Week 3, HW 4: Causal Inference with Difference in Differences (DiD)

Matt Goldman
mattgold@microsoft.com

June 28, 2018

1. **Picking Control Variables:** Imagine you have one data point from every county in the US that gives the unemployment rate in that county in 2010 and the amount of money spent by that country on job training programs in 2008. You want to know the treatment effect of money spent on job training programs onto the outcome of unemployment (as measured two years later). List as many examples as you can think of for each category of control variable:

   (a) Confounding variables:

   (b) Variance-reducing variables:

   (c) Unrelated variables:

   (d) Bad Controls:

2. **DiD with Card and Krueger (1993)**. Download the dataset and definitions file from the coursework repo. Note: (1) that the first interview occurred before the onset of the minimum wage and that the second occurred after and (2) this dataset is paired (each restaurant has a before and after observation) so the steps you take will be a little bit different than those discussed in class.

   (a) Before looking at the data, look just at the list outcome metrics that were tracked in each interview (EMPFT EMPPT NMGRS WAGEST INCTIME FIRSTINC BONUS PCTAFF MEALS OPEN HRSOPEN PSODA PFRY PENTREE NREGS). List the ones that you think could be impacted by the minimum wage change and give a brief explanation of why.

(b) Compute the 'diffs'. Check to see if each outcome metric changed in a statistically significant way between the two interviews. Do this analysis separately for PA and NJ.

(c) Now compute the "diff-in-diff" via the regression method. As I mentioned previously, this step will be a little different because the data is paired. Think carefully about the best way to do this and I'll come and chat with you guys about it after about an hour.

(d) Do you believe this analysis provides definitive evidence on the impact of the minimum wage on employment? Can you imagine a scenario that would lead to this type of estimation giving biased results?

(e) Card and Krueger create an Independent Variable called $GAP = (5.05 - W_{pre})/W_{pre}$, this is the percentage wage increase that New Jersey restaurants needed in order to meet the minimum wage. Use the variable $WAGEST$ (from before the interview) to create this variable. How might restaurants with very large or small values of this variable differ in their response to the minimum wage. Why do you think this variable is interesting? Run any other interesting analysis you can think of using this variable.