

Political Science 207

Problem Set 1

Professor: Daniel Masterson

Due Monday April 12th at 5:00pm

Submit your completed assignment as a single PDF or HTML file (including your write-up and all R code) on Gauchospace under the assignment for Week 2. I strongly recommend that you write your problem sets in R Markdown. ([I recommend this tutorial.](#))

Make sure you follow [good coding style \(see tutorial in link\)](#), and show your work for problems that require calculations.

You are encouraged to work in groups, but you should write up the problem set alone, and you should note at the top of your problem set who you worked with.

Problem 1

This problem provides a quick review of the basic potential outcomes notation we discussed in class.

Suppose the government of Pakistan institutes a new reform that provides Rs. 1,000 per month to households in the country whose monthly income is below Rs. 5,000. The objective of the reform is to improve child enrollment in schools. It is believed that reducing household budget constraints (through the income supplement) will allow households to put more children in school.

- Suppose we label this reform as a binary treatment $D \in \{0, 1\}$, for all households in Pakistan indexed by i . We call the share of children in school in every household Y . What is the meaning of Y_{1i} and Y_{0i} ?
- Define the average treatment effect (ATE). Describe in words what the ATE means for the program. Let's say that analysts collected some survey data of a sub-sample of potential beneficiaries and estimated the average treatment effect. Let's call this \widehat{ATE} . What is the difference between the ATE and \widehat{ATE} ?
- What is the Average Treatment on the Treated (ATT)? Describe in words what the ATT means for the program.
- When will the ATT and the ATE be equal to each other? Prove it.
- Show formally that $\tau_{ATE(x)} = E[Y_1 - Y_0 | X = x]$ (i.e. a subgroup average treatment effect) is identifiable given random assignment.

Problem 2

To reinforce the intuition behind the potential outcomes framework, consider the fictional data set “PO-data.csv.” In these fictional data, we observe an outcome for each unit both under treatment and under control (which, again, is usually impossible in the real world).

- (a) Write down the formula for individual level treatment effects and explain the fundamental problem of causal inference.
- (b) Define the Average Treatment Effect (ATE) and calculate the ATE in these data.
- (c) Plot the distribution of the individual treatment effects. Does the treatment seem to have an effect? How well is it captured by the ATE?