POLS0012 Causal Analysis: Tutorial Exercise 3

Question 1

In this question, we study a field experiment on voting carried out in Benin in 2001 by Leonard Wantchekon. He investigated what types of campaign messages are most effective in increasing voter turnout in this developing African nation.[1] He persuaded three presidential campaigns to randomly use different types of campaign messages in different villages throughout the campaign. There were 16 villages in this part of the experiment, and randomisation was carried out by blocks. Villages were divided into blocks of 2 villages based on their geographic locations. Within each block, the two villages were assigned to one of two conditions:

- *Public Policy*: Wantchekon describes this treatment condition as: "It was decided that any public policy platform would raise issues pertaining to national unity and peace, eradicating corruption, alleviating poverty, developing agriculture and industry, protecting the rights of women and children, developing rural credit, providing access to the judicial system, protecting the environment, and/or fostering educational reforms."
- *Clientelist*: Wantchekon describes this treatment as: "A clientelist message, by contrast, would take the form of a specific promise to the village, for example, for government patronage jobs or local public goods, such as establishing a new local university or providing financial support for local fishermen or cotton producers."

The data is contained in the file "benin.Rda" and consists of:

- *block*: a variable indicating the block number
- *reg.voters*: a background covariate, the number of registered voters in the village
- *vote.pop*: the outcome variable, the proportion of village that voted
- *treatment*: =1 if in the "clientelist" condition, 0 otherwise

a) Using a suitable regression, perform a balance test with the *reg.voters* variable between villages in the two treatment conditions (ignoring the blocking). What do you conclude?

b) Estimate the average treatment effect and its standard error, using the difference in means estimator (ignoring the blocking). What does the ATE tell you? Is it significant at the 5% level?

[1] For the original paper, see `http://www.nyu.edu/gsas/dept/politics/faculty/wantchekon/research/WP_0331.pdf`

c) Using a suitable regression, estimate the same average treatment effect while controlling for block membership. Are there any differences in the results compared to part (b)? Why or why not?

Code Hints:

- Linear regression can be used to calculate Average Treatment Effects
- `factor()` can be used to turn a categorical variable into a full set of dummy variables , 1 for each category

Question 2

For this question, we will examine the original data from the Tennessee STAR experiment. Recall that in the original study, within schools students were randomly assigned to "small", "regular" or "regular plus aide" classes for four years. To keep things simple, the dataset for this problem only contains students in the "small classes" or "regular classes" conditions for kindergarten and 1st grade only (the first two years of the study). Remember that a key problem with the experiment was that some children left the study early, and others did not comply with their assignments to treatment and control. In this problem we are going to assess both attrition and non-compliance.

The data is contained in the file "star.Rda" and includes the following variables for each student:

- *gkclasstype* - Class type enrolled in kindergarten, the beginning of the study
- *g1classtype* - Class type enrolled in first grade
- *gender* - gender (1 if female, 0 if male)
- *race* - race (1 if black, 0 if not black)
- *gkfreelunch* - gender (1 if child qualifies for free lunches, 0 if not). This is often used a proxy for family poverty
- *gktreadss* - SAT test score in reading at end of kindergarten
- *gktmathss* - SAT test score in maths at end of kindergarten
- *g1treadss* - SAT test score in reading at end of first grade
- *g1tmathss* - SAT test score in maths at end of first grade
- *gkschid* - Unique code identifying each school in the study

a) Create a variable called "treat" that equals 1 if a child was assigned to a small class in kindergarten and 0 otherwise

b) Using t tests, obtain p-values to assess the null hypotheses of no imbalance between the "small class" and "regular class" groups in terms of gender, race or free lunches. What do you conclude about balance between the two groups?

c) Calculate Average Treatment Effects for both maths and reading scores in kindergarten, for children in small classes compared to regular classes. Estimate results with and without controlling for the school attended. Interpret the results, including their statistical significance

d) Calculate the proportion of children who left the experiment (i.e., attrition) between kindergarten and first grade, accounting for both children who are missing from the sample in first grade altogether, and children who did not report outcome data in first grade.
Code Hints:
- Start by creating an indicator variable equalling 1 if a child was missing and 0 otherwise
- `is.na()` selects missing observations of a variable

e) Assess the extent of non-compliance amongst children in first grade by calculating:

i) The proportion of children who were assigned to regular classes in kindergarten that were enrolled in small classes in first grade

ii) The proportion of children who were assigned to small classes in kindergarten that were enrolled in regular classes in first grade

Hint: You need to exclude missing responses in first grade

What type of non-compliance do we have here? Did non-compliance differ by class assignment in kindergarten?

f) Assuming missingness-at-random, calculate ATEs and standard errors for maths and reading scores in the first grade, controlling for school attended. Interpret your results
Code Hint: Note that the `lm()` function automatically drops missing values that are labelled as "NA" in R

g) **[Challenging Question]** Calculate extreme-value bounds for the ATEs in part (e), using the highest and lowest observed test scores. Start by filling in missing values for class types in first grade by assuming that they would have been the same as in kindergarten. Do these bounds suggest that attrition between kindergarten and first grade threatens the validity of the ATEs calculated in (f), or not?

h) **[Challenging Question]** Briefly, explain (using words only) why the two sets of standard errors obtained in part (c) differ depending on whether school is controlled for