Lecture 7 Notes: Instrumental Variables and Regression Discontinuity

Econ 4523: Economics of Education
Prof. Tyler Ransom
University of Oklahoma

Fall 2019

Last Time

- Natural experiments
- Difference in differences
- Panel data / fixed effects

Brief Review

- Review diff in diff, briefly
- Show where OLS parameters reside in the 3×3 table

Today's Big Questions

- How can we learn about causality if we can't run an experiment, we don't have panel data, and we don't want to assume unconfoundedness?
- Instrumental variables, Regression discontinuity
- Under what conditions are these alternative approaches valid?

1 Instrumental variables (IV)

Instrumental variables focus on random variation in life that acts like a randomized experiment. Examples:

- Lotteries (either to determine charter school enrollment or to determine military service)
- Date of birth (determines grade level in school)
- Place of birth (determines social interactions)

In IV studies, there are now three (or four) variables of interest:

1. Outcome variable
2. Treatment variable
3. Pre-treatment variables / control variables
4. Instrument

1.1 Example: date of birth and compulsory schooling

Angrist and Krueger (1991) estimate the effect of education on earnings by leveraging when students were born (relative to state mandated schooling ages), which determines what age a student starts school, and, combined with compulsory schooling laws, what age he or she finishes school. This is borne out in the data: students born between January and June obtain less schooling than those born between July and December.

How is this helpful? If date of birth is effectively random, then we can compare earnings differences of those who were born in the first half of the year with those who were born in the second half of the year, and we know that the only reason for these differences is because of differences in their level of schooling.

How the instrument gives us a plausible counterfactual In other words, the instrument acts like the experimental randomizer:

units that happen to be treated are good counterfactual measures for units in the control group that are exactly the same in all ways except for treatment status.

1.1.1 Instrumental variables conditions

We now review the conditions discussed in the videos by Prof. Matt Masten, applied to the compulsory schooling example:

1. Observe the instrument (quarter of birth), which is correlated with earnings
2. *Assume* that quarter of birth does not cause higher earnings
3. *Assume* that quarter of birth has a causal effect on the amount of education students receive
4. *Assume* that quarter of birth is randomly assigned

5. Step (4) implies that correlation between quarter of birth and education is causal (by random assignment)

6. Step (4) also implies that quarter of birth is not correlated with any other possible confounder.

Thus, the only reason quarter of birth is correlated with earnings is because it affects educational attainment.

We can boil the previous 6 conditions down to 2 assumptions, which are discussed in Lovenheim and Turner:

1. The instrument must be correlated with treatment (this is often called the "first stage")

2. The instrument is randomly assigned

From these assumptions, we get all of the 6 conditions from the YouTube video. The advantage of the video is that it traces through the logic much more clearly.

1.1.2 Testing the IV assumptions

The main difficulty with instrumental variables is that we can only test one of the assumptions!

- That is, we can easily see in our data if the instrument and treatment are correlated.
- It is impossible to tell, however, if the instrument is randomly assigned!
  - This is for the usual reason: even if the instrument appeared to be random among all of the variables in our data set, there are many other variables that we don't have in our data, and the instrument might be correlated with at least one of those unobserved variables.

For this reason, a causal claim from an instrumental variables design typically requires strong logical reasoning and, in many cases, qualitative evidence that the instrument is operating in the way the researcher is claiming it is.

2 Regression discontinuity (RD)

A method related to IV is regression discontinuity, in which a researcher examines how outcomes differ on either side of a sharp break. Examples include:

- Financial aid policies where eligibility is satisfied by being below a certain income threshold (Denning, Marx, and Turner, 2017)
- University admission policies wherein the university admits all students above some test score threshold
- Public school boundaries that are strictly enforced

You'll know you're dealing with RD if you hear the words "threshold" or "cutoff."

2.1 The three variables in RD designs (RDDs)

RDs consist of three variables:

1. **Running variable.** (a.k.a. the forcing variable) This is the variable that the cutoff is based on (e.g. standardized test score, income level, geographical location)

2. **Treatment variable.** Treatment is defined as being above the cutoff.

3. **Outcome variable.** This is, as usual, the variable of interest (e.g. future earnings, graduation from college, admission to college, etc.)

2.2 How RDs work

One can compute a treatment effect from an RD design as follows:

1. Compute the outcome for units just above the cutoff

2. Compute the outcome for units just below the cutoff

3. The difference in the outcome is the average treatment effect.

What constitutes "just above" or "just below"? This is a topic we won't get into, as it is highly technical. Just know that the interpretation of the ATE in an RD is only valid near the cutoff value. Hence, the causal effect computed from an IV or RD study is called a **Local Average Treatment Effect (LATE)** because it only applies near the cutoff, or only to units whose treatment status is affected by the instrument.

2.3 RDs' relationship to IV

In fact, you can think of the cutoff in an RD as being an instrument: it is related to the treatment, but completely unrelated to the outcome (except through the treatment).

2.3.1 Sharp vs. fuzzy RD

There are actually two kinds of RDs: **sharp** and **fuzzy**.

- Sharp RDs are characterized by 100% compliance. This means that all units above the threshold get treated.

- Fuzzy RDs are just the opposite: they have less than 100% compliance.

Why might there be less than 100% compliance? Well, maybe students who score above a certain level would be admitted to University X, but they don't bother applying (either because they didn't know they would be admitted, or because they didn't want to attend University X). Or maybe the test score is one of many conditions required for admission, and some students who score higher than the cutoff don't satisfy other conditions.

2.3.2 IVs in fuzzy RDs

Because treatment is not 100% determined by being above the cutoff in fuzzy RD designs, we need to use instrumental variables to compute a causal effect. In this case, being above the cutoff acts as an instrument for treatment, and we use the instrumental variables techniques to compute the ATE (essentially, this is ATE = Reduced form / 1st stage).

2.4 When RDs might not be valid

Validity of the RD requires randomness in the cutoff. If you see bunching in the data just above or just below the cutoff, this is evidence of units attempting to manipulate their treatment status. RDs are usually valid if:

1. The cutoff is difficult to manipulate (i.e. test scores or birthdays)
2. The cutoff is unknown to individuals

On the other hand, if individuals can easily manipulate their treatment status (e.g. if there is an income threshold that is publicly known), then the RD is likely to be invalid for the usual reasons: selection bias. (In this case, units "selected" which side of the cutoff they would be in.)

2.4.1 A note on external validity

Remember that treatment effect recovered from an RDD is not valid outside of a narrow band around the cutoff. This means that many RDs are not externally valid, because they only capture treatment effects for the very narrow subpopulation that happens to be near the cutoff.

Summary

	Instrumental Variables	Regression Discontinuity
"Fourth" variable	Instrument	Running variable
Relevance condition	Instr. corr. w/treatment	Treatment changes across cutoff
Randomness condition	Instr. randomly assigned	Cutoff not manipulable
How to test relevance?	Corr (Instr, Trtmnt) $\neq 0$	$\text{Trtmnt}_a - \text{Trtmnt}_b \neq 0$
How to test randomness?	Appeal to qualitative argument	No discontinuity in dist'n of running variable at cutoff
(Local) Treatment effect?	$\frac{\text{corr(outcome,instrument)}}{\text{corr(treatment,instrument)}}$	$\frac{outcome_a - outcome_b}{treatment_a - treatment_b}$
External validity?	LATE only applies to units affected by instrument	LATE only applies near the cutoff

Video links

- https://www.youtube.com/watch?v=4xF_DMbL14w
- https://www.youtube.com/watch?v=nsr9eh-qVPg
- https://www.youtube.com/watch?v=TfKwgGT2fSM
- https://www.youtube.com/watch?v=QvrFpByeJxc
- https://www.youtube.com/watch?v=cP5BVFvWo64 (first 5 minutes)

References

Angrist, Joshua D. and Alan B. Krueger. 1991. "Does Compulsory School Attendance Affect Schooling and Earnings?" *The Quarterly Journal of Economics* 106 (4):979–1014.

Denning, Jeffrey T., Benjamin M. Marx, and Lesley J. Turner. 2017. "ProPelled: The Effects of Grants on Graduation, Earnings, and Welfare." Working paper, Brigham Young University.

Lovenheim, Michael and Sarah E. Turner. 2017. *Economics of Education*. New York: Worth Publishers.