Political Science 210: Introduction to Empirical Methods

Week 6: Small N

Advantages of large-N studies

Over the last two weeks, we've talked about how we can study large sets of cases in order to describe populations:

- Experiments
  - Great for determining causal relationships
  - Lets us compare averages between control and treatment groups
- Regression
  - A tool for studying covariance (when X changes, does Y change along with it?)
  - Can control for potential confounders
  - Gives us covariance between X and Y that best describes our set of cases

Limits of large-N and experimental studies

When we learned about experiments in Week 4, we asked about the causal effect of an extra hour of study on your grade:

Limits of large-N and experimental studies

…so we took a group of students and randomly assigned some of them to study an extra hour than they normally do:

Extra hour of study

Avg. final grade

Control group (no extra study)

Avg. final grade

…leading us to conclude that an extra hour of study increased our performance by half a letter grade, *on average*.

Limits of large-N and experimental studies

But does that answer the original question? Would *your* grade improve by half a letter?

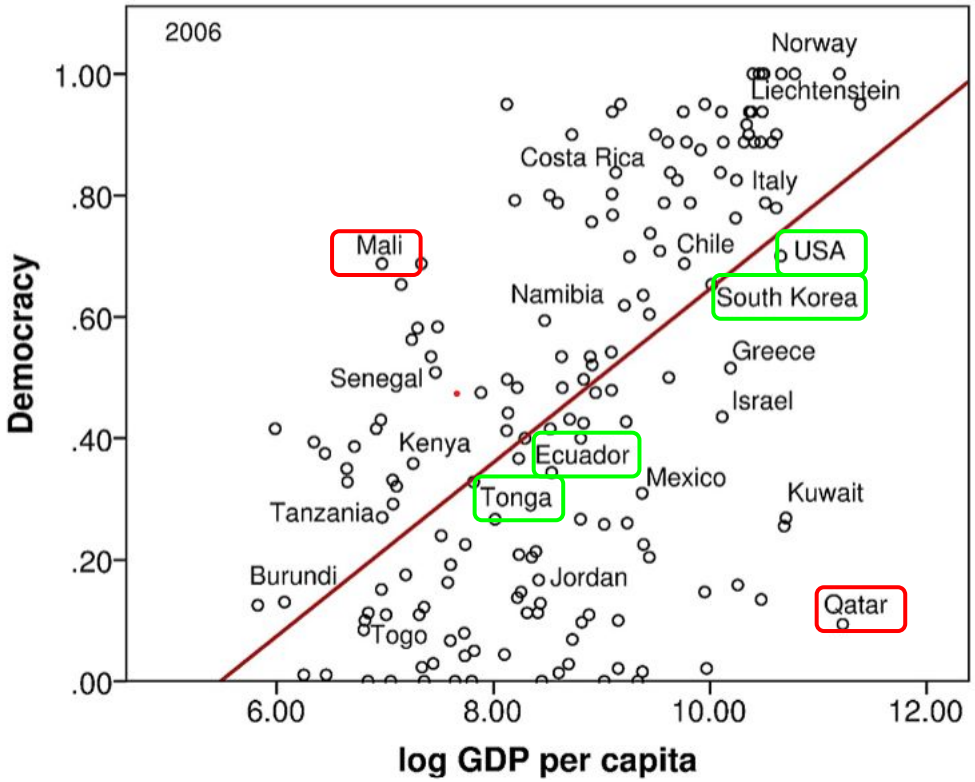
?

We don’t know! You might be different from the population average in many ways:

- Better (or worse) study habits
- Stronger (or weaker) prior knowledge of the material
- A weaker midterm grade (more room for improvement? Lower ceiling on final?)

Most of the population’s grade increase might differ from the average, if the population is heterogenous enough.

Limits of large-N and experimental studies



…and last week, we fit a regression line to find a relationship between X and Y that best explained a large set of cases.

But this didn't tell us *why* X and Y were related, even for cases that fit the line best…

…and we have even less idea why some cases diverged from the rest.

Advantages of small N / qualitative studies

- Can study complex processes put forward by a theory
  - X-Y relationship relies on complicated processes with different mediating mechanisms
    - Not just X -> Y, but X -> A -> B -> C -> Y
  - Even if a large-N study supports our theory, we may want to confirm a process is taking place
- Can probe why existing theories can't explain a given case
  - X and Y don't behave in the way we would expect given observations of other cases
  - If existing theory doesn't hold, does the outlier case
    - Discredit a critical piece of existing theory?
    - Provide scope conditions where the theory doesn't hold?
    - Allow us to test new theories about the X-Y relationship?
- But be careful of…
  - Whether the case is generalizable to a population of interest
    - (although sometimes cases might be of particular interest to us)
  - Whether our generated or tested theory is falsifiable (how would we know if we were wrong?)

Selecting cases

We might select as little as one single case to study the relationship between X and Y within that case.

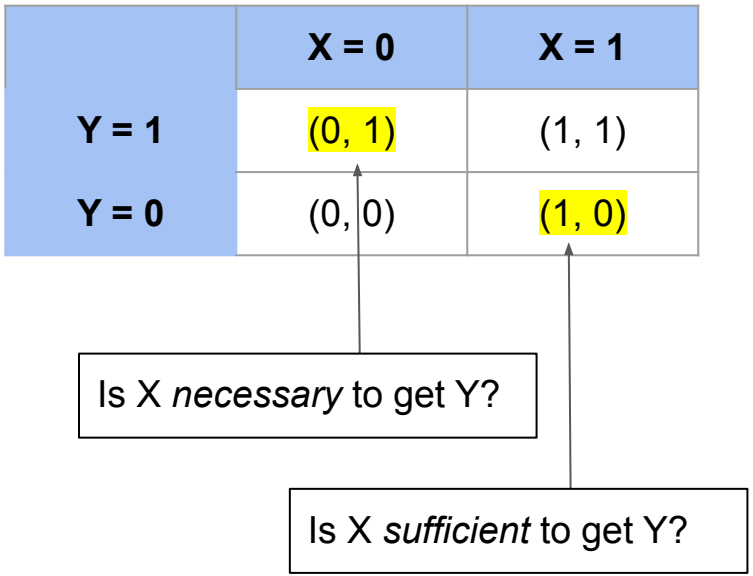
- Typical case selection
  - Where the case is thought to be representative of our overall population
  - Used to confirm our theory about X and Y
  - Useful to explore complex causal mechanisms in greater detail
  - *Why* does X cause Y?

	X = 0	**X = 1**
Y = 1	(0, 1)	(1, 1)
Y = 0	(0, 0)	(1, 0)

Selecting cases

We might select as little as one single case to study the relationship between X and Y within that case.

- Deviant case selection
  - Where the case behaves differently from the rest of the population
  - Used to either confirm existing theory or explore new theories
  - When X = 0, Y = 1:
    - Do we *need* X to get Y?
  - When X = 1, Y = 0:
    - Is X *enough* to get Y?



Cross-case selection

We might also compare two or more cases to better isolate which variable(s) X lead to Y.

- Most-similar design
  - Both cases are as similar to each other as possible on all independent variables, but not the outcome.
  - If cases are very similar on different possible X variables, can better isolate which X factors were *necessary* in leading to Y while “controlling” for other variables.

	Case 1	**Case 2**
X values	Mostly 1	Mostly 1
Y	1	0

Cross-case selection

We might also compare two or more cases to better isolate which variable(s) X lead to Y.

- Most-different design
  - Cases are as dissimilar to each other as possible, but both cases led to the outcome.
  - Lets us identify which potential X variables were *sufficient* (that is, could cause on their own) in leading to Y.
  - Can eliminate other variables as necessary causes of Y

	Case 1	**Case 2**
X values	Mostly 1	Mostly 0
Y	1	1

Read the following scenarios and discuss (1) the type of case selection used in study (typical, deviant, most similar, most different) and (2) whether you think anything could be done to improve the selection method.

- After a country is hit by a severe recession, most of the incumbent mayors in its major urban centers are voted out of office in the following elections. However, the mayor of one city is re-elected by overwhelming margins. A polling firm interviews voters in the city to find out why the mayor remains popular.

- Beset by political and economic upheaval, an authoritarian regime finds that rebel groups are beginning to emerge in different sections of a country. One group is based in a rural, poverty-stricken mountainous region, while another group has emerged in a relatively prosperous urban center. A team of ethnographers from outside the country gains the trust of the rebel groups and spends time with them to learn about their shared grievances.

- The modernization thesis (that is, that wealth causes democracy) was initially created by observing the histories of European countries. A professor of comparative-historical studies takes a closer look at the history of England, tracking how economic development led to the rise of the merchant classes, which generated demand for fewer restrictions on trade and discontent towards the monarchy's power.

- The Scandinavian countries of Northern Europe tend to spend generously on universal benefits, but another wealthy democracy, the United States, has fewer universal benefits. Researchers put together a set of focus groups in each country and find that Americans are more likely to blame poverty on individual failings.

Small N and causal inference

While we can’t prove causality with small N studies, we can still strengthen our confidence in the causal mechanism or causal pathway behind our theory.

This can be done through *process tracing*, where we study the sequence of events in our causal process.

If our theory is falsifiable, we should be able to look for evidence that supports our theory and discounts rival theories.

Small N and causal inference

Evidence that tests our theory are called *causal process observations*.

		Sufficient to affirm causal inference?	
		No	**Yes**
Necessary to affirm causal inference?	**No**	*Straw-in-the-wind* ● Affirms theory is relevant, but doesn't confirm it. ● Failing this test *slightly weakens* our theory, but doesn't discredit it	*Smoking gun* ● Confirms hypothesis ● Failing this test *slightly weakens* our theory, but doesn't discredit it.
	Yes	*Hoop test* ● Needed to affirm theory is relevant, but doesn't confirm it. ● Failing this test discredits our theory.	*Doubly decisive* ● Affirms our theory and eliminates rival theories ● Failing this test discredits our theory.

Theory: Times of national crisis or conflict produce a “rally around the flag” effect, where public support for the incumbent government increases as citizens rally around their leaders.

Potential causal process observations (pieces of evidence in a qualitative study):

1. When tensions rise between two countries prior to a conflict, media coverage of the tensions increases.
2. Public support of the conflict rises in opinion polling as the conflict escalates.
3. After the government announces a successful operation against the opposing country, citizens in major urban areas take to the streets in impromptu celebration.
4. The governing administration in a country in conflict is more likely to be re-elected than administrations in times of peace.

- Straw-in-the-wind: “Weak” test, neither necessary nor sufficient; slightly affirms theory
- Smoking gun: Sufficient to affirm the theory, but not necessary (can look for other tests)
- Hoop test: Needed to affirm the theory, but not sufficient on its own
- Doubly decisive: Both necessary and sufficient to affirm the theory

Read the following example observations and consider what type of causal observation they represent:

- *Straw-in-the-wind: "Weak" test, neither necessary nor sufficient; slightly affirms theory*
- *Smoking gun: Sufficient to affirm the theory, but not necessary (can look for other tests)*
- *Hoop test: Needed to affirm the theory, but not sufficient on its own*
- *Doubly decisive: Both necessary and sufficient to affirm the theory*

- A researcher studying social desirability bias (that is, the reluctance to share viewpoints that may be unpopular with one's audience) observes a group discussion about a controversial political issue where several participants appear hesitant to share their views.

- "Democratic peace theory" holds that democratic countries are less likely to go to war with one another. A historian is studying a case where two democratic countries held hostile relations toward each other, although the hostilities did not lead to armed conflict. While digging through archival records, the historian finds a government memo where one country's chief executive declines to invade the other country, citing how the move would prove unpopular among voters and noting the ethical implications of attacking another democratically-elected country.

- Is a country with higher levels of natural resources more likely to experience greater instability? A researcher studies one particular country that experienced a sudden boom in natural resource wealth, and observes how much of the profits were kept by corrupt officials while workers involved in resource extraction toiled under substandard labor conditions.