

Instrumental Variables

1. (short answer) If ability is omitted from a return of education on wages regression, which way do you expect the OLS estimate to be biased. Assuming the 2SLS estimate from Card (1993) identifies a causal effect, why is the 2SLS estimate larger than the OLS estimate?
2. (short answer) Intuitively, list and explain the five assumptions required to estimate a LATE with heterogeneous treatment effects. Which of the five assumptions are testable?
3. (short answer) List the three instrumental variable causal equations of interest by name and write them out using variables, coefficients and error terms appropriately.
4. (short answer) Which of the three instrumental variable causal equations of interest has a causal interpretation if monotonicity is violated?
5. (numerical) Given the provided estimates, show that the structural equation is the ratio of the reduced form and the first stage
6. (short answer) What is the problem with a weak instrument?
7. (short answer) Regarding compliance, what are the four sub populations? What is the empirical problem with the group that does not comply?
8. (short answer) If you are running a randomized experiment with one-sided non-compliance, what does your OLS estimate return, and what is the name of this instrumental variable causal equation? How can you get a version of a local average treatment effect.

Regression Discontinuity Design

1. (short answer) What is the specific causal effect estimated with sharp RDD and fuzzy RDD?
2. (short answer) List the three RDD causal identification violations discussed and provide an example for each.
3. (short answer) How are sharp RDD and fuzzy RDD similar? How are they different?
4. (short answer) What are the identifying assumptions for sharp RDD and fuzzy RDD?
5. (short answer) How is sharp RDD analogous to 2SLS?
6. (short answer) When fitting the trends on the running variable for an RDD, do you want to fit a flexible or rigid trend? What are the concerns of using the kind of trend you do not want?

Panel Data

1. (short answer) If you are using two-way fixed effects (e.g. year and state), do you need to cluster your standard errors? If so, by which variable(s)? Why?
2. (short answer) What are the two panel caveats discussed in lecture (and in the textbook)? Describe each caveat with an example.
3. (short answer) Describe how computers efficiently estimate fixed effect regressions? What theorem (discussed earlier in the semester) validates this technique?
4. (short answer) Intuitively, what kind of bias do fixed effects remove from regression estimates?

Difference-in-Differences

1. (short answer) Why can it be problematic to estimate an SDO between treatment and control groups in the post-treatment period? What about an SDO for pre- and post-treatment periods for the treatment group? What types of biases are present?
2. (short answer) What is the main identifying assumption for difference-in-differences? Is it directly testable? Why or why not?
3. (short answer) What are the two purposes of producing an event study figure (*hint: one is from the pre-treatment period and the other is from the post*)?
4. (short answer) What is the purpose of estimating a lagged placebo test? Does it provide direct or indirect evidence of what it is used to support? Why or why not?
5. (short answer) In differential timing difference-in-differences, the estimated causal coefficient is modified from the ATT by two components. What are the two components? For each component, what increases their value?
6. (numerical) Using the given 2×2 table of averages, calculate the SDO between treatment and control groups in the post period, calculate the SDO between pre-treatment and post-treatment periods for the treated group, and calculate the ATT.