Economic Costs of Running the Political System

In this exercise, we evaluate the argument on economies of scale in the economic costs of running political systems.

This exercise is based on: Blom-Hansen, Jens, Kurt Houlberg, and Soren Serritzlew. 2014. "Size, Democracy, and the Economic Costs of Running the Political System." *American Journal of Political Science* 58(4): 790-803.

For many years, economists have argued that small-sized municipalities create the potential for welfare gains because public sectors can be better tailored to local preferences. However, a counterargument holds that the large-sized municipalities are more cost-effective due to the economies of scale in the production of many public functions.

If scale effects exist in municipalities, we should, all things equal, observe a negative correlation between municipality size and per capita administrative costs. Testing this argument is not straightforward because of the differences between small and large jurisdictions and omitted variables bias.

The Danish structural reform of 2007 provides an unique opportunity of natural experiment. The reform was quick and radical, merging municipalities within six months, and created a variation between merged (treated) and not merged (control).

We will examine the impact of municipal merger happened in 2007 on administrative costs per capita. Since merging took six months to complete, we consider 2007 as the last year of the pre-treatment period.

The CSV data file, `admin_costs.csv`, contains the variables in the table. Each row is a municipality-year observation and the municipality is based on the period *after* amalgamation.

Note that amalgamation results in a change in the unit of analysis. The control group (not experienced amalgamation) has the same municipalities before and after the structural reform. However, the units in the treatment group change due to merging. In this exercise, we use later merged municipalities as treated units and the outcomes *before* amalgamation are based on their average. For example, suppose a municipality A in this dataset is a result of merging the municipalities B, C, and D. During the pre-treatment period (2005-2007), the outcome `Y` is the average of the outcomes from B, C, and D.

Name	Description
`year`	Year. (pre-treatment period is 2005-2007, post-treatment period is 2008-2011)
`Y`	The outcome variable. Administrative costs per capita.
`treatment`	The municipality is a result of amalgamation (1, treated group) or not (0, control group).
`region`	Region ID (1-5)
`municipality_id`	Municipality ID (there are 98 unique municipalities)

Question 1

We begin by loading the data and complete the following tasks. First, check if the dataset contains any missing values. Next, compute the mean of outcome (`Y`) by region without repeatedly using the `mean()` function.

Question 2

The authors employed the Difference-in-differences (DiD) approach to estimate the change in administrative costs in control and treatment municipalities, before and after the reform.

Create a time-series plot of the mean outcome by year. The horizontal axis is the years whereas the vertical axis is the mean outcome. You should have two lines connecting data points (each data point should be marked by a symbol): one for the treatment group and the other for the control group. Add a vertical dashed line on the year 2007 so that the pre- and post-treatment periods are clear. Lastly, properly label x and y axies. Do not forget to use different colors and add text labels to differentiate two lines.

Question 3

What is the identification assumption for the DiD approach? Does the plot you created in the previous question confirm that the assumption is likely to hold in this application?

Question 4

Using the data from years 2007 (the last year of pre-treatment period) and 2008 (the first year of post-treatment period), compute the DiD estimate of the effect of municipal merger on administrative costs. Report a point estimate as well as its 95% confidence interval. Briefly interpret the result.

Question 5

In Questions 2 and 3, we examined the validity of using the DiD estimator by a visual inspection. This question asks you to investigate the validity by calculating DiD estimates in pre-treatment periods (that is, conduct a placebo test). If the outcomes from both treated and control groups moved in tandem before the amalgamation happened, the DiD estimate should be zero. Follow the following steps to complete this question:

1. Define a function named `calc_did()` that takes the following arguments: `data` (the dataframe used), `pre_year` (the pre-treatment time period used for the DiD estimation), and `post_year` (the post-treatment time period used for the DiD estimation). The `calc_did()` function returns a vector with three elements that correspond to the lower bound of the 95% confidence interval, the point estimate, and the upper bound of the 95% confidence interval.
2. Use the function that you defined in the Step 1 and confirm that your answer is the same as the previous question. Run `calc_did(data = data, pre_year = 2007, post_year = 2008)`.
3. Check the following outcomes and briefly comment on the results: `calc_did(data = data, pre_year = 2006, post_year = 2007)` and `calc_did(data = data, pre_year = 2005, post_year = 2006)`.