POLS 201: Data Analysis and Politics

Professor Elena Llaudet

Lecture 15 | Internal vs. External Validity

Plan for Today

- Causal Studies
- Internal Validity vs. External Validity
- Randomized Experiments vs. Observational Studies
- The Role of Randomization
- Let's Evaluate Causal Studies We Have Seen

Causal Studies

- So far we have learned how to estimate the average change in the outcome caused by the treatment
  - with experimental data: by computing the difference-in-means estimator directly or by fitting a simple linear regression model where X is the treatment variable (chapter 2 + chapter 5)
  - with observational data: by controlling for all confounding variables using a multiple linear regression model (chapter 5)
- There are more issues we must consider when conducting or evaluating a scientific causal study, including the internal and external validity of the study

Internal Validity

- Refers to the extent to which the causal assumptions are satisfied
- It asks, **is the estimated causal effect valid for the sample of observations in the study?**
- The answer depends on whether treatment and control groups are comparable, that is, on whether we have:
  (a) eliminated all confounding variables by running a randomized experiment OR
  (b) controlled for all potential confounding variables when using observational data

External Validity

- Refers to the extent to which the conclusions can be generalized
- It asks, **is the estimated causal effect valid beyond this particular study?**
- The answer depends on:
  (i) whether the sample of observations in the study is representative of the population to which we want to generalize the results AND
  (ii) whether the treatment used in the study is representative of the treatment for which we want to generalize the results

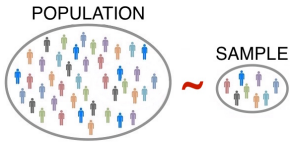
- **Randomized experiments** tend to have strong internal validity but relatively weak external validity
  - random treatment assignment eliminates all potential confounding variables, BUT
  - sample of participants might not be representative of population and/or treatment might be realistic and not comparable to real-world treatments
- **Observational studies** tend to have strong external validity but relatively weak internal validity
  - sample is usually representative of the population and treatment is usually realistic, BUT
  - possibility of uncontrolled confounding variables can't be ruled out

- This dynamic explains why scholars use both types of studies to estimate causal effects; they often have complementary strengths
- Nonetheless, some studies based on experimental data have strong external validity and some studies based on observational data have strong internal validity
- **We should pay attention to the study details when evaluating them**

Randomization Gives Us Super Powers

The Role of Randomization

1. When selecting observations from the population into the sample, **random sampling**
   - ensures sample is **representative** of target population
   - ensures strong **external validity** (assuming the treatment is realistic)



The Role of Randomization

2. When deciding who receives the treatment and who doesn't, **random treatment assignment**
   - eliminates **confounders**, making treatment and control groups **comparable**
   - ensures strong **internal validity**

Randomization Gives Us Super Powers

- Let's evaluate the causal studies we have seen thus far!

Does Social Pressure Affect Turnout?

(Based on Alan S. Gerber, Donald P. Green, and Christopher W. Larimer. 2008. "Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment." *American Political Science Review*, 102 (1): 33-48.)

- Experiment conducted in Michigan, where postcards with get-out-to-vote messages were randomly sent to registered voters
- Internal validity: Strong
  - Why? This is a randomized experiment. Random treatment assignment should have eliminated all confounding variables. Registered voters who received the postcard should be comparable to registered voters who did not.
- External validity: Depends on population and treatment we want to generalize the results to
  - Could we generalize the results to receiving the same type of postcard in the whole of Michigan? Yes if the sample of registered voters used is representative of all registered voters in Michigan.
  - Could we generalize the results to receiving the same type of postcard in Massachusetts? Depends on how different registered voters in Massachusetts are from those in Michigan.

Do Women Promote Different Policies than Men?

(Based on Raghabendra Chattopadhyay and Esther Duflo. 2004. "Women as Policy Makers: Evidence from a Randomized Policy Experiment in India." *Econometrica*, 72 (5): 1409–43.)

- Experiment conducted in India, where female politicians were randomly assigned to rural villages
- Internal validity: Strong
  - Why? This is a randomized experiment. Random treatment assignment should have eliminated all confounding variables. Villages that were assigned to have a female politician should be comparable to villages that were not.
- External validity: Depends on population and treatment we want to generalize the results to
  - Could we generalize the results to having a female politician in the whole of India? Not really, probably only to rural areas (i.e., the population the sample of rural villages is representative of).
  - Could we generalize the results to having a female politician in the U.S. towns? Absolutely not. Rural villages in India are not representative of U.S. towns.

Does the Death of the Leader Increase the Level of Democracy?

(Based on Benjamin F. Jones and Benjamin A. Olken. 2009. "Hit or Miss? The Effect of Assassinations on Institutions and War." *American Economic Journal: Macroeconomics*, 1 (2): 55-87.)

- Observational data from assassination attempts of leaders around the world
- Internal validity of study without controls: Weak
  - Why? This is NOT a randomized experiment. We might worry about potential confounding variables such as *politybefore*. Countries where the leader died were more democratic to begin with than countries where the leader did not die.
- Internal validity of study with controls: Stronger
  - Why? Controlling for *politybefore* should help make treatment and control groups more comparable.

- External validity: Depends on population and treatment we want to generalize the results to
  - Could we generalize the results to the death of the leader in all countries? Probably not. We should probably only generalize to countries with assassination attempts (which tend to be less democratic to begin with).

Today's Class

- Internal vs. External Validity

Next Class

- Review in Preparation for the Midterm
- **No computers needed**