

# EC 421 Classwork 5 (analytical)

## 1. Causal Inference and Bias

### 1a) Causal effect of education on earnings

Suppose we took survey data that included people's earnings and the number of years of education they have, and we estimated this model:

$$\text{Earnings}_i = \beta_0 + \beta_1 \text{Education}_i + u_i$$

And our estimate for  $\beta_1$  is very large and positive.

Think of one unobservable variable that is absorbed in  $u_i$  that correlates with both Earnings and Education. By omitting that variable, do we expect  $\hat{\beta}_1$  will be biased? What would an ideal experiment look like to find the causal effect of education on earnings?

### 1b) Causal effect of health insurance on health

Suppose we took survey data that included measures of people's health and whether they have health insurance or not, and we estimated this model:

$$\text{Health}_i = \beta_0 + \beta_1 \text{Health Insurance}_i + u_i$$

And our estimate for  $\beta_1$  is large and negative.

Think of one unobservable variable that is absorbed in  $u_i$  that correlates with both Health and Health Insurance. By omitting that variable, do we expect  $\hat{\beta}_1$  will be biased? What would an ideal experiment look like to find the causal effect of health insurance on health?

### 1c) Causal effect of Migration on Earnings

Suppose we took survey data that included people's earnings and whether they have recently moved to a new city or not, and we estimated this model:

$$\text{Earnings}_i = \beta_0 + \beta_1 \text{Migration}_i + u_i$$

And our estimate for  $\beta_1$  is large and positive.

Think of one unobservable variable that is absorbed in  $u_i$  that correlates with both Earnings and Migration. By omitting that variable, do we expect  $\hat{\beta}_1$  will be biased? What would an ideal experiment look like to find the causal effect of migration on earnings?

### 1d) Causal effect of friends of the opposite sex on a high school student's GPA

Suppose we took survey data that included high school students' GPAs and the number of friends of the opposite sex they have, and we estimated this model:

$$\text{GPA}_i = \beta_0 + \beta_1 \text{Opposite Sex Friends}_i + u_i$$

And our estimate for  $\beta_1$  is large and negative.

Think of one unobservable variable that is absorbed in  $u_i$  that correlates with both GPA and Opposite Sex Friends. By omitting that variable, do we expect  $\hat{\beta}_1$  will be biased? What would an ideal experiment look like to find the causal effect of Opposite Sex Friends on GPA?

**1e) Think of your own example!**

Make sure it's very different from any example we've used. It will also be useful to make your dependent (y) variable be continuous, not discrete. What's the difference? Continuous variables are measured while discrete variables are counted.

Causal effect of \_\_\_\_\_ on \_\_\_\_\_. Think of one unobservable variable that is absorbed in  $u_i$  that correlates with both variables. By omitting that variable, do we expect  $\hat{\beta}_1$  will be biased? What would the ideal experiment look like to find the causal effect?

**2. Maria and Khudzar**

Suppose the tables below report the health scores for two people (Maria and Khudzar) in parallel universes:  $Y_1$  is the individual's health if they are treated (that is, they decide to buy health insurance), and  $Y_0$  is the individual's health if they are not treated (they decide to not buy health insurance).

Calculate the individual treatment effects  $\tau_i$  and the average treatment effect  $\bar{\tau}$ . What would we calculate  $\hat{\beta}_1$  to be if we observe only Maria in  $Y_0$  and Khudzar in  $Y_1$ ?

Using the formula  $\hat{\beta}_1 = \bar{\tau} + \text{selection bias}$ , calculate the selection bias in the estimate for  $\hat{\beta}_1$ .

Using the definition of selection bias (the difference in the expected healths of the treatment and control groups, in a universe where the level of treatment is the same for both groups), show another way to get the same value for selection bias in the estimate for  $\hat{\beta}_1$ .

	Maria	Khudzar
$Y_1$	9	5
$Y_0$	8	4

**Extra Credit: Varying Individual Treatment Effects**

Answer the same questions as before, but with a new table of health scores:

	Maria	Khudzar
$Y_1$	9	4
$Y_0$	8	2

How can selection bias be calculated when the individual treatment effects are not constant?