EXERCISE SHEET 9: FUJIWARA (2015): VOTING TECHNOLOGY

Main learning goals: 1. Understanding the basic idea behind Regression Discontinuity Design (RDD)

2. Computing and thinking about bandwidths for the forcing variable
3. Investigating discontinuities visually and statistically
4. Estimate the treatment effect

Hand-in: As usual: do-file and the corresponding log-file (saved in .log-format) containing the following steps (numbered in the do-file).

1. Open the dataset `munic.dta` and get familiar with the variables and data structure. Then, generate the following variables:
   - A treatment indicator taking the value of 1 if the forcing variable is larger than 40,500, and 0 otherwise;
   - A variable containing the additional voters for the municipalities above the threshold of 40,500 voters;
   - A variable containing the distance to the threshold, which can be used later on to decide about the bandwidth.
2. Following Fujiwara (2015), cut the forcing variable into equally distant (4000 voters) bins between 500 and 200,000 voters. Then, compute for all bins the average values needed for the variables in figure 2 and 3. Which trade-offs do you face when deciding about the bin size?
3. Replicate figure 2 and figure 3. To do so, combine scatter plots of the binned variables with quadratic fits of the un-binned data using `qfit`. Why do we see the discontinuity only in figure 2 and not in 3? What does this imply for the story and the RD setup? Also check the distribution of voters in 1996 across municipalities of different size and discuss your outcome.
4. Run the regressions for columns (1) to (3) of table 1 and compare the municipalities closely above and below the threshold. Is it a problem in case you find significant differences? Why are we using different bandwidths?
5. Investigate the treatment effect by re-estimating columns (1) to (3) of panel A from table 2. How do you interpret these results?
6. Implement a placebo test of the following form. Assume that you have reason to believe that voters above a cut-off of either 30,000 or 70,000 voters were treated, while those below the cut-off were not treated (just choose one of the two thresholds). Repeat the previous points selectively (focus only on a bandwidth of 10,000 voters) and discuss your findings.
7. Re-estimate panel C of table 2 to investigate the treatment effect on the ideological composition of the legislative. How does the pattern of treatment effects changes along the bandwidth size? Does it differ from what you expected?
8. Set up two dummy variable that split the sample into municipalities with an illiteracy rate below and above the median illiteracy rate. Then, re-estimate column (3) of table 3 for the median dummy. Test whether the coefficients above and below the threshold are significantly different from each other by using `suest` and `lincom`. Discuss your findings. To do so, also check if the results hold when changing the bandwidth to compute the median illiteracy rate.

Further comments: Enjoy your Christmas break – you have earned it!