# Problem Set 3

Tao Wu

October 8, 2023
Professor Suleyman Taspinar
ECON 387: Advanced Econometrics

## 1 Question 1

**(20 pts) Define the following terms in your own words.**

(a) An Ideal randomized controlled experiment is the type of design that can effectively find causes and effects in a scientific study. Randomization of assignment helps ensure that the two groups, the treatment group and control group, are directly comparable and achieve zero conditional mean for unbiased results.

(b) Potential outcomes are multiple possible outcomes that exist simultaneously in the hypothetical world. However, these outcomes cannot be all observed at the same time (we, humans, don't live in a multiverse). In an experiment, outcomes are determined by the treatment or non-treatment.

(c) Causal effect is the difference that is attributable to a change in outcome based on treatment or intervention. If one thing would not have happened, had something else not happened, we say there is a causal effect here. We use the counterfactual concept to assess causal effects. How this plays out is through experiment. Observe the outcome of a group that receives treatment and then compare the outcome against the outcome of another group that does not receive treatment. If the difference is meaningful, a causal effect could be established.

(d) Average treatment effect is the **average** of individual_$i$'s treatment effect $Y_i(1) - Y_i(0)$, which can be written as such $E[Y_i(1) - Y_i(0)]$

(e) Internal validity is how well the estimation model can make the changes in the outcome variable solely attributable to manipulation of the independent variable. In other words, if the predicted estimator is free of bias and precise, we say the model is internally valid. External validity is about the replicability, inferences, and conclusions of a model. If we conduct the same experiment on another population and get the same result, we say the experiment is externally valid.

(f) Threats to internal validity are as follows:

i. Failure to randomize can result in selection bias. We do not want any unobserved characteristics of the participants to play a role in the changes in the outcome variable. Randomization can ensure differences in outcomes are solely due to treatment or intervention.

ii. Failure to follow the treatment protocol can skew the treatment effect. Full compliance is necessary for casual effect assessment.

iii. Attrition means participants drop out in the middle of a study where data is still being collected. This will produce bias and threaten the internal validity of an experiment.

iv. Experimental effects mean participants' behavior changes because of their awareness of being observed.

v. Small sample sizes are more vulnerable to random sample variability. Small sample sizes are less likely to be representative of the population and thus less powerful in explaining the results.

(g) Threats to external validity

i. Nonrepresentative sample means that one of the goals of every experiment is to generalize the results to that broader population. However, if the sample that an experiment is based on is not representative of the population, to begin with, the experiment's findings are hard to be generalized to the larger population.

ii. Nonrepresentative program or policy refers to a treatment or intervention in an experiment that does not accurately reflect the real-world program or policy to which the researchers want to apply the findings. If we want to find out whether painkillers can help patients alleviate pain but the treatment in an experiment is cold medicine. Even if the results are accurate, they won't be applicable in the right context.

iii. General equilibrium effects means the changes in the broader economic, social, or environmental context can affect the outcomes of an intervention in ways that are not captured by the experiment.

# 2 Question 2

I like the fact that there are levels to the treatment because causal effect is the difference that is attributable to changes in potential outcomes based on treatment or intervention. To create the differences, we need to manipulate independent variables. What I don't like is that there is an arbitrary order in assigning different lands to different levels of treatment, which could represent a threat to the internal validity of the experiment. First, the human assigner might have a bias in the assignment process. Second, each assignment of 25 one-acre lands might correlate with different unobserved factors such as weather and

bugs issues. Moreover, not all lands are created equal. Some lands are just more fertile or poor than others due to randomness. One major suggestion to make here is to introduce randomization of the assignment of lands in this study. Randomization minimizes pre-existing differences between 4 groups of 25 one-acre lands. We want the "participants" (lands in this case) to be equal in their characteristics and to avoid selection bias by a human assigner as much as we possibly can. Randomization helps in that regard.

# 3 Question 3

The group that receives a placebo is made up of people who are similar to the people in the treatment group in all aspects. So all else being equal, the only difference between 2 groups is one that gets active treatment and the other one that receives a placebo. For the groups, the difference in means of 2 groups' effect is sometimes called the differences estimator and the difference-in-means estimator. On the individual level, the average treatment effect is the **average** of individual_$i$'s treatment effect $Y_i(1) - Y_i(0)$, which can be written as such $E[Y_i(1) - Y_i(0)]$. If the study begins by having individuals randomly assigned to the treatment and placebo (control) groups, then $X_i$ is distributed independently of all personal attributes and in particular is independent of $Y_i(1)$ and $Y_i(0)$. This way, we can achieve zero conditional mean and know that the OLS estimator $\beta_1$ will be unbiased. To estimate the causal effect i.e. $\beta_1$, let's use a difference-in-means estimator which essentially calculates the difference in mean outcome in the treatment vs control group.

If researchers had data on the weight, age, and sex of each patient in the experiment, for the improvement of the model's estimate, we could introduce these factors, as control variables, into the model. Although we are probably already getting an unbiased $\hat{\beta}_1$, with the help of control variables, the overall fit of the model will improve by having a smaller standard error and reduced error variance. Lastly, if the randomization of the assignment happens to depend on one of the control variables, then not accounting for them will lead to omitted bias. All in all, it's always a good idea to include control variables for better estimation.

# 4 Question 4

**(20 pts) Suppose that, in a randomized controlled experiment of the effect of an SAT preparatory course on SAT scores, the following results are reported:**

(a) Average treatment effect is the **average** of individual_$i$'s treatment effect which can be written as such $\beta_1 = E[Y_i(1) - Y_i(0)]$. To find $\beta_1$, we first need to reasonably assume that the treatment effect is constant for all students, so we can use $\beta_1$ instead of $\beta_i$. Then by linearity expectation, we can re-write the ATE equation as $\beta_1 =$

$E[Y_i(1)] - E[Y_i(0)]$ and plug the given respective average SAT scores of the treatment group and control group into it. So the average treatment effect (ATE) on test scores is 40.

(b) Although we have a nonzero average treatment effect, we don't yet know if this number is significant enough to make a difference or its magnitude in the distribution of SAT scores (dependent variable). So we need to formulate a null hypothesis testing on this ATE. Because two sample groups are being compared here, we will use a two-tailed t-test in this case.

$$t = \frac{\overline{X}_{treatment} - \overline{X}_{control}}{\sqrt{\frac{s^2}{n_1} + \frac{s^2}{n_2}}} \sim t_{n-1}$$

$H_0$ is that $ATE = 0$. $H_1$ is that $ATE \neq 0$. Let's choose the significance level to be 5%, the common threshold of probability. After comparing the calculated t-statistic of 2.971974 to the critical values of $\pm 1.9842$ at 5%, we reject the null hypothesis. In other words, the Average treatment effect of 40 is statistically significant and an SAT preparatory course has a positive impact on students' SAT scores.

# 5 Question 5

**(20 pts) Consider a study to evaluate the effect of dorm room Internet connections on college student grades. In a large dorm, half the rooms are randomly wired for high-speed Internet connections (the treatment group), and final course grades are collected for all residents. Which of the following pose threats to internal validity, and why?**

(a) scenario a) poses an "attrition" type of threat to the internal validity of the study because participants drop out in the middle of a study where data is still being collected. This will produce bias and threaten the internal validity of an experiment.

(b) scenario b) poses a "failure to follow protocol" type of threat to the internal validity of the study because full compliance by both groups is necessary for casual effect assessment.

(c) scenario c) poses another "failure to follow protocol" type of threat to the internal validity of the study because some students don't do what the treatment guides them, which can skew the treatment effect data leading to an imprecise result.

(d) scenario d) poses an "experimental effect" type of threat to the internal validity of the study because students in the treatment group, knowing being observed, take advantage of those in the control group. Morality aside, both the control group effect will be skewed.

# 6 Question 6

(a) "A randomized controlled experiment is an experiment where you control to account for the factors you know about and then randomize to account for those you don't."," explains Redman. The article talks about the design of an experiment and its purpose. It basically says that the world in which we live is a complex system in which many things happen at the same time. The relationship in which one thing influences the others can be very difficult to correctly identify. A randomized controlled experiment is a powerful and reliable tool that helps us achieve that goal and learn about casual effects between things. Running an experiment is a tradeoff. So is everything in life. The "controlling" aspect of an experiment comes at a price that researchers have to balance before a sufficiently good result can be obtained.

(b) The key aspects of an experiment to keep in mind are randomization and control from which a couple of basic steps in conducting a randomized controlled experiment are derived.

i. Decide what your dependent variable of interest is

ii. Determine what the population of interest is

iii. Ask yourself about what question the experiment is trying to answer

iv. Think through all of the factors that could spoil your experiment

v. Write up a research protocol

vi. Run a pre-trial small-scale experiment to test if the protocol will work

vii. Revise the protocol based on what you learned in your pilot study

viii. Conduct the experiment

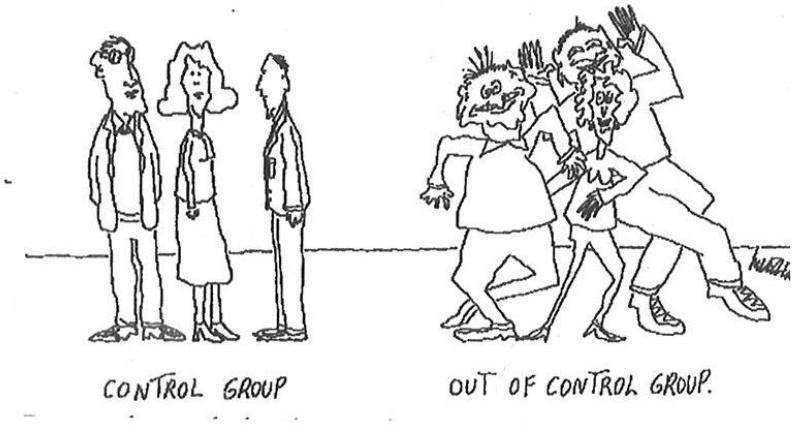
ix. Analyze the results. After that, as Redman says, "You don't make money in a lab. You make money in the real world. So move out of the lab quickly."

(c) According to the article, mistakes people make when doing randomized controlled experiments are

i. Managers or people in charge have not done enough experiments to verify their hunches about answers.

ii. Once they do, they haven't put enough controls in place to isolate the variable(s) they are interested in.

iii. Not involving data analyst who may have more domain knowledge than the data scientist who is the expert in setting up an experiment. In other words, collaborating with data analysts from the get-go can yield great benefits.


CONTROL GROUP
OUT OF CONTROL GROUP.