Week 9: RD

JJ Chen

March 13, 2015

Today's Plan

1. A midterm exam question and related issues
2. Recap: Shape and fuzz RD
3. Some RD examples

Exam Question

- We are interested in the impact of school spending on student achievement
  - Reg: $\text{Scores}_{ics} = \alpha + \rho \text{Spending}_{cs} + \text{X}'_{ics}\boldsymbol{\gamma} + e_{ics}$
  - i: student; c: school; s: states
- We worry about the sorting of parents and children into schools, so we aggregate to state level
  - Reg: $\overline{\text{Scores}}_s = \alpha + \rho \overline{\text{Spending}}_s + \bar{\text{X}}'_s\boldsymbol{\gamma} + \bar{e}_s$
- **Q: How to use "an IV framework" to think about it? What test would you provide?**

Grouped-data Version of IV

- Aggregation as a grouped-data version of IV
  - Each $\overline{\text{Spending}}_s$ is a conditional mean defined by a state dummy, say IL_i
  - Just like in Visual-IV, each dots is defined by a draft number dummy
  - Fitting an OLS model is like running a 2SLS procedure
- The instruments are a full set of mutually exclusive dummy variables: States
  - So the grouped-data estimator (GLS) is a linear combination of all state-specific Wald estimators

A Few More Grouped-Data Examples

- Angrist (1991): To deal with measurement error in measures of hours worked
  - In fact, this is related to the original Wald estimator, where grouped-data is defined by a categorical variable that is correlated with the true regressor but independent of measurement error
- Rivkin, Hanushek, and Kain (2005): Aggregate to grade level to overcome sorting at classroom level
- Many peer effects models, see Angrist (2014)
  - For examples, $s_{ij} = \alpha + \rho \bar{s}_j + e_{ij}$
  - What're the instruments? What's the first stage?
  - This is a bad regression, don't run it

Over-id Tests

- Once we become more familiar with the link between grouped-data and IVs, the logic of over-id test should be natural
  - Just back to Visual-IV
- Given the precision of group means, how well does the line fits the data points (which are about the same as many just-identified estimates)
- So, the question asks for a Chi-sq test. But many of you jump to the endogeneity test.

Some Questions

- The RD idea is so simple that I think reviewing basic things like ID assumption, tests for RD assumptions, and RD regressions is kind of dull. So let's do questions:
  1. What's the ID assumption?
  2. What're the implications of the ID assumption? (Or how would you test the ID assumption?)
  3. The relationship between the outcome variable and the running variable can be nonlinear. So would you consider using a high-order polynomials regression? (Gelman and Imbens (2014))

More Questions

4. The treatment dummy is determined by the running variable, so should we worry about perfectly collinearity?
5. What's the trade-off of using non-parametric RD (narrowing the window around the cutoff)?
6. By definition, there should be no OVB for the treatment D_i. That is, once we conditional on the running variable X_i, there is no variation left for the treatment. How can there be any OVB that's correlated with $\tilde{\mathrm{D}}_i$, where $\tilde{\mathrm{D}}_i$ is the residual of running a regression of D_i on X_i? What's the catch? (There're at least two...)

More Questions

7. How bad is manipulation/heaping around the cutoff?
8. How does RD differ from matching?
9. In one sentence, how would you distinguish shape RD and fuzzy RD?

Some RD People Used

- Let's think about which one of these cutoffs is easily manipulated
  1. The second Sunday in March
  2. Age: 21, 65
  3. 1500 grams
  4. School district border
  5. GPA: 2.0, 3.5, 3.666, 3.75
  6. IQ: 137, 113

References I

Angrist, Joshua D. 1991. "Grouped-Data Estimation and Testing in Simple Labor-Supply Models." *Journal of Econometrics* 47 (2). Elsevier: 243–66.

———. 2014. "The Perils of Peer Effects." *Labour Economics* 30. Elsevier: 98–108.

Gelman, Andrew, and Guido Imbens. 2014. *Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs*. National Bureau of Economic Research.

Rivkin, Steven G, Eric A Hanushek, and John F Kain. 2005. "Teachers, Schools, and Academic Achievement." *Econometrica* 73 (2). Wiley Online Library: 417–58.