PS 0700
Nuts and Bolts of Political 'Science': Causality

Political Science Research Methods

Professor Steven Finkel

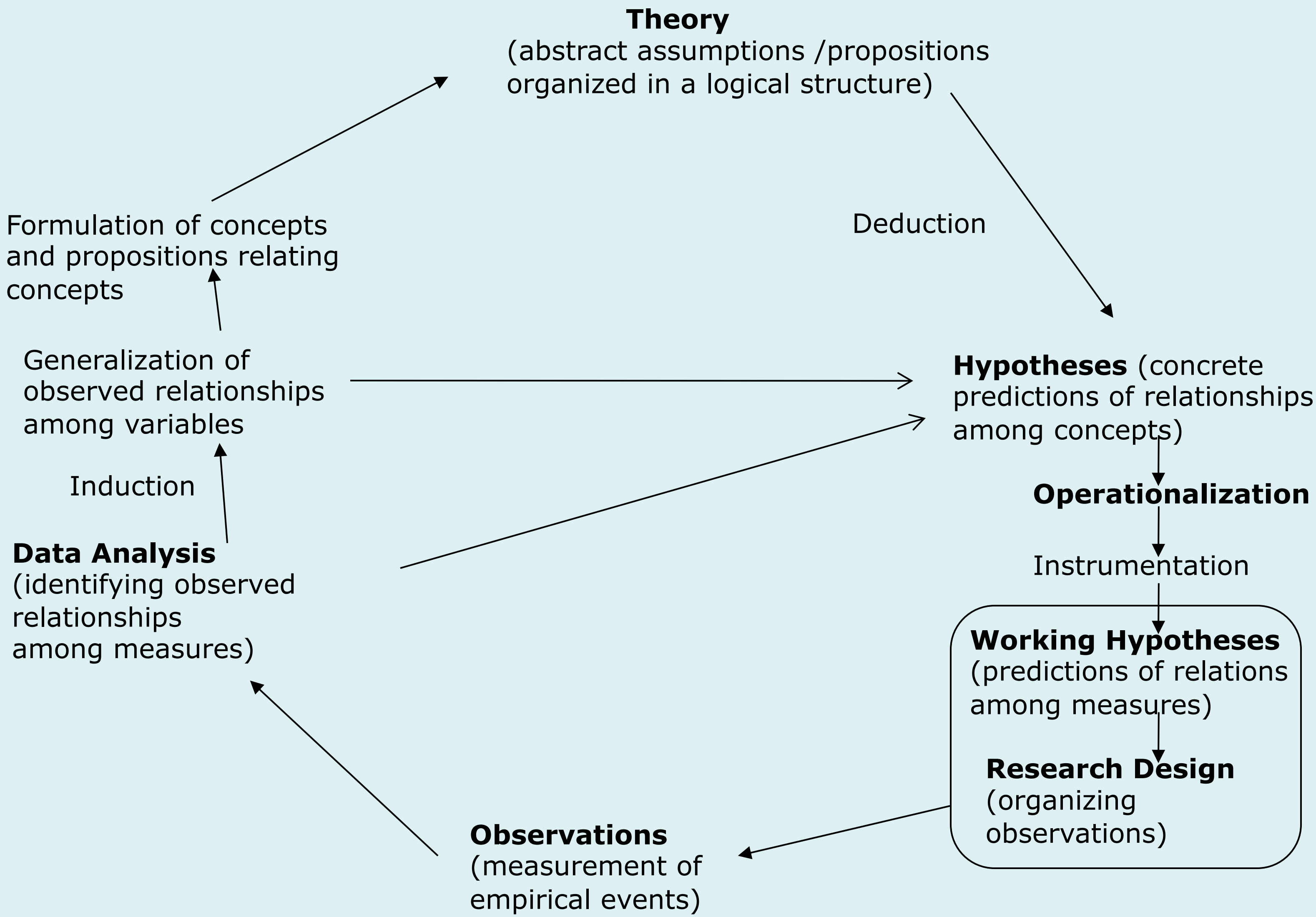
Fall Semester 2022

Week 3b

Goals for the Session

- Discuss how concept of causality is used in the social sciences
- Discuss criteria used to establish causal relationships in political science

A Model of the Research Process



What is a “Causal Explanation”?

- Recall that our goal in social science is to explain observable phenomena by appealing to general *empirical regularities* of the form:
  - If X (independent variable), then Y (dependent variable)
    - If financial conditions increase or stay the same since last election, individuals vote for incumbent party candidate for President; if financial conditions decrease, individuals vote for opposition party candidate
    - If countries have high levels of social trust, they will have more democratic systems; if countries have low levels of social trust, they will have less democratic political systems
  - We want to use these regularities as *causal* explanations, such that we can explain an individual’s vote (Y) by appealing to that individual’s change in financial situation (X); we want to explain a given country’s political system (Y) by appealing to its level of social trust (X)
- Under what conditions can we make a valid claim that X “causes” Y? This is one of the most important – if not the most important -- issue faced by political scientists (if not the most)

The "Necessary and Sufficient" Approach to Causality

- The strictest notion of causality – or perhaps the "ideal" notion of causality – is one where X (independent variable) is both a *necessary* and a *sufficient* cause of Y (dependent variable)
- What is a "necessary" cause of something?
  - If Y, then there *must be* X
  - In order to obtain some outcome on the dependent variable, there *has to be* a particular value on the independent variable
  - *Without* that value of X, there will *never be* the particular outcome on Y
- **"Necessary" does not mean that there will *always* be Y whenever there is X, it only means that *whenever there is Y*, there will also be *X***

- Example: Having a lottery ticket is a necessary condition for winning that lottery.
  - You cannot win the lottery (Y) without having a ticket (X)
  - But there are many (virtually all) people with tickets who don't win
- *Possible* Additional Examples?
  - To get lung cancer (Y), you *have* to smoke (X): ???
  - To get an opposition party candidate elected for President (Y), you *have* to have economic distress (X) : ???
  - To get a more democratic political system (Y), you *have* to have high levels of social trust (X) : ???
- **None of these examples fulfill the "necessary" condition for causality. (Why?)**

- What is a "sufficient" cause of something?
  - If X has a certain value, then Y *will be* a certain value
  - X brings about Y no matter what

- Example: Pushing a working accelerator is a "sufficient cause" for making a car go faster. If X (PUSH ACCELERATOR) equals "YES,", then Y (CAR SPEED)="INCREASE"

- **"Sufficient" does not mean that there are not *other* ways to get Y aside from X; it only means that *if* you have X at a certain value, you *will* get Y.**
  - There are other ways to make a car go faster (e.g., putting it in neutral and rolling down a hill).

- *Possible* Political Science Examples?
  - If a country has high inequality (X), it will have a revolution (Y)
  - If a state has a corporate income tax (X), it will have high unemployment (Y)
- Sufficiency" means that *whenever* you have high inequality or a corporate income tax (X), you will see a revolution or high unemployment (Y).
- **But none of these examples fulfills the "sufficient" condition either! (Why?)**

- The most powerful kind of causal statement would be one that said that X is a *necessary and sufficient* cause of Y. That is:
  - To get Y, you have to have X (necessary)
  - Whenever you have X, you also get Y (sufficient)
  - So X *always* leads to Y, and Y *always* implies X
- Increasing the temperature to 212^0 Fahrenheit (X) boils water (Y). Whenever you have X, you also have Y; whenever you have Y, there had to be X
- Can you think of *any* political science examples of a "necessary and sufficient" causal relationship**? (I can't!!!)**
- This definition is too strict for what we can accomplish in social or political science. It is even too strict for what is usually accomplished in the medical and natural sciences (e.g., you can get lung cancer in other ways besides smoking, and not all smokers get lung cancer)

- Why does the necessary and sufficient model fail for political/social science?
  - Social/political phenomena can stem from multiple potential causes, i.e., multiple independent variables, so we need to recognize that we will rarely have a *necessary* condition for some outcome. There may be several different routes to some outcome, or some combination of factors that lead to a particular outcome
    - Opposition party victories may result from economic distress *or* from scandal involving the incumbent candidate *or* from an unsuccessful war
    - Liberal democratic systems may result from high social trust *or* from advanced economic development *or* from elite bargaining
  - We cannot explain *every* individual or *every* country's outcomes; social and political life is too complex, and to some degree, intrinsically unpredictable. So we will rarely have "sufficient" conditions for Y such that X *always* brings it about
    - When there is inequality in a country, it is *likely* that there will be a revolution, but it won't *always* happen
    - When there is a corporate income tax in a country, it is *likely* that there will be high unemployment, but it won't *always* happen

- So we need to accept that causal explanations in social and political science are going to be:
  - "Probabalistic", not "Deterministic". We will say that having X increases the likelihood of Y, or that X *tends* to cause Y, not that it *always* does
  - **Social science causality stems from a *probabalistic* version of the sufficiency condition:**

    If X occurs, then Y is *likely* to occur; if X does not occur, then Y is *unlikely* to occur
    - If high trust, then democracy is *likely* ; if low trust, then democracy is *unlikely*
    - If economic distress, then opposition party victory is *likely*; if no economic distress, then opposition party victory is *unlikely*

Establishing Causality in Social Science

- Even with this looser *probabilistic* conception of causality, however, it is still *very* difficult in the social sciences to establish that there is a causal relationship between X and Y
- At least three specific criteria need to be met to be able to say with confidence that "X causes Y", even probabilistically
- *Research Design* is crucial for setting up the research and data collection process so that we can satisfy these criteria and make causal claims with the highest possible degree of confidence

Three Criteria for Establishing Causality in the Social Sciences

1. **"Covariation"**: There is an association between X and Y

   – Most basic criterion for causality: there must be some kind of covariation between X and Y to begin with. If you see no association between X and Y, there is almost certainly no causal connection.

   – This criterion is important because it may be the case that you take a *sample* of observations and see that X is related to Y, but there is not a true relationship between X and Y in the overall *population*. We will discuss this more fully in later lectures

   – But showing covariation between X and Y does not by itself establish causality ("correlation does not equal causality!!") (Or, covariation is a "necessary" but not "sufficient" condition for establishing causality!!!)

A Note on Non-Meaningful or Chance Relationships

- With a small number of observations, you may find some relationships that covary, yet are just due to random chance. For example, when the Washington Commanders (formerly Redskins) win their last home game before a Presidential election, the incumbent party tends to win, while a Commander loss means the opposition party tends to win. This prediction has held for 18 out of the last 19 election years. But there is no credible reason for this relationship to hold. It is (almost certainly) the product of chance patterns in the universe!!!

- This is one reason why textbooks often list another criteria for causality as "there is a theoretical reason for the effect" or "there are causal mechanisms to explain the effect". We need a reason to think that a causal relationship exists, not just chance patterns.

2. **"Time Precedence"**: X Comes Before Y

- For X to "cause" Y, it must be the case that X occurs first, and then Y comes later
- We live in a universe where this is true – backwards causation in time isn't possible (outside of the movies or the TV show "LOST")
- In some instances, we can be pretty sure that X comes before Y in time (e.g. exposure to civic education classes in elementary school and later voter turnout), but in many instances in social science we are not so sure which comes first
- For example: does belonging to voluntary associations lead to higher levels of social trust, or do higher levels of social trust lead to joining more voluntary associations? Does U.S. democracy assistance lead to higher levels of democracy, or does higher levels of democracy make the U.S. more likely to provide assistance?
- WE ALWAYS NEED TO CONSIDER THE POSSIBILITY THAT Y→X, AND DESIGN STUDIES TO RULE THIS OUT!!!

3. **"Non-Spuriousness"**: There is not another variable(s) "Z" that *confounds* the causal process and renders the relationship between X and Y *spurious*. In that case, the observed relationship between X and Y results completely from their joint association with Z, not because X → Y

X (Number of Storks in Area)

↓

Y (Birthrate of Area)

- X and Y covary, but is this covariation "causal"? Probably not!!! It is probably due to some confounding variable that is related to both X and Y. What could this be?

Z (Rural/Urban Area) → X (Number of Storks in Area)

Z (Rural/Urban Area) → Y (Birthrate of Area)

- Rural areas tend to have more storks than urban areas, and rural areas tend to have higher birthrates than urban areas (regardless of how many storks they have). If you don't consider Z in the model, it appears that storks "cause" births. Once you "control" for Z, you realize that this is an *illusory* causal claim, and in fact the original X—Y relationship was "*spurious" due to the joint relationship with Z*
- Many observed relationships are spurious – the trick is to find the Zs!!!
  - Children's Shoe Size (X) "causes" Cognitive Ability (Y). Z?
  - Ice cream consumption (X) "causes" Drowning Deaths (Y). Z?

- Political Science Example

X (Attending an Adult Civic Education Workshop)

$\downarrow$

Y (Political Tolerance)

- We observe covariation between X and Y – people who attend civic education workshops in developing democracies are more tolerant than non-attendees. But the relationship may be spurious due to: ???

- Political Science Example

$$\begin{tikzcd} \text{Z (Education)} \arrow[r] \arrow[dr] & \text{X (Attending Civic Education Workshop)} \\ & \text{Y (Political Tolerance)} \end{tikzcd}$$

- We observe covariation between X and Y – people who attend workshops are more politically tolerant. But the relationship may be spurious due to education: more educated individuals tend to attend the workshops (Z→X), and more educated people tend to be more tolerant, *regardless of whether or not they attend the workshops* (Z→Y).
- If we would "control" for Z, we would see that the X→Y relationship was *spurious.*
- "Controlling" means examining the X→Y relationship *within* every category of Z, as we will see in our statistics sessions later

- Thus, we need to be sure that any observed relationship between X and Y is not due to some outside Z variable's influence on the process. This is the most important criteria for causality, and the one that is most difficult to rule out in actual research!!!
  - Does social trust (X) **cause** countries to have more democracy (Y), or do richer countries (Z) have higher levels of trust and higher levels of higher democracy, regardless of X? (If so, wealth is what matters, not trust)
  - Does having a proportional representation electoral system in a country(X), compared with a single member district electoral system, **cause** the country to have more political parties (Y)? Or do more socially fragmented countries (Z) tend to adopt proportional representation and also have more parties, regardless of X? (If so, social fragmentation is what matters, not the type of electoral system)
- **WE ALWAYS NEED TO CONSIDER THE POSSIBILITY OF SPURIOUS RELATIONSHIPS DUE TO Z VARIABLES, AND DESIGN STUDIES TO RULE THEM OUT!!!!!**

Causality and "Counterfactual" Logic

- Final approach to causality rooted in "counterfactual" logic
  - Think about independent variables as "treatments" and dependent variables as "outcomes" of those treatments.
  - Virtually all empirical political analysis can in principle be viewed as an attempt to estimate the causal effects of some kind of "treatment" on a particular outcome or set of outcomes. E.g., the effect of going to college on voting, effect of the effect of changing electoral laws on number of parties, the effects of a state corporate income tax on unemployment, etc.
  - For example, a unit either *receives* or *doesn't receive* a treatment of some kind (e.g., attending a civic education workshop). This is the level of "X".
    - When X=0, the individual doesn't attend the workshop ("control")
    - When X=1, the individual attends the workshop ("treatment")

- Now imagine the *same unit's* level of some outcome (Y) under each of these conditions. That is, for the people who attend the workshop (i.e., were "treated"), we need to imagine what Y would have been *had they not attended*; and for the people who don't attend (i.e., were "not treated" or were "control"), we need to imagine what Y would have been *had they attended* the workshop
- In this framework, the difference in the outcome Y that a unit would obtain under "treatment" (Y_{1i}) versus "no treatment" (Y_{0i}) is the *causal effect of the treatment*
- So: $Y_1 - Y_0$ ***for each unit *i**** is the causal effect we are after; what each unit's difference on the outcome is in the world where they attended the workshop versus the world where they did not attend the workshop
- And we could estimate these quantities across an entire sample or population and calculate the "Average Treatment Effect" ("ATE")
- **This makes good sense, BUT there is a HUGE PROBLEM!!!!**

The “Fundamental Problem of Causal Inference”

- Problem: The quantity we are after is *unobservable*!!! We only see one of the two values of Y for a given unit --- Y_0 for the control group (X=0), and Y_1 for the treatment group (X=1). This is what is known as the **“fundamental problem of causal inference”**!!!
  - We don’t know what the treatment group would have looked like at a given point in time if it had not gotten the treatment (Y_0| X=1), and we don’t know what the control group would have looked like at a given point in time if it had gotten the treatment (Y_1| X=0). **These “counterfactual” outcomes are unobserved, so we cannot directly calculate the causal effect of the treatment for any individual (or the average ATE)**.

	Treatment Group	Control Group
Treated	**Y_1|X=1 Observed**	***Y_1|X=0 Counterfactual***
Untreated	***Y_0|X=1 Counterfactual***	**Y_0|X=0 Observed**

- Nearly all modern empirical social science research is concerned with developing ways of **identifying** and **estimating** the unobservable quantity (Y_1- Y_0).
- Under what conditions can we use differences between *observed* treatment and control groups as proxies for the true *unobserved* causal effects of interest? That is, we need to think about when groups that we *do* observe in the real world, i.e., those who receive and do not receive treatment, can serve as appropriate *counterfactuals* or proxies for one another
- When can this occur? when **all potential Z variables which may produce "baseline" differences between the groups are controlled or accounted for!!!**
- This leads us back to needing to satisfy the "non-spuriousness" criteria we discussed earlier!
- It also leads us to think about strategies for observing or *finding* the right counterfactual group in the first place for a given research design. We'll see lots of different strategies for finding appropriate "counterfactual" control groups over the course of the semester

- **The counterfactual model has become the most prevalent model of causality in the social sciences**
- It subsumes the "3 criteria" model and gives the ability to make more, and more precise, assessments of causal quantities of interest
- Example: in longitudinal analyses there may be Z variables but it is not **as clear** from the "3 criteria" model exactly how the Zs interfere with causal inference (we'll see this in the research design section)
- Example: with the counterfactual model, you can see that there are many kinds of causal effects that the researcher could estimate
  - For example, if you restrict the model to the first column of the matrix – the treatment group – you would be estimating the "Treatment Effect on the Treated" (or ATT")
  - If you restrict the model to the second column – the control group – you would be estimating the "Treatment Effect on the Controls" (or "ATC")
  - These quantities may sometimes be different (and important for substantive and policy reasons). The effects of treatments on units that *are* treated in the real world may be different from what the effect of treatments *would be* on units that don't ordinarily get the treatment in the real world.