Jump in outcome is only due to treatment!

Other discontinuities

- **RDD Assumption**: Treatment has discontinuity at the cut-off, but everything else continues smoothly
- Jump in outcome is only due to treatment!
- **Other discontinuity**: If something else "jumps" at the cut-off, then we can't distinguish between effect of treatment and the other thing

Other discontinuities

- **RDD Assumption**: Treatment has discontinuity at the cut-off, but everything else continues smoothly
- Jump in outcome is only due to treatment!
- **Other discontinuity**: If something else "jumps" at the cut-off, then we can't distinguish between effect of treatment and the other thing
- National Merit example:
  - Suppose students above cut-off also receive tutoring on how to write better personal statements
  - Cannot distinguish between effect of Certificate of Merit and tutoring

Other discontinuities

- Requires knowledge about problem context
- Same as drawing the right DAG

Other discontinuities

- Requires knowledge about problem context
- Same as drawing the right DAG
- Can check with placebo tests
  - Pick an outcome that shouldn't be affected by treatment
  - See if RDD "detects" a causal effect of treatment on unrelated outcome

Other discontinuities

- Requires knowledge about problem context
- Same as drawing the right DAG
- Can check with placebo tests
  - Pick an outcome that shouldn't be affected by treatment
  - See if RDD "detects" a causal effect of treatment on unrelated outcome
  - Ex: scores on writing sample)

Fuzzy RDD

Fuzzy RDD

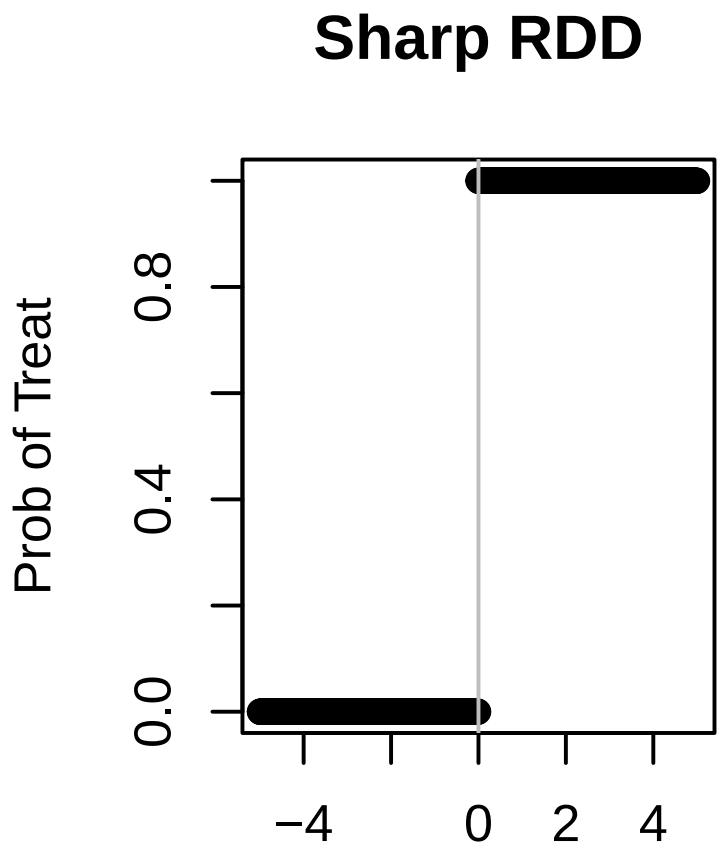
- **Standard setting**: everyone with running variable above threshold is treated, everyone with running threshold is not treated

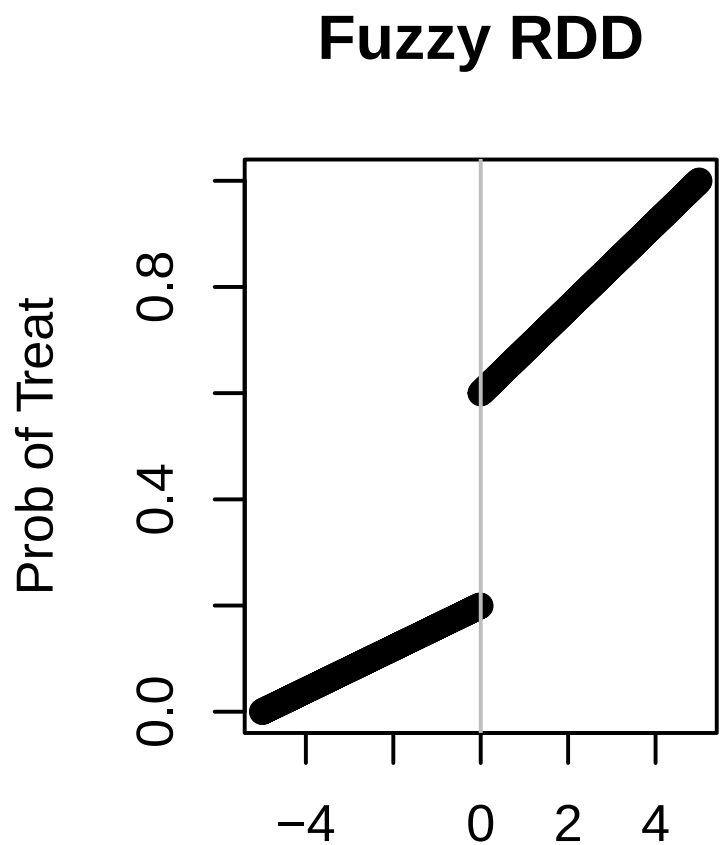
Fuzzy RDD

- **Standard setting**: everyone with running variable above threshold is treated, everyone with running threshold is not treated
- **Fuzzy setting**: Probability of receiving treatment jumps at the threshold

Fuzzy RDD

- **Standard setting**: everyone with running variable above threshold is treated, everyone with running threshold is not treated
- **Fuzzy setting**: Probability of receiving treatment jumps at the threshold





Fuzzy RDD

What is the effect of retirement on spending?[1]

[1] The Retirement Consumption Puzzle: Evidence from a Regression Discontinuity Approach. Battistin et al (2009)

Fuzzy RDD

What is the effect of retirement on spending?[1]

- Spending habits depend on age
- Probability of retirement jumps at 65 because of benefits elegibility
- Some people retire before 65, some people continue to work

[1] The Retirement Consumption Puzzle: Evidence from a Regression Discontinuity Approach. Battistin et al (2009)

Fuzzy RDD

What is the effect of retirement on spending?[1]

- Spending habits depend on age
- Probability of retirement jumps at 65 because of benefits elegibility
- Some people retire before 65, some people continue to work
- Treatment: Retirement
- Outcome: Spending
- Running variable: Age, cutoff: 65

[1] The Retirement Consumption Puzzle: Evidence from a Regression Discontinuity Approach. Battistin et al (2009)

Fuzzy RDD

What is the effect of retirement on spending?[1]

- Spending habits depend on age
- Probability of retirement jumps at 65 because of benefits elegibility
- Some people retire before 65, some people continue to work
- Treatment: Retirement
- Outcome: Spending
- Running variable: Age, cutoff: 65
- Something that is essentially random (being above or below threshold), encourages treatment uptake... sound familiar?

[1] The Retirement Consumption Puzzle: Evidence from a Regression Discontinuity Approach. Battistin et al (2009)

Fuzzy RDD

- Something that is essentially random (being above or below threshold), encourages treatment uptake... sound familiar?

Fuzzy RDD

- Something that is essentially random (being above or below threshold), encourages treatment uptake... sound familiar?
- Around the cut-off, being above/below is like an instrumental variable
- Effect of being above threshold on outcome

$$\lim_{r \to c^+} \mathrm{E}(Y \mid R = r) - \lim_{x \to c^-} \mathrm{E}(Y \mid R = r)$$

- Effect of being above threshold on treatment

$$\lim_{r \to c^+} \mathrm{E}(A \mid R = r) - \lim_{r \to c^-} \mathrm{E}(A \mid R = r)$$

Fuzzy RDD

- Something that is essentially random (being above or below threshold), encourages treatment uptake... sound familiar?
- Around the cut-off, being above/below is like an instrumental variable
- Effect of being above threshold on outcome

$$\lim_{r \to c^+} \mathrm{E}(Y \mid R = r) - \lim_{x \to c^-} \mathrm{E}(Y \mid R = r)$$

- Effect of being above threshold on treatment

$$\lim_{r \to c^+} \mathrm{E}(A \mid R = r) - \lim_{r \to c^-} \mathrm{E}(A \mid R = r)$$

- Dividing gives us the **local ATE for compliers**:

$$\frac{\lim_{r \to c^+} \mathrm{E}(Y \mid R = r) - \lim_{x \to c^-} \mathrm{E}(Y \mid R = r)}{\lim_{r \to c^+} \mathrm{E}(A \mid R = r) - \lim_{r \to c^-} \mathrm{E}(A \mid R = r)}$$

Manipulation

[2]Evaluating the effect of an antidiscrimination law using a regression-discontinuity design. Hahn et al. (1999)

Manipulation

- **Standard setting** For people close the cutoff, being above or below the threshold is essentially random
- Within a small neighborhood of the cut-off, conditional exchangebaility holds

[2] Evaluating the effect of an antidiscrimination law using a regression-discontinuity design. Hahn et al. (1999)

Manipulation

- **Standard setting** For people close the cutoff, being above or below the threshold is essentially random
- Within a small neighborhood of the cut-off, conditional exchangebaility holds
- **Manipulation**: People choose to be above/below threshold

[2] Evaluating the effect of an antidiscrimination law using a regression-discontinuity design. Hahn et al. (1999)

Manipulation

- **Standard setting** For people close the cutoff, being above or below the threshold is essentially random
- Within a small neighborhood of the cut-off, conditional exchangebaility holds
- **Manipulation**: People choose to be above/below threshold
- What is the effect of the 1964 civil rights act on hiring discrimination?[2]

[2] Evaluating the effect of an antidiscrimination law using a regression-discontinuity design. Hahn et al. (1999)

Manipulation

- **Standard setting** For people close the cutoff, being above or below the threshold is essentially random
- Within a small neighborhood of the cut-off, conditional exchangebaility holds
- **Manipulation**: People choose to be above/below threshold
- What is the effect of the 1964 civil rights act on hiring discrimination?[2]
  - Federal EEOC law prohibits discrimination and applies to firms with 15 or more employees
  - Firms with 14 should be essentially the same as firms with 15 employees

[2] Evaluating the effect of an antidiscrimination law using a regression-discontinuity design. Hahn et al. (1999)

Manipulation

- **Standard setting** For people close the cutoff, being above or below the threshold is essentially random
- Within a small neighborhood of the cut-off, conditional exchangebaility holds
- **Manipulation**: People choose to be above/below threshold
- What is the effect of the 1964 civil rights act on hiring discrimination?[2]
  - Federal EEOC law prohibits discrimination and applies to firms with 15 or more employees
  - Firms with 14 should be essentially the same as firms with 15 employees
  - Firms have direct control over how many employees they hire
  - Those wanting to avoid EEOC law may decide to stay under 15 employees

[2] Evaluating the effect of an antidiscrimination law using a regression-discontinuity design. Hahn et al. (1999)

Manipulation

- In many settings, units have control over their running variable to some extent

Manipulation

- In many settings, units have control over their running variable to some extent
- Manipulation becomes a problem when units can choose precisely to be above/below threshold
- Units above the threshold are no longer essentially the same as units below the threshold

Manipulation

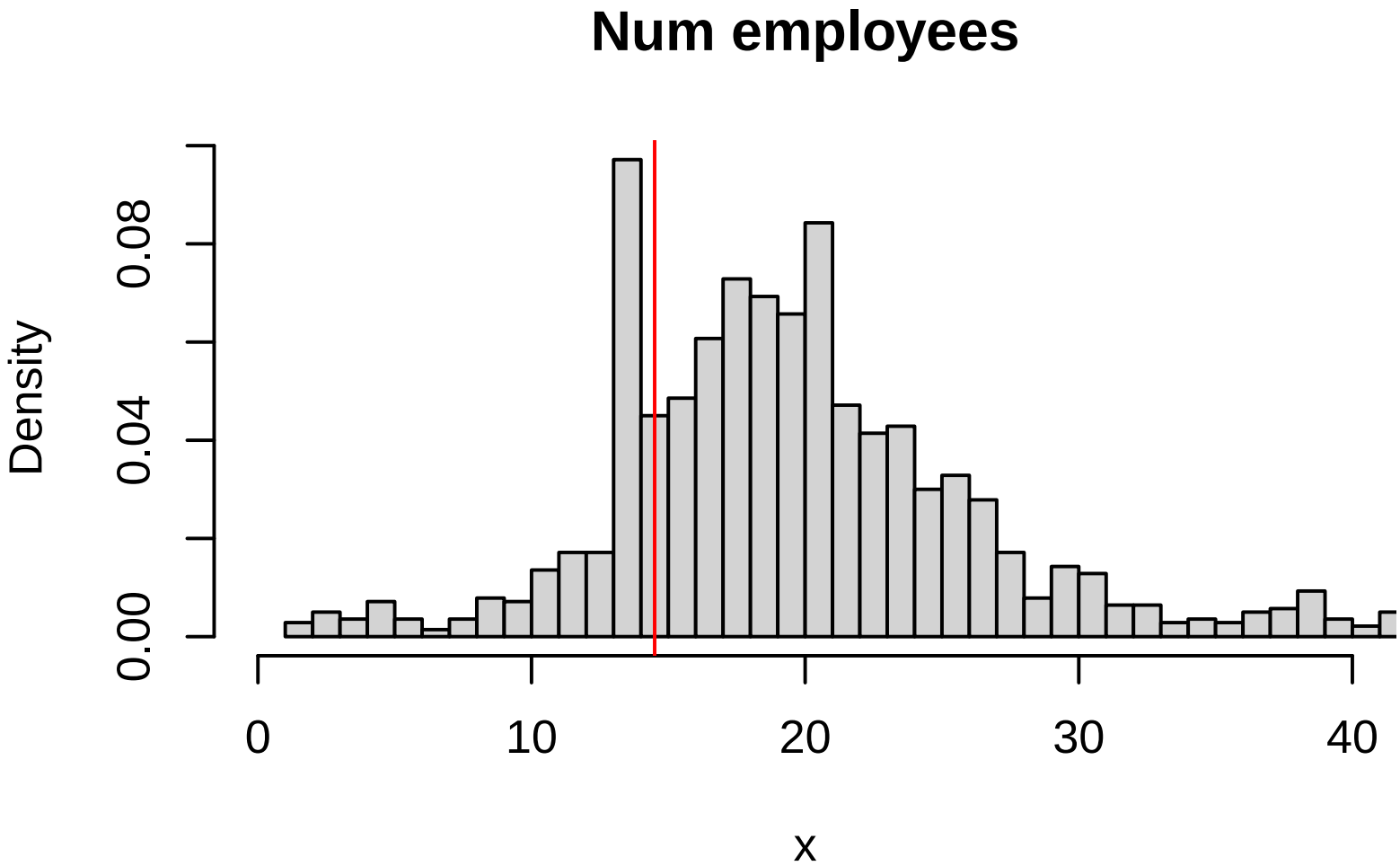
- Hard to test for manipulation directly

Manipulation

- Hard to test for manipulation directly
- Check for balance in covariates

Manipulation

- Hard to test for manipulation directly
- Check for balance in covariates
- If manipulation is occurring, we would expect to see "heaping" on one side of the cut-off



Code

Learning goals for today

At the end of class, you will be able to:

1. Explain how fuzzy discontinuities can be used for causal identification
2. Understand how manipulation or other discontinuities may post potential threats to identification