Problem Set 10
Instructor : Professor Yujung Hwang

1. Use pset4Q1data.dta to solve this question. The data includes observations from one treated group and one control group. There are two time periods, pre- and post-treatment.

   (a) Explain an identifying assumption to use difference-in-difference research design.
   (b) Draw an average outcome graph by group and year. Based on a graph, explain whether you expect a positive / close to zero / negative difference-in-difference estimate.
   (c) Compute a difference-in-difference estimate and report its standard error clustered at the treated group level.

2. Use pset4Q2data.dta to solve this question. The data includes observations from multiple treated groups, that are treated at different timings.

   (a) Draw an event study plot and discuss whether the identifying assumption for two-way fixed effect research design is likely to hold. Normalize using a year before treatment year.

   (b) Compute a (standard) two-way fixed effect estimate and its standard error, clustered at the group level.

   (c) Use de Chaisemartin and D'Haultfeuille (2020) "did_multiplegt" command and compute dynamic treatment effects pre- and post-treatment.

3. Use Pset5Q1data.dta to solve this question. The data includes a treatment variable D, a running variable X, another covariate Z, an outcome variable Y. The discontinuity of treatment D is known to happen at X=1 cutoff.

   (a) Bin X variables using the width 0.05. Draw a scatterplot of the averages of D in each X bin. State whether you can use either sharp RD or fuzzy RD design.

   (b) Using the same binned X in (a), draw a scatterplot of the averages of Y in each X bin. State whether you see discontinuous jump in Y around the cutoff X=1.

   (c) Set the bandwidth=0.2. Estimate the RD estimate using 4^{th} order polynomial.

Report the estimate with the SE. Show the RD plot.

(d) Change the bandwidth to 0.5. Estimate the RD estimate using 4^{th} order polynomial. Report the estimate with the SE. Discuss how the estimate changes from (b). Show the RD plot.

(e) Estimate the RD estimate using the `rdrobust' command which computes the optimal bandwidth. Use the epanechnikov kernel. Show the `rdplot' graph.

(f) Examine whether a variable Z is balanced around the cutoff. State whether the evidence supports continuity assumption of RD.

(g) Examine whether there is any evidence of self-selection around the cutoff using the density test (rddensity). Report the p-value.

(h) Using a different cutoff point 1.3, do the placebo test. That is, redo the RD estimation using a different cutoff 1.3 State whether you find any significant treatment effect.