GOV1368 Section 7

Mohit Karnani

Harvard

Fall 2024

Recap

Last time (before the mock midterm) we covered:

- Endogeneity
- Instrumental Variables (IV)
- IV Identifying Assumptions

Today we will learn about another *flavor* of Instrumental Variables (IV) estimation. It is called the **Regression Discontinuity** (RD) design.

Agenda

Refresher: Instrumental Variables

Regression Discontinuity (RD)

Local Average Treatment Effect (LATE)

Application

Contents

Refresher: Instrumental Variables

Suppose something *exogenously* affects preschool enrollment and does not *directly* impact test scores (e.g. a random voucher program).

Refresher: Instrumental Variables

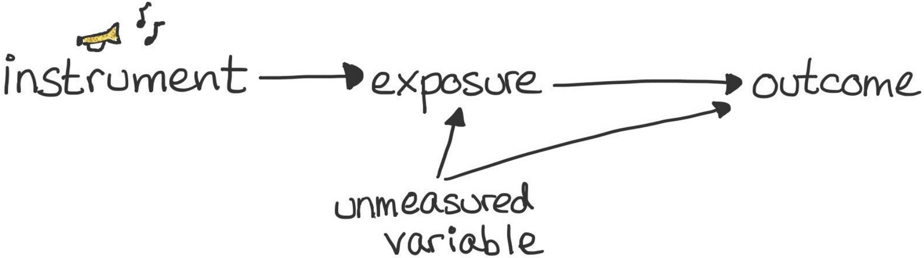
Suppose something *exogenously* affects preschool enrollment and does not *directly* impact test scores (e.g. a random voucher program).

We will call this an **instrumental variable** (or simply an **instrument**).

Refresher: Instrumental Variables

Suppose something *exogenously* affects preschool enrollment and does <u>not</u> *directly* impact test scores (e.g. a random voucher program).

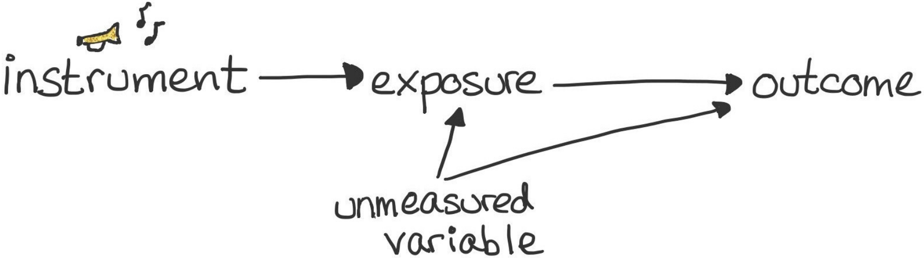
We will call this an **instrumental variable** (or simply an **instrument**).



Refresher: Instrumental Variables

Suppose something *exogenously* affects preschool enrollment and does not *directly* impact test scores (e.g. a random voucher program).

We will call this an **instrumental variable** (or simply an **instrument**).



How can we use this to obtain a causal estimate?

Refresher: The IV-Wald Estimator

Recipe for computing the causal effect of preschooling D_i on test scores Y_i using a randomly assigned voucher Z_i as an instrument:

Refresher: The IV-Wald Estimator

Recipe for computing the causal effect of preschooling D_i on test scores Y_i using a randomly assigned voucher Z_i as an instrument:

1. Calculate the difference between the average test score of those who received the voucher and those who didn't.

$$\text{IV-Wald} \rightarrow \frac{\mathbb{E}[Y_i|Z_i = 1] - \mathbb{E}[Y_i|Z_i = 0]}{\mathbb{E}[D_i|Z_i = 1] - \mathbb{E}[D_i|Z_i = 0]} = \frac{\mathbb{E}[Score_i|Voucher] - \mathbb{E}[Score_i|NoVoucher]}{\mathbb{P}[PreK_i|Voucher] - \mathbb{P}[PreK_i|NoVoucher]}$$

Refresher: The IV-Wald Estimator

Recipe for computing the causal effect of preschooling D_i on test scores Y_i using a randomly assigned voucher Z_i as an instrument:

1. Calculate the difference between the average test score of those who received the voucher and those who didn't.
2. Calculate the difference between the probability of attending preschool for those who received the voucher and those who didn't.

$$\text{IV-Wald} \rightarrow \frac{\mathbb{E}[Y_i|Z_i=1] - \mathbb{E}[Y_i|Z_i=0]}{\mathbb{E}[D_i|Z_i=1] - \mathbb{E}[D_i|Z_i=0]} = \frac{\mathbb{E}[Score_i|Voucher] - \mathbb{E}[Score_i|NoVoucher]}{\mathbb{P}[PreK_i|Voucher] - \mathbb{P}[PreK_i|NoVoucher]}$$

Refresher: The IV-Wald Estimator

Recipe for computing the causal effect of preschooling D_i on test scores Y_i using a randomly assigned voucher Z_i as an instrument:

1. Calculate the difference between the average test score of those who received the voucher and those who didn't.
2. Calculate the difference between the probability of attending preschool for those who received the voucher and those who didn't.
3. Divide the difference in outcomes by the difference in exposure.

$$\text{IV-Wald} \rightarrow \frac{\mathbb{E}[Y_i|Z_i = 1] - \mathbb{E}[Y_i|Z_i = 0]}{\mathbb{E}[D_i|Z_i = 1] - \mathbb{E}[D_i|Z_i = 0]} = \frac{\mathbb{E}[Score_i|Voucher] - \mathbb{E}[Score_i|NoVoucher]}{\mathbb{P}[PreK_i|Voucher] - \mathbb{P}[PreK_i|NoVoucher]}$$

Refresher: The IV-Wald Estimator

Recipe for computing the causal effect of preschooling D_i on test scores Y_i using a randomly assigned voucher Z_i as an instrument:

1. Calculate the difference between the average test score of those who received the voucher and those who didn't.
2. Calculate the difference between the probability of attending preschool for those who received the voucher and those who didn't.
3. Divide the difference in outcomes by the difference in exposure.

$$\text{IV-Wald} \rightarrow \frac{\mathbb{E}[Y_i|Z_i = 1] - \mathbb{E}[Y_i|Z_i = 0]}{\mathbb{E}[D_i|Z_i = 1] - \mathbb{E}[D_i|Z_i = 0]} = \frac{\mathbb{E}[Score_i|Voucher] - \mathbb{E}[Score_i|NoVoucher]}{\mathbb{P}[PreK_i|Voucher] - \mathbb{P}[PreK_i|NoVoucher]}$$

Refresher: The IV-Wald Estimator

Recipe for computing the causal effect of preschooling D_i on test scores Y_i using a randomly assigned voucher Z_i as an instrument:

1. Calculate the difference between the average test score of those who received the voucher and those who didn't.
2. Calculate the difference between the probability of attending preschool for those who received the voucher and those who didn't.
3. Divide the difference in outcomes by the difference in exposure.

$$\text{IV-Wald} \rightarrow \frac{\mathbb{E}[Y_i|Z_i=1]-\mathbb{E}[Y_i|Z_i=0]}{\mathbb{E}[D_i|Z_i=1]-\mathbb{E}[D_i|Z_i=0]} = \frac{\mathbb{E}[Score_i|Voucher]-\mathbb{E}[Score_i|NoVoucher]}{\mathbb{P}[PreK_i|Voucher]-\mathbb{P}[PreK_i|NoVoucher]}$$

Another interpretation: it is the Intent To Treat (ITT) divided by the "first stage".

Refresher: Local Average Treatment Effect (LATE)

Our target estimand is the **Local Average Treatment Effect** (LATE).

Refresher: Local Average Treatment Effect (LATE)

Our target estimand is the **Local Average Treatment Effect** (LATE).

This is the average treatment effect on the *compliers*, i.e. the causal impact of preschooling on test scores for those who would attend preschool *only* if they received the voucher, and not otherwise:

$$LATE := \mathbb{E}[Y_i(1) - Y_i(0) | D_i(Z_i = 1) > D_i(Z_i = 0)]$$

Refresher: Local Average Treatment Effect (LATE)

Our target estimand is the **Local Average Treatment Effect** (LATE).

This is the average treatment effect on the *compliers*, i.e. the causal impact of preschooling on test scores for those who would attend preschool *only* if they received the voucher, and not otherwise:

$$LATE := \mathbb{E}[Y_i(1) - Y_i(0) | D_i(Z_i = 1) > D_i(Z_i = 0)]$$

A taxonomy of the population based on the potential outcomes:

- Compliers: $D_i = Z_i$, i.e. those who attend preschool only if they receive the voucher.
- Always-takers: $D_i = 1$, i.e. those who attend preschool regardless of the voucher.
- Never-takers: $D_i = 0$, i.e. those who don't attend preschool regardless of the voucher.
- Defiers: $D_i = 1 - Z_i$, i.e. those who attend preschool only if they don't receive the voucher. Assumption: there are <u>no defiers</u>.

Refresher: Identification Assumptions

Our identification strategy relies on the following assumptions:

1. **Relevance**: The instrument Z_i is *relevant* for the treatment D_i, i.e.

$$\mathbb{P}[D_i = 1 | Z_i = 1] \neq \mathbb{P}[D_i = 1 | Z_i = 0].$$

Refresher: Identification Assumptions

Our identification strategy relies on the following assumptions:

1. **Relevance**: The instrument Z_i is *relevant* for the treatment D_i, i.e.

$$\mathbb{P}[D_i = 1 | Z_i = 1] \neq \mathbb{P}[D_i = 1 | Z_i = 0].$$

2. **Exclusion**: The instrument Z_i affects the outcome Y_i *only* through the treatment D_i:

$$\mathbb{E}[Y_i(1) - Y_i(0) | D_i = d, Z_i = 1] = \mathbb{E}[Y_i(1) - Y_i(0) | D_i = d, Z_i = 0].$$

Refresher: Identification Assumptions

Our identification strategy relies on the following assumptions:

1. **Relevance**: The instrument Z_i is *relevant* for the treatment D_i, i.e.

$$\mathbb{P}[D_i = 1|Z_i = 1] \neq \mathbb{P}[D_i = 1|Z_i = 0].$$

2. **Exclusion**: The instrument Z_i affects the outcome Y_i *only* through the treatment D_i:

$$\mathbb{E}[Y_i(1) - Y_i(0)|D_i = d, Z_i = 1] = \mathbb{E}[Y_i(1) - Y_i(0)|D_i = d, Z_i = 0].$$

3. **Monotonicity**: There are no defiers, i.e. $D_i \neq 1 - Z_i$.

Refresher: Identification Assumptions

Our identification strategy relies on the following assumptions:

1. **Relevance**: The instrument Z_i is *relevant* for the treatment D_i, i.e.

$$\mathbb{P}[D_i = 1 | Z_i = 1] \neq \mathbb{P}[D_i = 1 | Z_i = 0].$$

2. **Exclusion**: The instrument Z_i affects the outcome Y_i *only* through the treatment D_i:

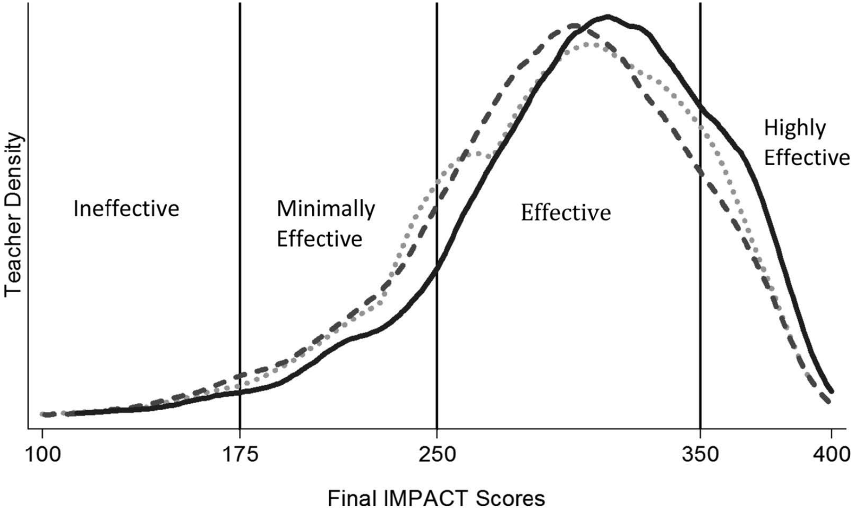
$$\mathbb{E}[Y_i(1) - Y_i(0) | D_i = d, Z_i = 1] = \mathbb{E}[Y_i(1) - Y_i(0) | D_i = d, Z_i = 0].$$

3. **Monotonicity**: There are no defiers, i.e. $D_i \neq 1 - Z_i$.
4. **Independence**: The instrument Z_i is *independent* of the potential outcomes $Y_i(0)$ and $Y_i(1)$.

Contents

Threshold-Crossing as IV

Let's think of a special scenario where there is a cutoff/threshold c such that a variable S_i crossing c determines the take up of a treatment D_i. For example, in the IMPACT study, teachers with scores S_i over $c = 350$ points could receive substantial bonuses D_i.



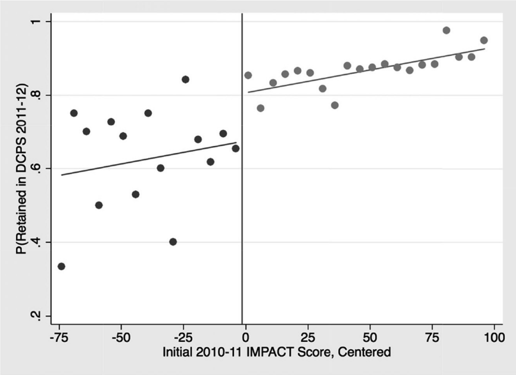
Regression Discontinuity

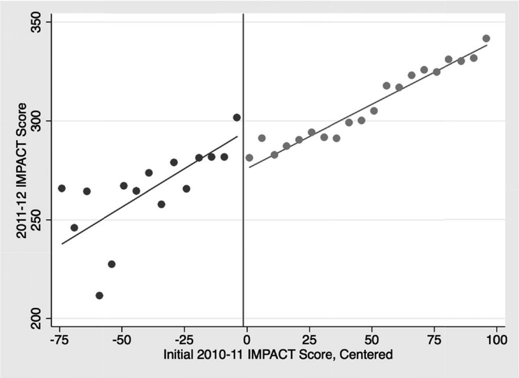
We can measure the causal impact that this treatment D_i (e.g. bonus or dismissal threat) has on some outcome variable Y_i (e.g. future retention or teacher performance).

Regression Discontinuity

We can measure the causal impact that this treatment D_i (e.g. bonus or dismissal threat) has on some outcome variable Y_i (e.g. future retention or teacher performance).

You basically fit "two regressions" and measure the *discontinuity* around the cutoff c:





A simple Wald-like RD formula

Before:

$$\text{IV-Wald} \rightarrow \frac{\mathbb{E}[Y_i|Z_i = 1] - \mathbb{E}[Y_i|Z_i = 0]}{\mathbb{E}[D_i|Z_i = 1] - \mathbb{E}[D_i|Z_i = 0]}$$

Now:

$$\text{“RD-Wald”} \rightarrow \frac{\mathbb{E}[Y_i|S_i \rightarrow c^+] - \mathbb{E}[Y_i|S_i \rightarrow c^-]}{\mathbb{E}[D_i|S_i \rightarrow c^+] - \mathbb{E}[D_i|S_i \rightarrow c^-]}$$

A simple Wald-like RD formula

Before:

$$\text{IV-Wald} \rightarrow \frac{\mathbb{E}[Y_i|Z_i = 1] - \mathbb{E}[Y_i|Z_i = 0]}{\mathbb{E}[D_i|Z_i = 1] - \mathbb{E}[D_i|Z_i = 0]}$$

Now:

$$\text{“RD-Wald”} \rightarrow \frac{\mathbb{E}[Y_i|S_i \rightarrow c^+] - \mathbb{E}[Y_i|S_i \rightarrow c^-]}{\mathbb{E}[D_i|S_i \rightarrow c^+] - \mathbb{E}[D_i|S_i \rightarrow c^-]}$$

We want to compare the average outcome for units *just above* the cutoff vs units *just below* the cutoff. It is *as if* we had a *local randomization* of units into a treatment group or a control group. We can't compare units too far away from the cutoff, because we would confound the effect of the treatment D_i with the effect of having a very different score S_i.

A simple Wald-like RD formula

Before:

$$\text{IV-Wald} \rightarrow \frac{\mathbb{E}[Y_i|Z_i = 1] - \mathbb{E}[Y_i|Z_i = 0]}{\mathbb{E}[D_i|Z_i = 1] - \mathbb{E}[D_i|Z_i = 0]}$$

Now:

$$\text{“RD-Wald”} \rightarrow \frac{\mathbb{E}[Y_i|S_i \rightarrow c^+] - \mathbb{E}[Y_i|S_i \rightarrow c^-]}{\mathbb{E}[D_i|S_i \rightarrow c^+] - \mathbb{E}[D_i|S_i \rightarrow c^-]}$$

We want to compare the average outcome for units *just above* the cutoff vs units *just below* the cutoff. It is *as if* we had a *local randomization* of units into a treatment group or a control group. We can't compare units too far away from the cutoff, because we would confound the effect of the treatment D_i with the effect of having a very different score S_i.

If there is "imperfect compliance", i.e. crossing the threshold does not determine the treatment D_i in a one-to-one fashion, we need to adjust this difference by the change in take-up of the treatment (just as with our basic IV estimator). When this is the case, we talk about a *fuzzy* RD. Otherwise, if crossing c implies $D_i = 1$ for all units, and not crossing it implies $D_i = 0$ for all units, we call this a *sharp* RD (the denominator is 1).

Contents

RD estimates a LATE

From the paper (Dee and Wyckoff, 2015, page 269):

> *We are also careful to emphasize the stylized nature of the causal estimands that result from these RD designs. In particular, it should be noted that the "localness" of these RD estimates implies that they do not necessarily identify the average treatment effect (ATE) associated with the introduction of IMPACT. However, these results do provide credible evidence on the effects of the types of novel performance incentives IMPACT introduced.*

RD estimates a LATE

From the paper (Dee and Wyckoff, 2015, page 269):

> *We are also careful to emphasize the stylized nature of the causal estimands that result from these RD designs. In particular, it should be noted that the "localness" of these RD estimates implies that they do not necessarily identify the average treatment effect (ATE) associated with the introduction of IMPACT. However, these results do provide credible evidence on the effects of the types of novel performance incentives IMPACT introduced.*

Indeed, the estimand estimated by this estimator is a LATE, just like in IV estimation. The difference is that the *compliers* are now defined as units arbitrarily close to c that would change their treatment adoption if and only if they cross c.

Local Average Treatment Effect

Our target estimand is still the Local Average Treatment Effect (LATE):

$$LATE := \mathbb{E}[Y_i(1) - Y_i(0) | D_i(S_i > c) > D_i(S_i < c), S_i \approx c]$$

One key *assumption* for an RD estimate to be valid is that there is **no manipulation** of the running variable S_i around the cutoff c. We also need other covariates to evolve *smoothly* around the cutoff (exclusion restriction).

Contents

Application: School Spending and Educational, Economic Outcomes

`Stata`