# GEQ1000 Asking Questions
# Economics (Social Science) Segment
# Video 1.3
# Randomized Trials

In our previous video, we talked about how controlled experiments are impossible to do in social sciences, because you cannot find two people or two units that are identical in every way, except that you treat one and you don't treat the other.

What we can do is to run a randomized experiment. The social science community prefers to call them **randomized controlled trials** or **randomized trials** in short.

**Description of the method**

Here's how a randomized trial works. We assemble a group of people, and we randomly put some in a treatment group where they will be treated, and put the rest in a control group where they will not be treated. We then compare the outcomes of the treatment group with those of the control group.

Now if we only had two people, and we randomly assign one to the treatment group, leaving the other in the control group, that wouldn't be helpful because the two people are likely to be different in many ways, and the problem of confounders will come in.

But if we have many people in each group, the *average* value of any possible confounding variable in the treatment group is going to be the same as the *average* value of the same variable in the control group.

So what randomization does is that it creates two groups of people that are essentially identical except that one group is treated and the other is not. Then, when we observe the outcome of the control group, it is as if we are observing the counterfactual outcome of the treatment group. We therefore take the average outcomes in the treatment group, subtract the average outcomes of the control group, and obtain the **average treatment effect**.

Thus, the identification assumption in a randomized trial is that the treatment group is identical to the control group except for the treatment. This assumption will only hold if you have a sufficiently large number of people.

**Examples of randomized trials**

Now we will look at a few examples of randomized trials in education research and development economics.

First, the Tennessee Class Size study.

In 1985 the state government of Tennessee in the United States commissioned a randomized trial called Project STAR to test the effects of smaller class size. About 6,500 students in over 300 classes in nearly 80 different schools participated. Students were randomly assigned either to a small class of 13 to 17 students or a regular class of 22 to 25 students, lasting from kindergarten to Grade 3, the equivalent of Primary 3 in our education system. When they reached Grade 4, all students returned to regular class sizes. The effects turned out to be very large. By the time the students reached Grade 5, the small-class students were performing as well in tests as regular-class students that were on average five months older. That is, the small-class students were ahead of regular-class students by five months!

Next, the Perry Preschool Project.

One of the earliest examples of the use of randomized trials for evaluating social policies was started in 1962 in the vicinity of the Perry Elementary School in Michigan. The study aimed to find out if free preschool education for disadvantaged children could lead to better outcomes – test scores in primary school, and so on. 128 children from poor African American families, who had scored below average in IQ tests, were enrolled in the study, and randomly assigned to treatment and control groups. Those in the treatment group took free preschool lessons for five mornings a week. What is remarkable about this study was that not only were the students followed as they continued school, but data was also collected about them when they turned 27 years old, and again when they turned 40 years old. The researchers were therefore able to look at a whole range of different outcomes, such as whether the subjects went to college, how much they earned, whether they went to jail and so on. And the results were astounding.

At age 27, women who had gone to preschool were 26 percent less likely to have children out of wedlock, and had fewer teen pregnancies.

At age 40, men were 46 percent less likely to have served time in prison. Men and women earned substantially higher incomes, and were far less likely to have received government poverty assistance from age 30 to 40. (source: http://evidencebasedprograms.org/1366-2/65-2)

Next, the use of randomized trials in development economics.

Economists working on problems of economic development in poor countries have increasingly used randomized trials to test out policy interventions. Esther Duflo of MIT is one of the intellectual leaders of this movement, and has done randomized trials in many countries.

In Kenya, she conducted a trial which showed that providing a small subsidy just before harvest season could dramatically increase farmers' use of fertilizer by 10 to 20 percent. In another trial to reduce teacher absenteeism in rural India, she showed that a low-cost solution of having teachers take photos with their students increased student test scores and graduation rates.

Duflo's best selling book *"Poor Economics"*, co-written with fellow economist Abhijit Banerjee, argues that randomized trials should be the main tool in evaluating development policy. The two have set up the Abdul Latif Jameel Poverty Lab, or J-PAL for short, that helps to register and fund hundreds of randomized trials in developing countries. Her TED talk "*Social Experiments to fight poverty*" (https://www.ted.com/talks/esther_duflo_social_experiments_to_fight_poverty) is well worth watching.

**Practical problems with randomized trials**

There are many practical problems to be solved when designing a randomized trial.

For one, it can be hard to *ensure* that those who are assigned to the treatment group get treated, and those who are assigned to the control group do not get treated. In the Tennessee Project STAR, some parents whose children were assigned to the control group made appeals to schools on grounds of fairness, and managed to move their children into the smaller classes. This **contamination** of the treatment assignment likely added bias to the measured treatment effect.

Researchers may also be bound by **ethical concerns**. To have some disadvantaged children not get free preschool if funds are available, simply because of scientific inquiry, smacks of gross injustice. However, if funding is limited, randomized assignment becomes an asset instead, because it can be considered a fair way of handing out limited treatment slots, since every applicant gets an equal chance to be chosen. Budget limitations happen often in developing countries, which is perhaps why randomized trials are more commonly used there.

Another problem comes about if some people drop out of the study. It is likely that those who drop out are systematically different from those who do not drop out, and this would create biases in the estimates of the treatment effect.

The biggest practical problem is that for many questions of causal inference, randomized trials are either too expensive, take too long, or are simply infeasible because the policy has already been enacted. Thus, there is a need for alternative methods, perhaps less rigorous, to measure treatment effects. We will look at two methods that attempt in their own ways to create valid control groups to compare with treatment groups. They require stronger identification assumptions than randomized trials, but they can be used in situations where randomized trials cannot.

**Questions concerning Validity**

There are two questions researchers ask of any study that seeks to measure treatment effects. The first question is whether we are confident that the study can successfully identify the treatment effect. This is called **internal validity**, and depends on how solid the identification assumptions are.

The most appealing aspect of a randomized trial is that internal validity is strong if the trial is executed properly. The Perry Preschool Project showed conclusively that outcomes from a group of children that lived near the Perry Elementary School were improved by free preschool education. That's internal validity.

But can we then claim that making preschool free for *all* disadvantaged children would have the same treatment effect? That's the second question, whether we can generalize the results of the study to a wider population and wider context. This is called **external validity**. Unfortunately, randomized trials tend to have low external validity. Perhaps there is something particular about the location or history of the Perry School that results in large treatment effects. The only way we can find out for sure is to do randomized trials in many schools!

We have come to the end of this video. In the next two videos we will look at methods for causal inference that do not have random assignment.