AAE: Exercise Diff-in-diff, timing, Bacon-dec

by: Felix Weinhardt

In the lecture we learnt that the TWFE extimator averages different two-by-two comparisions into an overall estimate based on group size (squared) and remaining treatment variation. This is very similar to the OLS weighting results that we discussed. One particular in the context of TWFE is that the partialling out of variation (i.e. the fixed effects) can result in negative conditional variances, and hence negative group weights, if the model is mis-specified. The good news is that we can examine empirically the weight and estimates associated to the various two-by-two comparisions. Different estimators have been put forward by the literature that effectievly propose different ways in avoiding negative weights and for weighting two-by-two comparison estimates into an overall estiamte.

1. Read the QJE paper by Stevenson and Wolfers that examines the effects of divorce law on female suicide. Answer the following questions
   (a) What is their research question?
   (b) What type of diff-in-diff setting do we have here (think: how many groups/times, binary or continuous treatment)
   (c) What do they do to defend parallel trend assumptions?
2. So far so good. But in the end they just have a TWFE estimator, so lets examine the subgroups. For this, install the Bacon-decomposition package into stata by typing $ssc install bacondecomp$.
   (a) Load the Stevenson and Wolfers data provided on github.
   (b) Run the TWFE model by typing §xtreg asmrs post pcinc asmrh cases i.year, fe robust"
      i. what is the main estimate and its statistical significance?
      ii. How does this specification differ from the specification in the paper on page 276?
      iii. Estimate the TWFE model as above but for the average effect of the reform five years after implementation.
   (c) Check the Stata helpfile to apply the Bacon-decomposition to the TWFE model of the effects of the divorce law on female suicides five years after state-level implementation. How do you interpret the resulting scatterplot?
   (d) Request a detailed decomposition to better understand the how the different groups average into the overall estimate?