Event Study Graphs using Stata

Soumak Basumallik

A difference-in-differences (DiD) design is a quasi-experimental technique that compares the changes in the outcome of interest over time between a population which is affected by a policy/event/treatment (the treatment group) to a similar population which is unaffected by the policy/event/treatment (the control group). It is a crucial tool for policy analyses and is one of the primary techniques used by researchers for getting causal estimates. However, the DiD design hinges on a crucial assumption—the *parallel trends assumption*.

The parallel trends assumption states that the treatment and control group prior to the policy/event/treatment would have followed the same trend had there been no policy/event/treatment that is, in the absence of a policy/event/treatment the difference between the treatment and control group would have remained unchanged over time. While there a few ways to check if the parallel trends assumption is satisfied like, running a DiD for the pre-period and see if the difference between the treatment and control group were insignificant; the better or more appealing way to check it is through event study graphs.

Event study graphs are used to depict whether there existed differences between the treatment and control groups at different time points before and after the policy or event. Event study designs are even more helpful when policy/event/treatment do not happen at a single time-point. To construct an event study graph, one has to create lags and leads on the policy variable and then regress the outcome on these lags and leads along with the other relevant covariates and then plot the coefficients of the lags and leads on the outcome along with the associated confidence intervals.

In this article, I have tried to replicate the first event study graph (Figure. 1) from one of the papers by Andersen et al. 2016 where they examined whether plausibly exogenous increases in the number of establishments licensed to sell alcohol by the drink were related to violent crime using county-level data from Kansas for the period of 1977–2011. I have also showed how this event-study graph changes when you consider a different reference/base period.

Also, using Fatality Analysis Reporting System (FARS) data based on the paper by Anderson et al. 2013, I have computed event study graphs to look at the impact of legalization of medical marijuana laws on traffic fatalities. From these graphs, it can be seen that adoption of MML laws do have a significant negative impact on traffic fatalities which is similar to what the researchers found in their study. These event study graphs also point out that this negative effect only becomes significant at the fifth period and remains insignificant at the earlier periods. Additionally, I have shown how to compute event study graphs in Stata using two popular ways namely '*coefplot*' and '*twoway*'.

Lastly, when you have multiple time points and cannot show all the time points in a single graph, I have shown how to compute lags/leads for that. In this case the end time point could be something like 5+ or 5- where in your total periods might be 35 time-points or years (1977–2011 for Anderson et al. 2016) and you only want to show 5 years before and 5 years after the event happened (as in year 0) in your graph.

All these analyses have been done in Stata version16. The code, generated event-study graphs, related literatures and data have been added in this GitHub repository: https://github.com/sbmeco1991/Event-Study-Graphs.