**Day 2:**

Briefly review

Event studies, triple differences, falsifications and covariates as well as “selection and parallel trends”

* Simulations in R and Stata
* Replications in R and Stata to learn new specifications for estimating causal effects with diff-in-diff with covariates

Review:

Fundamentals of diff-in-diff.

1. The equation. Diff in diff equation. “Four averages and three subtractions” or what is more generally called technically “simple 2x2”. I like to call it “four averages and three subtractions” bc: a) it is that and b) I like to quote Orley Ashenfelter.   
     
   2x2 or DiD eq: { E[Y|D=1,Post] – E[Y|D=1, Pre] } – { E[Y|D=0, Post] – E[Y|D=0, Pre] }  
     
   You can estimate that directly or run this regression:  
     
   Y = a + gamma Post + lambda Treat + delta (Post x Treat) + e  
     
   OLS estimation of those coefficients, \widehat{\delta} is four averages and three subtractions
2. Parameter and Identification. If you assume no anticipation and you use a never treated comparison group, then the DiD equation “identifies” the sum of two terms:  
     
   DiD = ATT + “Non-parallel trends bias”   
     
   That second term is this:  
     
   “Non-parallel trends bias expression”  
   = { E[Y(0)|D=1,Post] – E[Y(0)|D=1, Pre] } – { E[Y(0)|D=0, Post] – E[Y(0)|D=0, Pre] }  
     
   Notice interestingly, parallel trends as a concept is a diff-in-diff equation, just not on Y. It's on Y(0).
3. Repeated cross sections (Hong paper to deal with compositional changes), and we discussed the chained diff-in-diff
4. Discuss weights and the role of sorting and heterogenous treatment effects as they interact with weights and how that changes causal parameters and can accidentally lead to misinterpretation of estimates