Applied Microeconometrics I
Discussion of Problem Set 1

Atte Pudas

Aalto University

20/09/2023

Exercise 1.1

What type of research questions can in principle be empirically addressed? What type of questions do typically economists care about? In case you characterize one type of question as higher in the priority list of economists, can the other questions still be considered relevant?

- Descriptive, predictive, and causal
- Causal questions are those of greatest interest to economists (they have the potential of being policy relevant)
- Other questions can be useful. E.g., descriptive evidence of new phenomena often is as a preliminary step before casting the question in a causal inference setting.

Exercise 1.2

Consider the following questions:
a) *The Ministry of Environment wants to know the expected number of electric vehicles used by 2030.*
b) *The Ministry of Environment wants to know whether providing subsidies to consumers to buy new electric cars promotes electric vehicle adoption.*
c) *The Ministry of Environment wants to know whether electric car sales are larger in urban than rural areas.*

Would you characterize these questions as descriptive, predictive, or causal?

- a) predictive; b) causal; c) descriptive

Exercise 1.3

Explain the ideal randomized controlled trial (RCT) that would allow one to analyse the question(s) above that you characterized as causal. Briefly discuss potential problems with the implementation of RCT in the given setting of the exercise.

- Lottery that randomly distributes vouchers to buy an electric car.
- Potential problems: Hawthorne effect, general equilibrium effects, etc.

Exercise 2.1

Exlain the notation of $Y_i(0)$.

- The **potential** outcome for subject i if this subject were untreated.
- Another way to put it: the untreated potential outcome for subject i.

Exercise 2.2

Contrast the meaning of $Y_i(0)$ *with* Y_i.

- $Y_i(0)$ is the potential outcome for subject i if this subject were untreated.
- Y_i is simply the **observed** outcome for subject i.

Exercise 2.3

Contrast the meaning of $Y_i(0)$ with the meaning of $Y_i(1)$. Is it ever possible to observe both at the same time? Why?

- $Y_i(0)$ is the potential outcome for subject i if the subject were untreated. $Y_i(1)$ is the potential outcome for subject i if the subject were treated.
- In any moment, only one of the two potential outcomes for i can be realized. A subject cannot treated and untreated simultaneously, so observing both potential outcomes is not possible. This is known as the "fundamental problem of causal inference".

Exercise 2.4

Explain the notation $E[Y_i(0)|D = 1]$, *where* D_i *is a binary variable that gives the treatment status for subject* i, 1 *if treated*, 0 *if control.*

- The expected value of the **potential** outcome for subject i if the subject were untreated, given that this subject actually receives treatment.
- Another way to put it: the expected value of the untreated potential outcome for a subject in the treatment group.

Exercise 2.5

Contrast the meaning of $E[Y_i(0)]$ *with the meaning of* $E[Y_i|D = 0]$.

- $E[Y_i(0)]$ is the expected value of the untreated **potential** outcome for subject i
- $E[Y_i|D = 0]$ is the expected value of **observed** outcome for subject i, given that they were untreated.

Exercise 2.6

Contrast the meaning of $E[Y_i(0)|D = 1]$ *with the meaning of* $E[Y_i(0)|D = 0]$.

- $E[Y_i(0)|D = 1]$ is the expected value of the untreated **potential** outcome for a subject who actually receives treatment.
- $E[Y_i(0)|D = 0]$ is the expected value of the untreated **potential** outcome for a subject who is in fact untreated.

Exercise 2.7

Which of the following four quantities (that you explained in parts 2.4. through 2.6.), can be identified from observed information? Why?

1. $E[Y_i(0)|D = 1]$
2. $E[Y_i(0)]$
3. $E[Y_i|D = 0]$
4. $E[Y_i(0)|D = 0]$

- We can identify 3. and 4.
- 3. is the expected **observed** value for the untreated subject i, and under SUTVA (= potential outcomes for each subject i respond only to own treatment status and are unrelated to the treatment status of others) 3. and 4. are equal.

Exercise 2.8

Now, assume that D_i (the treatment) is randomly assigned to the units in this sample. Which of the below quantities can be identified from the observed data? Why?

1. $E[Y_i(0)|D = 1]$
2. $E[Y_i(0)]$
3. $E[Y_i|D = 0]$
4. $E[Y_i(0)|D = 0]$

- All of them.
- Due to random assignment, the expected potential outcomes are the same for the treated and untreated.

Exercise 3

Eisenberg et al. (2004) study whether teenagers in families who eat together more often have better health and well-being, as measured by, among others, tobacco, alcohol and marijuana use; academic performance and depressive symptoms. To this aim, they survey 4746 teenagers from the Minneapolis/St Paul, Minnesota metropolitan area in the US. They run logistic regressions and control for family connectedness and sociodemographic variables. They find that frequency of family meals was inversely related with tobacco, alcohol and marijuana use, low grade point average and depressive symptoms.

Exercise 3.1

Consider a policy maker who must decide whether to encourage families to eat together. Does the paper by Eisenberg et al. (2004) provide support for this decision?

- No. Evidence descriptive, not causal, may be driven by selection.

Exercise 3.2

Now let us consider the health perspective. Based on this evidence, would it be correct to conclude that family meals are good for adolescent health (relative to no family meals)? The observed difference in health for family meals vs. no family meals can be characterized with the potential outcomes framework below. Explain briefly which assumption (if any) might not be satisfied in this specific example and why.

$$\underbrace{E[Y_i|D_i=1]-E[Y_i|D_i=0]}_{Observed\ difference\ in\ health}$$
$$=\underbrace{E[Y_{1i}|D_i=1]-E[Y_{0i}|D_i=1]}_{Average\ effect\ on\ the\ treated}+\underbrace{E[Y_{0i}|D_i=1]-E[Y_{0i}|D_i=0]}_{Selection\ bias}$$

- Selection bias. Those who choose to have family meals would probably have different health outcomes in the absence of family meals compared to the health outcomes of those who choose not to have family meals (maybe more positive?)
- For example, the authors note that "youth who engage in substance use or perform poorly in school may avoid eating meals with family members to avoid discussion of 'problem' behavior."

Exercise 3.3

In this context where family meals is the treatment and adolescent health is the outcome variable, would you expect the population average treatment effect (ATE) to be similar to the average treatment effect on the treated (ATET)?

- No. The treatment is not randomized.
- Those who have chosen (or those whose parents have chosen) to have family meals (the treated) probably react differently to family meals than those who have chosen not to have family meals (the reaction could be negative).

Exercise 3.4

Suppose that the policy maker has started a collaboration with a school to promote the use of family meals. What would be the ideal randomized control trial (RCT) that you would conduct in order to find out if family meals are good for adolescent health (relative to no family meals) in this setting? Are there specific problems that you envision?

- Encouragement design where people/families are randomly selected to be informed about how good it is to have family meals. This could happen by mail, phone interview, invitation to participate, etc.
- Potential problems: low take up rate of the treatment, GE effects, Hawthorne effect, etc.

Exercise 4

From 2015-2019, Iceland ran two large-scale trials, involving 2,500 workers – over 1% of Iceland's entire working population – of a reduced working week of 35-36 hours with no reduction in pay. The results led to important lessons for both employees and businesses. Since the completion of the trial, 86% of the country's workforce are now working shorter hours or gaining the right to shorten their hours. The analysis of the RCT showed that productivity and service provision remained the same or improved across most trial workplaces. At the same time, worker well-being increased across a range of indicators, from perceived stress and burnout to health and work-life balance.

Exercise 4.1

Discuss the potential existence of a Hawthorne effect.

- People may behave differently when observed. The treated may care about others' perceptions or, especially in this case, may want to affect the results of the study.

Exercise 4.2

Discuss the potential existence of general equilibrium effects

- The treatment assigned to other units may affect the potential outcomes of those who are not treated. Maybe work became more stressful for the control group because some of the co-workers worked less?