Appendix B
Identification Problem and Randomized Experiments

1. Introduction

Most empirical questions in economics are simple cause and effect questions, such as: Does education increase earnings? Does training reduce poverty? Do minimum wage laws increase unemployment? Does class size affect students' grades? Do divorce laws affect domestic violence?

Each of these questions is of the type: Does *X* cause *Y*? The challenge in answering this question is that correlation does not imply causation. For example, education and earnings tend to be positively correlated. Yet, we cannot conclude from this correlation that education increases earnings because something else, such as ability, may cause both higher educational attainment and higher earnings. As another example, health of people in hospitals is usually worse than health of people outside hospitals. Yet, it would be absurd to conclude from this that hospitals make people sick!

If you are interested in influencing *Y* (e.g. earnings, poverty, and unemployment), it is important to know whether *X* causes *Y* or whether *X* and *Y* are just correlated. For example, suppose that you are interested in achieving a high grade in this course. If the class attendance has a causal effect on grades, then attending as many classes as possible will help you achieve your goal. However, if the class attendance and grades are merely correlated, then attending the class will not influence what grade you get in the course.

The objective of this appendix is to help you better understand whether and how it may be possible to establish that *X* causes *Y*. We start in the following section by asking what does it exactly mean to say that *X* causes *Y* and why a simple correlation between *X* and *Y* is not sufficient to establish causality.

2. Identification Problem

To focus ideas, suppose you are a manager in a large company in which workers are paid by salary. You wish to increase productivity of your employees, and you've heard that piece rate workers tend to be more productive than salary workers. Could this work for your company?

Let's start with some definitions. The **outcome** you are interested in is the workers' productivity. The **treatment** you are considering is switching to the piece rate system. The question is whether the treatment has a causal effect on the outcome.

Intuitively, we are asking whether the same salary worker would be more productive if, at this very moment, she was a piece rate worker. If she would, we can say that switching to the piece rate system has a causal effect on this worker's productivity.

Formally, let y_i^1 represent the outcome if individual *i* receives the treatment and let y_i^0 be the outcome if individual *i* does not receive the treatment. The causal effect of the treatment for individual *i* is then defined as $y_i^1 - y_i^0$. The **average treatment effect** (ATE) for the entire population of workers is defined by $E[y_i^1 - y_i^0]$, where *E*[.] denotes the expectation operator.

Example 1. Suppose that the worker's productivity can be represented by

$$y_i^0 = \alpha + u_i$$
$$y_i^1 = \alpha + \beta + u_i$$

y_i^0 *denotes the worker's productivity if she is paid by salary and* y_i^1 *is her productivity if she is paid by piece rate.* α *represents the part of productivity that is common to all workers (e.g. productivity due to simply being present in the office during regular work hours).* u_i is *the individual-specific determinant of productivity, such as ability. The causal effect of switching to the piece rate pay for individual i is then* $y_i^1 - y_i^0 = \beta$*. The ATE is also equal to* β *because it does not vary by individuals.*

Observed and Potential Outcomes

For any given individual *i*, we observe only y_i^0 or y_i^1, but not both. For example, we observe the productivity of a salary worker (y_i^0) but we don't observe her productivity, at this very same moment, if she was a piece rate worker (y_i^1). For this worker, y_i^0 is the **observed outcome** while y_i^1 is the **counterfactual**. Similarly, we observe the productivity of a piece rate worker (y_i^1) but not her productivity, at this very same moment, if she was a salary worker (y_i^0). In this case, y_i^1 is the observed outcome while y_i^0 is the counterfactual.

Therefore, y_i^0 and y_i^1 are sometimes called **potential outcomes**. The observed outcome *y* can be written as:

$$\begin{aligned} y_i &= d_i y_i^1 + (1 - d_i) y_i^0 \\ &= y_i^0 + (y_i^1 - y_i^0) d_i \end{aligned}$$

where *d* is an indicator function equal to $d_i = 1$ if the individual receives the treatment and $d_i = 0$ if the individual does not receive the treatment.

Example 2. Given the assumptions about y_i^0 *and* y_i^1 *from Example 1, the observed productivity of worker i can be written as:*

$$\begin{aligned} y_i &= y_i^0 + (y_i^1 - y_i^0) d_i \\ &= \alpha + u_i + (\alpha + \beta + u_i - \alpha - u_i) d_i \\ &= \alpha + \beta d_i + u_i \end{aligned}$$

Selection Bias

The fact that we cannot observe the counterfactual outcome implies that the causal effect of treatment cannot be observed for any individual. Given this difficulty, we may try to use the observed outcome of the non-treated individuals to approximate the counterfactual outcome for the treated individuals. In other words, we can compare observed outcome for treated and untreated individuals. Can this approach help us learn about the causal effect of treatment?

In general, the answer is no. To see this, suppose that y_i^0 and y_i^1 are as in Example 1. The observed outcome is then:

$$y_i = \alpha + \beta d_i + u_i$$

The comparison of outcomes for treated and untreated individuals is then:

$$E[y_i \mid d_i = 1] - E[y_i \mid d_i = 0]$$
$$= \beta + E[u_i \mid d_i = 1] - E[u_i \mid d_i = 0]$$

The term $E[u_i \mid d_i = 1] - E[u_i \mid d_i = 0]$ is known as **the selection bias** and represents the difference in outcomes between treated and untreated individuals that is not caused by the treatment. If the treated and untreated individuals are sufficiently similar, this term will be zero. In other words, the observed outcome for the untreated individuals would be a good approximation to the counterfactual outcome for the treated individuals. The comparison of outcomes for treated and untreated individuals would then yield the true causal effect of treatment β.

However, in most cases individuals select whether they receive the treatment or not. In this case, the selection bias term will not be zero and the comparison of outcomes for treated and untreated individuals will not identify β.

To illustrate, more able individuals may be more likely to work for the piece rate firms than less able individuals. This is an example of positive selection, since more able individuals are likely to be more productive than salaried workers, irrespective of how they are paid. In other words, the difference in productivity between the piece rate and salaried workers captures not only the causal effect of switching to the piece rate pay but also the difference in ability between piece rate and salaried workers.

Example 3. Suppose you have data on productivity of piece rate workers and salaried workers. The potential and observed outcomes are shown in the following table, with the observed outcomes highlighted in bold:

	y_i^0	y_i^1
Piece Rate	20	**25**
Salaried	**5**	10

The average treatment effect in this example is 5 (25-20=10-5), but the difference in observed outcomes is 20 (25-5). The selection bias is equal to 15 (20-5), the difference that would exist even if both groups were paid using the same payment method.

Identification Problem

The problem of estimating the true causal effect of treatment is known as **the identification problem**. This problem arises because we cannot observe the counterfactual outcomes for **the treatment group** of individuals. To approximate this counterfactual, we may use the observed outcomes for the untreated or **control group**. This represents a problem when the treatment and control groups are not sufficiently similar.

The potential solution to the identification problem lies in finding creative ways to make the treatment and control groups similar. Each of these ways is known as an empirical strategy. In the following section, we will consider merits and shortcomings of randomized experiments as an empirical strategy.

3. Randomized Experiments

Randomized experiments represent one of the most credible empirical strategies.

This strategy is based on the observation that individuals in the treatment and control groups tend to be different because they select whether to receive the treatment or not. In randomized experiments, this choice is taken from individuals and they are randomly assigned to either the treatment or the control group. This randomization ensures that individuals in the two groups are sufficiently similar, or that:

$$E[u_i \mid d_i = 1] - E[u_i \mid d_i = 0]$$

Therefore, the **identification assumption** in the randomized experiments is that, conditional on randomization, the treated and untreated individuals are similar in all respects except treatment.

Example 4. To illustrate how the randomized experiments work, consider the data from Example 3. We "know" that the true causal effect is 5, but the observed difference is 20. The reason for the discrepancy is that more able workers are more likely to work for the piece rate firms than less able workers. But what would happen if we were able to randomly assign individuals to the type of firm? For example, assume that we flip a coin and then assign the worker to the piece rate firm if the coin turns head and to the salary firm otherwise. In this case, each worker has an equal probability to work for either firm. The expected productivity of workers assigned to the piece rate firms is then

$$0.5\times25+ 0.5\times10=17.5$$

The expected productivity of workers assigned to the salary firms is then

$$0.5\times20+ 0.5\times5=12.5$$

The difference in the observed productivity between the piece rate and salaried workers would then be 5 (17.5-12.5), that is, the true causal effect of switching to the piece rate pay.

The randomized experiments are widely used in many natural sciences, such as medicine, pharmacology, and chemistry. They are also sometimes used in economics, as the following example illustrates.

Example 5. Bruce Shearer (2004) studied the impact of piece rate pay on productivity of tree planters in British Columbia using experimental data. Each block of land was divided into two parts, one to be planted under piece rates and the other under fixed wages. Each planter was randomly assigned to a block of land within his compensation region. The number of trees planted by workers paid by piece rates was then compared to the number of trees planted by fixed wage workers.

The main limitation of the randomized experiments is that they may be difficult to implement, due to their high costs, ethical consideration, etc. For example, suppose you were interested in knowing whether smoking causes lung cancer. Because of possible health risks, it would be unethical to conduct an experiment in which participants are randomly assigned to the treatment group of smokers and the control groups of non-smokers.