GEQ1000 Asking Questions
Economics (Social Science) Segment
Video 1.2
Causal inference

One of the most important things social scientists do is to ask and answer questions about causation. Do changes in one variable cause changes in another variable?

For example, suppose the average class in schools has twenty-five students, and a researcher wants to find out if reducing the class size to fifteen improves student test scores. Now reducing class size is one of the most expensive ways of increasing education spending, because it involves building more schools and hiring more teachers. Unless the effect of making classes smaller on test scores is powerful, governments may choose to devote limited public funds to other objectives. Answers to this question and others like it can have big consequences for how governments set policies.

Social scientists have adopted the techniques and even the language of medical researchers in their evaluations of policy effectiveness. We call the policy intervention a “treatment”, and we want to measure the **treatment effect** on our outcome of interest. In our class size example, the subjects or units of analysis would be the students. The treatment variable is class size, and the outcome variable is (sic) the students’ test score.

How do we go about establishing and measuring treatment effects?

Observational studies

If there are a variety of class sizes in existing schools, we could make observations about how class size is **associated** or **correlated** with test scores. A positive correlation means that larger class sizes tends to go with higher test scores. A negative correlation means that higher class sizes tend to go with lower test scores. If reducing class size improves test scores, we would expect to see a negative correlation. You can plot test scores against class size to visualize the correlation.

But **correlation does not equal causation**. There are three possible explanations for the correlation, and our observational study cannot distinguish between them, at least without further investigation.

The first explanation is that indeed reducing class size causes test scores to rise. That is, there is a treatment effect.

The second explanation is that high test scores reduce class size, so there is **reverse causation**. Perhaps school administrators have decided students who do well in tests will be moved into smaller classes where they can be given advanced curricula. Thus, high test scores are causing small classes!

The third explanation is that there could be another variable that causes test scores to go up, which is also negatively correlated with class size. Perhaps the schools with smaller class sizes are in wealthy neighborhoods and accept students from relatively wealthy families. Thus, students may be doing well on tests because their family background gives them advantages. For example, the schools they go to tend to have more resources and recruit better teachers. Family background and teacher quality are called **confounding variables** or **confounders**, and the negative correlation is picking up the confounder's effect on the outcome.

We could use the statistical technique known as **regression** to help us take into account the effect of confounders. Indeed, regressions form the bulk of empirical studies in economics and other social sciences. However, regression requires that we are able to measure all the confounding variables and incorporate them into the analysis. That is often impossible because not all confounders are observed.

Thus, to establish causation, we will need to go beyond measuring correlation and doing regressions.

The fundamental problem of causal inference

Let's take a generic case where we have a single person who is our unit of analysis. What we would ideally like to do is to make two observations. The first observation is what happens to the person's outcome if he receives the treatment. The second observation is what happens to the person's outcome if he does not receive the treatment. We can then compare the two outcomes to obtain the treatment effect.

The problem is that we cannot see both observations. Once somebody is treated, we see that he's treated, and we see his outcome. We can never see what his outcome *would have been* if he was not treated. We call this unobservable outcome the **counterfactual**.

The fundamental problem of causal inference is the fact that *we cannot observe the counterfactual*. We must thus find substitutes for the counterfactual. This inevitably involves making assumptions that the chosen substitute is just like the counterfactual. So the question we will ask as we go along the following videos is "What assumptions do we make in using methods to *identify* the treatment effect, and are these assumptions credible?" We call such assumptions **identification assumptions** or **identifying assumptions**, and they are essential to any method of establishing treatment effects.

Now counterfactual thinking is also used with qualitative studies such as those found in history or anthropology. We could think of historical questions such as "Would Japan have surrendered if the United States did not use atomic bombs on Hiroshima and Nagasaki?". Historians will argue based on their knowledge of historical records of the time, whereas social scientists will argue based on experimental and statistical techniques. But both are essentially imagining what the counterfactual looks like, without being able to observe it.

In situations involving units other than humans, finding substitutes for the counterfactual can sometimes be quite easy. Take for example your typical washing detergent commercial. You want to find out if Brand X is better than Brand Y at removing stubborn stains. You put a stain on a shirt, wash it with Brand X. At the same time, you put a stain on another shirt of the same make, and wash it with Brand Y in the same type of washing machine. That's an example of a controlled experiment – you are able to keep everything the same in the two cases except for the treatment.

Well, you can find two shirts that are identical, but you cannot find two persons who are identical. Even the so-called identical twins are not truly identical. Thus we will never be able to measure treatment effects at the individual unit level. However, we may still be able to measure treatment effects for groups of people. That's the idea behind randomized trials, which we will talk about in our next video.