GV300: Quantitative Political Analysis

Problem Set 12

Due Thursday, February 28, 9.45am on Faser

(100 marks) **Difference in difference estimator**
Load data set `indicators.csv` which holds a range of economic indicators for 19 Eastern and Central European countries for the years 1992-2018. The data set also features a variable `euJoin2004`, which is coded `EU Member` for those countries that joined the European Union in 2004, and a variable `yearJoinEU` indicating when a country joined the EU if it did (some joined in 2004, 2007, or 2013, some did not join so far and are coded `9999`). The country-level economic indicators are `GDPPerCapita, exportsShareGDP`, `importsShareGDP`, and `taxRevenueShareGDP`. We are interested in whether joining the EU in 2004 exerts any influence on a country's GDP per capita (the variable is measured in Current $s).

1. (10 marks) From this country-year data set, generate the variables necessary to compute the differences-in-differences estimator. Which countries are in the treatment group, which countries are in the control group, which years should be coded as pre-intervention, which years should be coded as post-intervention?

2. (10 marks) Compute the mean of `GDPPerCapita` for treatment and control group pre- and post-intervention. Plot those numbers. Compute the differences-in-differences from those numbers.

3. (25 marks) Plot the mean of `GDPPerCapita` for treatment and control group over `year`. Add a line indicating the intervention year. Add the counterfactual `GDPPerCapita`. Evaluate whether the common trend assumption is met pre-intervention. Are the parts of SUTVA met that are relevant to the goodness of the differences-in-differences estimator? Why or why not?

4. (15 marks) Run a regression to compute the differences-in-differences estimator. Report and interpret the result. Speak to the three relevant coefficients.

5. (10 marks) Improve your regression in (d) by computing clustered standard errors. Check out the `lm_robust()` command in R (`estimatr`-package) or the entry on the `vce()` option in the helpfile on Stata's `regress` command.

6. (15 marks) Improve your estimate of the causal effect of joining the EU in (e) by including one relevant country-level covariate into the regression. Report and interpret your result. Was your estimate in (e) an over- or underestimate of the causal effect? Speculate why failing to include this covariate led to bias in your estimate in (e).

7. (15 marks) Can you think off an institutional change, exogenous shock, or geographic peculiarity that allows you to apply the differences-in-differences estimator in your own research. Describe what that intervention would be. Speculate about the direction of the causal effect you would estimate with applying the differences-in-differences estimator in your own research.