Problem Set 3
Econometrics 718
Chris Taber
Due: Tues. Nov. 8

Problem 1. Consider a model in which

$$Y_i = \beta_0 + \beta_1 T_i + \mu_i + \xi_i$$

Assume that

- ξ_i is mean 0 and independent of everything else in the model
- μ_i is 0 with probability 0.5 and 1 with probability 0.5.
- There is a further variable X_i which is

$$X_i = \frac{1}{2}\mu_i + \nu_i$$

  where ν_i is independent of everything else in the model and uniformly distributed on $[0, 1]$
- T_i is binary and determined according to the following rule
  - When $X_i < x^*$and $\mu_i = 0$ then $Pr(T_i = 1) = 0$
  - When $X_i \geq x^*$ and $\mu_i = 0$ then $Pr(T_i = 1) =p_0$
  - When $X_i < x^*$and $\mu_i = 1$ then $Pr(T_i = 1) = 0$
  - When $X_i \geq x^*$ and $\mu_i = 1$ then $Pr(T_i = 1) =p_1$
- The econometrician observes X_i, T_i, and Y_i but not the other variables.

a) Suppose you run a regression of Y_i onto T_i. What is the plim of $\widehat{\beta}_1$ (where $\widehat{\beta}_1$ is the OLS estimator of the coefficient on T_i)?

b) Suggest a methodology to get a consistent estimate of β_1 and sketch how identification works.

c) Now suppose the type 1 people (i.e. the people for whom $\mu_i = 1$) can manipulate their X_i. In particular assume they are endowed with X_{0i}. They can change their X_i to $X_{0i} + \delta$ at some cost. Assume that anyone with $x^* - \delta < X_{0i} < x^*$ chooses to pay this cost so that

$$X_i = \begin{cases} X_{0i} + \delta & \text{if } x^* - \delta < X_{0i} < x^* \\ X_{0i} & \text{otherwise} \end{cases}$$

How does this affect your estimates from you estimation method above?

d) Now suppose they can not manipulate it exactly but only with error so

$$X_i = \begin{cases} X_{0i} + \delta + \omega_i & \text{if } x^* - \delta < X_{0i} < x^* \\ X_{0i} & \text{otherwise} \end{cases}$$

where ω_i is independent of everything else and uniform. Sketch why this form of manipulation is not a problem.

Problem 2. Take a real data set with a real X and a real Y that are related somehow. Construct a placebo treatment T_i by choosing some rule for so that $T_i = 1$ when $X_i > x^*$ for some x^*. Now try estimating the model

$$Y_i = \beta_0 + \alpha T_i + u_i$$

by regression discontinuity in several different ways (i.e. kernel regression, local linear regression, using polynomials etc.) Compare the results (given that you should get an effect of zero)

Problem 3. Construct a dummy data set as I did in class in the rd.do file (which is on my website). Modify the simulated model and the estimation method for a fuzzy design rather than a sharp design and show what you get.

Problem 4. Take any data set you would like and consider using some instrument for some treatment. Control for a bunch of regressors.

a) First run IV and get estimates.

b) Now calculate the bias assuming that "selection on the observables is the same as selection on unobservables." In doing so you can assume that the unobservables are uncorrelated with the observables and that your answer in a) gave consistent estimates of γ (the regression coefficients on the observed variables). Given the formula for the bias of IV given in the lecture notes in discussing the bias using "Catholic as an instrument" calculate the bias.

c) Now construct for yourself a perfect instrument. That is construct a model with

$$\begin{aligned} T_i &= \beta_0 + \beta_1 Z_i + X_i' \beta_2 + u_i \\ Y_i &= \gamma_0 + \gamma_1 T_i + X_i' \gamma_2 + \varepsilon_i \end{aligned}$$

where Z_i is independent of X_i and ε_i, but u_i is correlated with ε_i. That will involve using a random number generator to generate new values of $u_i, \varepsilon_i, Z_i, T_i$, and Y_i, but using the $X's$ you used in a). Estimate the IV model and also estimate the bias in this case.