Problem Set 4
Econometrics 742
Chris Taber
Due: Wed. Feb. 29

**Problem 1.** Consider a model in which

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

Assume that

- $\xi_i$ is mean 0 and independent of everything else in the model
- $\mu_i$ 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 $\nu_i$ 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 econometricial 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 $\beta_1$ 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 they can change their $X_i$ to $X_i + \delta$ at some cost, and anyone with $X_i < x^* < X_i + \delta$ pay this cost and manipulate there $X_i$. How does this affect your estimates from you estimation method above?

**d)** Now suppose they can not manipulate it exactly but in reality

$$X_i = Z_i + \omega_i$$

and they can manipulate $Z_i$ but not $\omega_i$ where $\omega_i$ is independent of everything else and uniform. Sketch why this form of manipulation is not a problem. For simplicity (if you want) assume that $Z_i$ only takes on two values, but that the distribution of $Z_i$ depends on $\mu_i$.

**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.