QTM 385 - Experimental Methods

Assignment 05

Instructions

As in previous assignments, here are 10 questions related to our recent classes. They cover one-sided and two-sided non-compliance. Please answer all question and export the file as a PDF or html file. If you use R or Jupyter Notebook, make sure to include both the code and the output in the document.

If you have any questions about the assignment, feel free to email me at danilo.freire@emory.edu.

Good luck!

Questions

1. Name the four compliance types in two-sided non-compliance, and provide a brief description of each. Imagine a scenario in which each type of non-compliance could occur.

2. Explain the concept of Two-Stage Least Squares (2SLS) in the context of estimating CACE. Briefly describe the two stages.

3. Discuss the trade-off between internal validity and external validity when focusing on CACE rather than ATE in the presence of non-compliance.

4. The following three quantities are similar in appearance but refer to different things. Describe the differences.

$$E[Y_i(d_i(1))|D_i = 1]$$

$$E[Y_i(d_i(1))|d_i(1) = 1]$$

$$E[Y_i(d_i(1))|d_i(1) = d_i(0) = 1]$$

5. Imagine we have a dataset for a hypothetical experiment evaluating a "Financial Literacy Programme" aimed at improving personal financial management skills. The dataset is available at https://github.com/danilofreire/qtm385/blob/main/assignments/05-assignment/financial.csv. The dataset has the following variables:

- `id`: Unique participant identifier.
- `assignment`: Treatment assignment (0 = control group, 1 = treatment group). Participants assigned to the treatment group were offered access to the programme.
- `programme_Usage`: Actual usage of the Financial Literacy Programme (0 = no usage, 1 = used the programme). This indicates whether participants in either group actually used the programme (treatment). Note that control group members could theoretically access similar resources outside the study, leading to potential two-sided non-compliance aspects, though we will primarily focus on IV analysis suitable for non-compliance scenarios discussed in lectures.
- `financial_knowledge_score`: Score on a financial knowledge test post-experiment (outcome variable). Higher scores indicate better financial knowledge.

Estimate the intent-to-treat (ITT) effect of `assignment` on `financial_knowledge_score` using robust linear regression. Interpret, in detail, the ITT effect in the context of the Financial Literacy Programme.

6. Estimate the ITT effect of `assignment` on `programme_usage` using robust linear regression. Interpret this ITT_D in terms of programme uptake.

7. Use the IV approach to estimate the CACE. Interpret the estimates and compare with the results from the ITT analysis you conducted in questions 5 and 6.

8. In her study of election monitoring in Indonesia, Hyde (2010) randomly assigned international election observers to monitor certain polling stations.Here, we consider a subset of her experiment where approximately 20% of the villages were assigned to the treatment group. Because of difficult terrain and time constraints, observers monitored 68 of the 409 polling places assigned to treatment. Observers also monitored 21 of the 1,562 stations assigned to the control group. The dependent variable here is the number of ballots that were declared invalid by polling station officials.

Explain what the non-interference assumption means in the context of this experiment.

9. Download the sample dataset at https://github.com/danilofreire/qtm385/blob/main/assignments/05-assignment/hyde.csv and estimate the ITT and the CACE. Interpret the results.

10. Use randomisation inference to test the sharp null hypothesis that there is no intent-to-treat effect for any polling location. Interpret the results. Explain why testing the null hypothesis that the ITT is zero for all subjects serves the same purpose as testing the null hypothesis that the ATE is zero for all Compliers.