Potential Outcomes, Observed Outcomes, Treatment and Selection Effects

Yona Rubinstein

July 2016

Outcomes: Random Variables

- Outcomes, that is wages, profits, etc., are random variables.
- A *random variable*, usually written X, Y, D, etc., is a variable whose possible values are numerical outcomes of a random phenomenon.
- There are **two** *types* of random variables, *discrete* and *continuous*.
- The two main characteristics of a random variable are the expected value, the "mean" $E(Y_i)$, and the variance σ_Y^2.
- The mean of a random variable – more commonly referred to as its expected value – is the value you expect to obtain should you carry out some experiment whose outcomes are represented by the random variable.
- The outcome of the i observation – Y_i – can be written as the sum of the expected value and the deviation from the expected value (U_i):

$$Y_i = E(Y_i) + U_i. \quad (1)$$

- By construction the expected value of U_i, that is $E(U_i)$, equal to zero.

Conditional Means

- Suppose X_i and Y_i are random variables. Then, the mean of Y_i conditional on $X_i = x$ is defined as:

$$E\left(Y_i | X_i = x\right).$$

- We can use also the following notation:

$$E\left(Y_i | X_i = x\right) = Y_x.$$

- The conditional mean of Y_i is defined as th
- e mean of Y_i for a subset of the population for which X_i equals to x.
- The outcome of the i observation for whom $X_i = x$ can be written as the sum of the expected value and a zero mean deviation from the expected value (U_{xi}):

$$Y_{x,i} = E\left(Y_i | X_i = x\right) + U_i = Y_x + U_{xi}, \tag{2}$$

where $E\left(U_i | X_i = x\right) = U_{xi} = 0.$

Conditional Means: Example

- The population of workers consists of those who work in the *financial industry* and *all others*:

$$
\begin{aligned}
Y_i &= \text{log hourly wage of worker } i \\
F_i &= \text{sector, categorical, two values: 1 finance; 0 other}
\end{aligned}
$$

- The expected log hourly wage of a random worker employed in the financial sector and outside the financial sector are:

$$
\begin{aligned}
E\left(Y_i | F_i = 1\right) &= Y_1; \\
E\left(Y_i | F_i = 0\right) &= Y_0.
\end{aligned}
$$

- The log hourly wage of worker i in the financial sector and the log hourly wage of worker i outside the financial sector (not the same person) are:

$$
\begin{aligned}
Y_{1,i} &= Y_1 + U_{1i}; \\
Y_{0,i} &= Y_0 + U_{0i}.
\end{aligned}
$$

Observed Outcomes

- The expected log hourly wage of a random worker i can be written as:can be written as:

$$E\left(Y_i|F_i\right) = \left\{ \begin{array}{ccc} Y_0 & \text{if} & F_i = 0 \\ Y_1 & \text{if} & F_i = 1 \end{array} \right\}.$$

- Thus the expected value of Y - the mean log hourly wage in the population conditional on employment sector - equals:

$$E\left(Y_i|F_i\right) = F_i Y_1 + \left(1 - F_i\right) Y_0.$$

- This can be also written as:

$$\begin{array}{ccccccc} E\left(Y_i|F_i\right) & = & \underbrace{Y_0} & + & \underbrace{\left(Y_1 - Y_0\right)} & \cdot F_i \\ E\left(Y_i|F_i\right) & = & b_0 & + & b_F & \cdot F_i \end{array}$$

Observed Outcomes (cont.)

- Therefore the log hourly wage of worker i can be expressed as:

$$Y_{Fi} = b_0 + b_F F_i + U_{Fi}, \tag{3}$$

- Where:

$$U_{Fi} = \left(Y_{1i} - Y_1 \right) F_i + \left(Y_{0i} - Y_0 \right) \left(1 - F_i \right).$$

- What do b_0 and b_F mean? What is U_{Fi}?
- These are:

$$\begin{array}{lllll} b_0 & = & Y_0 & : & \text{the mean log hourly wages for those who work} \\ b_F & = & Y_F - Y_0 & : & \text{the gap in mean log hourly wages.} \\ U_{Fi} & = & & : & \text{the gap between worker's } i \text{ outcome and her g} \end{array}$$

Treatment and Outcomes

- Does finance pay higher wages?
- In other words, what is the *causal impact* of working on the financial industry on pay?
- Thinking of the about employment in finance as a "*treatment*" the *question of interest* is about the *treatment effect* of finance on wages.
- The finance treatment is described by a binary variable F_i. that equals 1 if worker i is treated – that is works at the finance industry – and 0 if worker i is not treated – that is does not work in the finance industry:

$$F_i = \{0, 1\}.$$

- The outcome of interest - in this example a measure of pay such as log hourly wages – is denoted by Y_i.

Treatments and Potential Outcomes

- Workers can be employed in the financial sector and outside the financial sector.
- For any worker i there are two potential outcomes – that is two potential log hourly wages – one in each sector:

1. The log hourly wage if she works in outside of the financial sector is Y_i^0;
2. The log hourly wage if she works in the financial sector is Y_i^1,

- Thus the potential outcome can be described as follows:

$$\text{potential outcome} = \begin{cases} Y_i^0 & \text{is the log hourly wage of worker } i \text{ in the n} \\ Y_i^1 & \text{is the log hourly wage of worker } i \text{ in the fi} \end{cases}$$

- The *causal impact* of finance on the log hourly wage of worker i is simply the difference between these two potential outcomes:

$$\text{the treatment effect for } i = Y_i^1 - Y_i^0.$$

The Average Causal Effect

- Let's denote by β_{Fi} the treatment effect of the financial sector (indicated by F) on the log hourly wage of worker i.
- Thus, the treatment effect of finance for worker i is:

$$\text{the treatment effect for } i = \beta_{Fi} = Y_i^1 - Y_i^0.$$

- The average causal effect, also known as the <u>*Average Treatment Effect*</u> (ATE), is the mean of β_{Fi} in the population:

$$E\left[\beta_{Fi}\right] = E\left[Y_i^1\right] - E\left[Y_i^0\right] = Y^1 - Y^0.$$

- Y^0 and Y^1 are the population mean potential wages outside of the finance industry and in the finance industry respectively.

Potential Outcomes and Causal Effects

- Can we express the *potential outcomes* in terms of *treatment effects*?
- Yes, we can.
- The potential outcome can be written in terms of the *potential outcomes without treatment* and the *treatment effect* using as a **switching outcomes model**:

$$\text{potential outcome of worker } i = Y_i^0 + \beta_{Fi} \cdot F_i \tag{4}$$

Actual Outcomes in Terms of Potential Outcomes

- The **observed outcome** Y_i can be expressed in terms of the *potential outcomes* and the *treatment effect*:

$$Y_i = \left\{ \begin{array}{ccccccc} Y_{0i} & = & Y_{0i}^0 & = & Y_i^0 & \text{if} & F_i = 0 \\ Y_{1i} & = & Y_{1i}^1 & = & Y_i^1 & \text{if} & F_i = 1 \end{array} \right\}.$$

- Therefore, the log hourly wage of worker i is:

$$Y_i = Y_i^0 + \beta_{Fi} F_i. \tag{5}$$

- This notation is useful because β_{Fi} is in fact the *causal effect* of the financial sector on worker's i log hourly wage.
- However, since we never observe **both** *potential outcomes* for each worker we won't be able to estimate β_{Fi} directly by comparing outcomes for each person in different states of treatment.

Constant Treatment Effects, Absolute Advantage and the Outcome Equation

- With **common treatment effect** – when the *causal impact* of working in the *finance sector* on log hourly wages is the same to all – the *treatment effect* equals to the **ATE**:

$$\beta_{Fi} = \beta_F .$$

- Note that this means: $U_i^0 = \left(Y_i^0 - Y^0\right) = \left(Y_i^1 - Y^1\right) = U_i^1$.
- In this case outcomes can be expressed as:

$$\begin{array}{ccccccccc} Y_i & = & \underbrace{Y^0} & + & \underbrace{ATE} & \cdot F_i & + & \underbrace{Y_i^0 - Y^0}; \\ & & \parallel & & \parallel & & & \parallel \\ Y_i & = & \beta_0 & + & \beta_F & \cdot F_i & + & U_{Fi}. \end{array}$$

- We can represent worker's *i* *outcomes* in terms of her **potential outcomes** and her **choices**:

$$Y_{Fi} = \beta_0 + \beta_F F_i + U_{Fi}. \tag{6}$$

Treatment and Selection Effects

- Since we never observe **both** *potential outcomes* for each worker we must **infer** regarding the *effects* of finance on log hourly wages by **comparing** the wages of those who work in and outside the financial sector.
- A good starting point is a **naive comparison** of averages by sector of employment. We can express the naive means using the actual outcomes equation in (6):

$$\begin{aligned} E\left(Y_i|F_i=0\right) &= Y_0=\beta_0+U_0 \\ E\left(Y_i|F_i=1\right) &= Y_1=\beta_0+\beta_F+U_1 \end{aligned} \tag{7}$$

- Note that $U_0 = E\left(U_i|F_i=0\right)$ and $U_1 = E\left(U_i|F_i=1\right)$ equal to:

$$\begin{aligned} E\left(U_i|F_i=0\right) &= U_0=Y_0^0-Y^0 \\ E\left(U_i|F_i=1\right) &= U_1=Y_1^0-Y^0 \end{aligned}$$

Treatment and Selection Effects

- The gap in mean log hourly wages in equation (7) can be expressed in terms of the *ATE* and the *selection effect*:

$$\begin{array}{ccccccc} Y_1 - Y_0 & = & b_F & = & \beta_F & + & (U_1 - U_0) \\ & & & & \shortparallel & & \shortparallel \\ & & b_F & = & ATE & + & \left(Y_1^0 - Y_0^0\right) \end{array}$$

- The first term is the *average treatment effect*.
- The second term is the *selection effect* $(U_1 - U_0)$, the gap in potential earnings between those who work in the financial sector and all others as measured outside of finance **without any treatment**.
- The selection effect - known also as the *selection bias* – can go both ways:

$$Selection\ Bias = (U_1 - U_0) = \left(Y_1^0 - Y_0^0\right) \gtreqless 0$$

- If workers in the financial sector posses better labor market skills then workers outside this sector we would expect the bias to be positive.

Observed and Unobserved by the Econometrician

- Can we estimate the bias term?

$$Selection\ Bias = Y_1^0 - Y_0^0$$

- Note that both Y_0 and Y_1 are observed in the data:

$$\begin{aligned} E\left[Y_i^0 \mid F_i = 0\right] &= Y_0^0 = Y_0; \\ E\left[Y_i^1 \mid F_i = 1\right] &= Y_1^1 = Y_1. \end{aligned}$$

- Yet Y_1^0 as well as Y_0^1 are not observed in the data.

The What If...

- The expected/average wage of those workers who are not employed in the finance industry, if they would have been working in the finance industry:

$$E\left[Y_i^1 \mid F_i = 0\right] = Y_0^1.$$

- The expected/average wages of those workers who work in finance, if they would have been working outside of finance:

$$E\left[Y_i^0 \mid F_i = 1\right] = Y_1^0.$$

- Note that both Y_0^1 and Y_1^0 are the *what if terms*. These are **NOT observed in the data**.
- Therefore since we do not observe the what if term the selection term is unobserved:

$$\begin{array}{ccccc} \textit{Selection Bias} & = & Y_1^0 & - & Y_0^0 \\ & & \shortparallel & & \shortparallel \\ & & \textbf{unobserved} & - & \textbf{observed} \end{array}$$

The Compound Effect

- We observe only the actual outcomes, that is Y_0 and Y_1.
- Using the above let us formulate the link between average wages *conditional* on industry choices and the average causal effect:

$$Y_1 - Y_0 = \left(Y_1^1 - Y_1^0\right) + \left(Y_1^0 - Y_0^0\right) \tag{8}$$

- What do these terms, $\left(Y_1^1 - Y_1^0\right)$ and $\left(Y_1^0 - Y_0^0\right)$, mean?
- Assuming constant treatment effect what is the term $\left(Y_1^1 - Y_1^0\right)$?

Heterogenous Treatment Effects, Comparative Advantage and the Outcome Equation

- So far we assumed *common treatment effect*.
- Now let's allow for **comparative advantage**. In this case, treatment effects **are not** *common* anymore. Some benefit more than others.
- A way to represent that - with a common coefficient model - is to allow the error term (U_{Fi}) to take the following form:

$$U_{Fi} = U_i^0 + \left(U_i^1 - U_i^0\right) F_i \tag{9}$$

- Note that if $\left(U_i^1 = U_i^0\right)$ then we are back at the **absolute advantage** model.
- The equation representing worker's i *outcomes* in terms of her **potential outcomes** and her **choices** (equation 6) does not change:

$$Y_{Fi} = \beta_0 + \beta_F F_i + U_{Fi}, \tag{10}$$

where:

$$\beta_F = Y^1 - Y^0.$$

The Gap in Mean Outcomes, Treatment and Selection Effects

- The gap in actual mean outcomes equals to:

$$Y_1^1 - Y_0^0 = Y_1 - Y_0 = \beta_F + \left(U_1^0 - U_0^0\right) + \left(U_1^1 - U_1^0\right) \tag{11}$$

- Note that this gap is always equal to:

$$Y_1^1 - Y_0^0 = Y_1 - Y_0 = \beta_F + \left(Y_1^0 - Y_0^0\right) + \left(U_1^1 - U_1^0\right) \tag{12}$$

- The **first term** (β_F) is the common treatment effect - the **a**verage **t**reatment **e**ffect - also known as the *ATE*.
- The **second term** is the selection on *absolute advantage* $\left(Y_1^0 - Y_0^0\right)$
- What is the **third term** $\left(U_1^1 - U_1^0\right)$? This term reflect the **comparative advantage** of the financiers in finance.

$$U_1^1 - U_1^0 = \left(Y_1^1 - Y_1^0\right) - \left(Y^1 - Y^0\right) \tag{13}$$

The Observed Gap, Treatment and Selection Effects

- Let's denote the causal impact of finance on those who work in finance by β_{1F}:

$$\beta_{1F} = \left(Y_1^1 - Y_1^0\right) - \left(Y^1 - Y^0\right)$$

- Then, the gap in observed wages is equal to:

$$Y_1 - Y_0 = b_F = \beta_F + (\beta_{1F} - \beta_F) + \left(Y_1^0 - Y_0^0\right) \tag{14}$$

- Hence, the gap in mean outcomes reflects three factors:

1. Average treatment effect: β_F.
2. Selection on *comparative advantage*: $(\beta_{1F} - \beta_F)$.
3. Selection on *absolute advantage*: $\left(Y_1^0 - Y_0^0\right)$.

Treatment Effects

- So far we assumed constant treatment effect. Let's realse this assumption.
- The first term in the right hand side $\left(Y_1^1 - Y_1^0\right)$ equals to the average causal effect of finance on the log hourly wages of those workers employed in the financial sector.
- This term is also known as the *Average Treatment Effect on the Treated* (ATET):

$$ATET = \left(Y_1^1 - Y_1^0\right).$$

- It is a "fair" question of interest. Note that it does not address/provide the returns to those who work outside finance if they would have been working in finance, the *Average Treatment Effect on the Not Treated*, which is:

$$ATENT = \left(Y_0^1 - Y_0^0\right).$$

Few More Words About the Selection Bias

- The *selection bias* term reflects *pre-determined* factors, unobserved by the econometrician, that influence potential wage rates.(or other outcomes of interest). For instance cognitive and non-cognitive traits.
- These latent factors may also explain who works where. For instance, the financial sector might have larger demand for "smart" workers than other sectors.
- In that case workers in the financial sector may posses, on average, better cognitive abilities than workers outside finance.
- If these traits are *observed* by the **employers** and the **labor market** *values* these traits, we may find that workers in finance earn more than other workers even if the financial sector **does not** pay higher wage rates per *effective unit of labor*.