Problem Set 2b: Panel Data and Safety Belt Legislation

This empirical portion of this problem set is based off of a data set used in Anderson (2008). That paper examined the effects of light trucks on traffic fatalities, but in doing so it also collected data on primary and secondary seat belt laws as additional covariates. This problem set will focus on the question of whether primary seat belt laws save lives. (A primary belt law stipulates that law enforcement can ticket a driver for not wearing a seat belt, regardless of whether he has broken any other laws. A secondary belt law stipulates that law enforcement can only ticket a driver for not wearing a seat belt if the driver has already been pulled over for simultaneously breaking a different traffic law.)

Unlike the data from the previous problem set, the data here have already been cleaned, with missing values dropped or replaced. The labels on the variables should be self-explanatory.

1. We now apply the synthetic control methods from Abadie et al (2010).

   Some preliminaries: Abadie et al have created a downloadable "canned" command to run the synthetic control method. To download the command for Stata you will need to have an updated version of Stata and be running Stata on a Mac, PC or Unix/Linux. R code is also available. In Stata type 'update all' and then 'update swap'. Next, go to the website below and follow the instructions. There is also downloadable code for Matlab at http://web.stanford.edu/∼jhain/synthpage.html (note:

blindly copying and pasting this URL may not reproduce the tilde)

(a) We created an aggregate "treatment" state (state number 99 or "TU") which combines the (population weighted) data from the first 4 states to have a primary seatbelt law (CT, IA, NM, TX). Please use this state as the "treatment" state in the synthetic control analysis.

   i. Compare the average pre-period log traffic fatalities per capita of the TU site to that of the average of all the "control" states. Next, graph the pre-period log traffic fatalities by year for the pre-period for both the TU and the average of the control group. Interpret.

   ii. Compare the dependent variable between the TU site and each control state for the year before the treatment. Which control state best matches the TU? Now compare this state's covariates with the TU covariates. Do they appear similar? What might this imply for in terms of using this state as the counterfactual state?

(b) Apply the synthetic control method using the available covariates and pre-treatment outcomes to construct a synthetic control group.

   i. Discuss the synthetic control method including its benefits

and potential drawbacks.

ii. Use the pre-programmed command provided by Abadie et al to apply the synthetic control method. (You are free to use either Stata, Matlab, or R but answers will be provided in Stata and R only). Please be sure to state precisely what the command is doing and how you determined your preferred specification.

(c) Graphical interpretation and treatment significance.

i. Generate graphs plotting the gap between the TU and the synthetic control group under both your preferred specification and a few other specifications you tried.

ii. Compare the graph plotting the gap between the TU and the synthetic control group under your preferred specification with the graphs plotting the gap between each control state and its "placebo" treatment. Do you conclude that the treatment was significant? Why or why not?

iii. Create a graph of the post-treatment/pre-treatment prediction ratios of the Mean Squared Prediction Errors (MSPE) for the actual and "plecebo" treatment gaps in (ii). [See Abadie et al for an example]. Do you conclude that the treatment was significant? Why or why not?

(d) How do your synthetic control results compare to your fixed effects

results from question (3) in the last problem set? Interpret any differences.