GEQ1000 Asking Questions
Economics (Social Science) Segment
Video 1.4
Regression Discontinuity

In the last video we looked at how randomized trials are used to identify treatment effects. If randomized trials are not feasible, and only observational studies can be done, can we still identify treatment effects with high internal validity?

Quasi-experiments

Well, yes, if we are lucky and if we are observant. Sometimes nature provides a situation where we get two groups that are very much the same, except that one group is treated and the other is not. The situation is therefore akin to a randomized trial. Smart and lucky researchers that come across such situations can use what we call quasi-experimental methods to measure treatment effects.

Regression discontinuity design

The first quasi-experimental method we will look at makes use of the fact that many social programs use cutoffs to decide if units will get treated or not. For example, suppose a school runs remedial classes during school holidays. The principal wants to know if remedial classes help to increase the subsequent test scores of the students.

Only students who have done poorly in their tests are assigned to do these remedial classes. This is obviously not a random assignment. Students who are required to go to remedial class are on average going to be different in important ways from students who are not required to go to remedial class. The two groups are likely to have differences in family background, attentiveness, discipline, ability to focus, and other possible ways that matter for test performance. Therefore, we cannot use students who do not go to remedial class as a control group to compare with our treatment group.

We also cannot simply compare the after-remedial test scores of the students against their pre-remedial test scores. Many things will have changed over the course of time between the two tests. Perhaps new teachers are hired, or the

school received new equipment. Such events are confounders that make it difficult to uncover the treatment effect of remedial classes.

How then can a treatment effect be measured? Here's the idea.

Suppose the passing mark is 60. Students with test scores that are *slightly below* 60 are going to be very similar to students with test scores that are *slightly above or equal to* 60. They are likely to pay the same amount of interest in the class, they are likely to study the same amount of time, they probably have equal abilities, and so on. The only difference is treatment status. Those who scored just below the 60-mark cutoff are treated and those who scored at or just above 60 are not treated.

We can therefore argue that we have a solid identification assumption - within a small band around the cutoff, our two groups are identical except for treatment status. We can thus observe the outcomes and uncover the treatment effect. This assumption can be tested by doing a balance check, to verify that average values of observed confounders are equal across the two groups.

We will have to decide how many marks around 60 constitutes "slightly below" and "slightly above" passing. The bigger the band, the less confident we are about our identification assumption. Thus, we will want to look only at students who, for example, score within 3 marks of 60, that is, between 57 and 63 marks.

A real example

This example is a simplified account of an actual study done by economists Brian Jacobs and Lars Lefgren. In the city of Chicago, school leaders had decided to make summer remedial classes mandatory for third and sixth grade students that did poorly on tests. Jacobs and Lefgren found that for third graders, summer school improved reading and math performance two years later by 12 percent, whereas for sixth graders, performance two years later increased by about six percent.

This exploitation of a cutoff to obtain treatment and control groups is usually combined with regression analysis, and is called a **regression discontinuity design**.

Here are the elements of a successful regression discontinuity design:

First, there is an assignment variable that ranks the subjects.

Second, there is a cutoff to decide eligibility for treatment.

Third, people on one side of the cutoff are treated. People on the other side are not treated.

As applied to our remedial class example:

The initial test scores serve as the assignment variable, ranking the students.

The cutoff is the test's passing mark of 60.

The school must ensure that students who fail the test will attend remedial class, even if they don't want to. Also, the school must ensure that those who pass the test will not attend remedial class, even if they want to.

Weaknesses of regression discontinuity design

The ability to measure a treatment effect makes the regression discontinuity design a powerful tool in the many instances where randomized controlled trials are not doable. However, because our treatment and control groups are drawn only from the portion of the students who are near the cutoff, there are some weaknesses as well.

First, the number of units used to estimate the treatment effect is effectively much smaller than the total number of units in the study. In our example, if the number of students who score within 3 marks of 60 is small, the treatment effect may not be precisely estimated. Increasing the band can increase the number of students, but runs the risk of reducing our confidence in the identification assumption. There is thus a trade-off.

Second, the treatment effect identified by regression discontinuity applies only to the students in the band near the 60-mark cutoff. We cannot generalize the result to the other students in the school, much less to other schools. Thus, the external validity of regression discontinuity is quite weak. Once again, we need to do many studies to be sure that remedial school raises test scores in a variety of contexts.

In our next video we will examine another quasi-experimental technique called Difference-in-differences.