PSCI2300 The Study of Politics
Research Design

Tetsuya Matsubayashi
University of North Texas
February 15, 2011
Our goal of scientific research project is to draw sound conclusions that are supported by evidence.

We need to specify a research design to test your hypothesis and offer evidence.

- What is a research design?
- What is a causal inference?
- Why is causal inference important?
- Causal inferences and controlled experiments
- Causal inferences in non-controlled experiments
Why is a Research Design Important?

- Suppose that you plan to examine the relationship between a literacy rate and democracy.
- “As the proportion of a country’s population that is literate increases, the country’s political process becomes more democratic.” ↔ A time-series design.
- “Countries with higher literacy rates tend to be more democratic than countries with lower literacy rates.” ↔ A cross-sectional design.
- Finding an appropriate research design is almost equivalent to deciding what data you collect for testing your hypothesis.
Research Design

- A plan that shows how a researcher intends to test a hypothesis

- Your research design specifies:
  1. What hypothesis will be tested.
  2. What the appropriate “unit of analysis” (e.g., individuals, organizations, states, countries) is for the tests.
  3. What data are needed and how they are collected.
  4. What would be the best analytical or statistical procedure.

- Your research design allows you to make a causal inference.
Causal inference is a key element of theory building in science.

To make a valid causal inference, a research design shows three things:

1. **Covariation**
   - Demonstrate that the cause \((X)\) does in fact covary with the supposed effect \((Y)\)

2. **Temporal Sequence**
   - Show that the cause \((X)\) preceded the effect \((Y)\)
   - \(X\) must come before \(Y\) in time

3. **Elimination of possible alternative causes**
   - Find confounding factors
   - Make sure that there is no spurious relationship
Suppose we study the effect of party mobilization in campaign on voter turnout

<table>
<thead>
<tr>
<th>Y (Voted?)</th>
<th>X (Not Mobilized)</th>
<th>X (Mobilized)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>0%</td>
<td>100%</td>
</tr>
<tr>
<td>No</td>
<td>100%</td>
<td>0%</td>
</tr>
</tbody>
</table>
Causal Relationship

\[ X \rightarrow Y \]

(Party mobilization) \hspace{1cm} (Decision to vote)

Spurious Relationship

\[ Z \rightarrow X \rightarrow Y \]

(Party supporter) \hspace{1cm} (Party mobilization) \hspace{1cm} (Decision to vote)
Party mobilization and Voter Turnout

- Party mobilization covaried with citizens’ decision to vote
- Party mobilization preceded citizens’ decision to vote
- Is this a causal relationship or spurious relationship?
- A **spurious relationship** arises because two things are both affected by some third factor and thus appear to be related. Once this additional factor has been taken into account, the original relationship weakens or disappears altogether.
- Distinguishing causal relations from spurious ones is an important part of scientific research.
Given the possibility of spuriousness, how do we make valid causal inferences?

Control the effect of being a party supporter using information from survey data

Not a party supporter

<table>
<thead>
<tr>
<th>Y (Voted?)</th>
<th>X (Not Mobilized)</th>
<th>X (Mobilized)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>No</td>
<td>100%</td>
<td>0%</td>
</tr>
</tbody>
</table>

A party supporter

<table>
<thead>
<tr>
<th>Y (Voted?)</th>
<th>X (Not Mobilized)</th>
<th>X (Mobilized)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>0%</td>
<td>100%</td>
</tr>
<tr>
<td>No</td>
<td>0%</td>
<td>0%</td>
</tr>
</tbody>
</table>
Experimental and non-experimental designs are used.

In experimental designs, random assignment solves the selection problem.

In non-experimental designs, the selection problem is solved by identifying and controlling the effects of important confounding variables.

The key issue is self-selection.
Randomized Controlled Experiment

Population → Sample → Treatment Group

Random Assignment

Random Assignment

Control group
Randomized Controlled Experiment

- The experimenter establishes a treatment group and a control group.
- Randomization ensures that at the outset the experiment and control groups are virtually identical in all respects.
- The researcher controls the administration or introduction of the experimental treatment.
- The response of interest is assessed before and after the stimulus is given.
- The environment of the experiment is under control.
Controlled Experiment on Negative Ads and Voter Turnout (Ansolabehere et al. 1994)


Groups of participants were exposed to one of the 3 advertisement treatments: positive political advertisements; no political advertisements; negative political advertisements.
Example: Hypothetical Media Experiment

Treatment/Control

Group 1: Positive Political Ad
Intention to vote: no change

Group 2: Non-Political Ad
Intention to vote: no change

Group 3: Negative Political Ad
Intention to vote: decrease
The table below reports the percentage of experimental subjects who intend to vote before and after the exposure.

<table>
<thead>
<tr>
<th>Group</th>
<th>Before Experiment (Pre-Test)</th>
<th>After Experiment (Post-Test)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1: Positive</td>
<td>71%</td>
<td>72%</td>
</tr>
<tr>
<td>(control)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 2: Non-Political</td>
<td>68%</td>
<td>66%</td>
</tr>
<tr>
<td>(control)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 3: Negative</td>
<td>70%</td>
<td>20%</td>
</tr>
<tr>
<td>(treatment)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Do We Make a Valid Causal Inference?

- Covariation?
- Time Order?
- Elimination of Alternative Possible Causes?
Advantages and Disadvantages of Experimental Designs

- High **internal validity**
- Low **external validity**

**Summary:** experimental designs are potentially useful because...

1. they allow researchers to isolate the effects of independent variables by controlling ...
2. the random assignment of subjects to experimental treatments
3. the introduction of the experimental stimulus itself

- Laboratory and field experiments
Because laboratory and field experiments are difficult to carry out (especially when one wants to study aggregates like cities, states, or countries), nonexperimental approaches are more practical. Importantly, although nonexperimental designs can be applied to a wide variety of topics, they do not lead to as strong as causal inference as experiments do.

- Time-series design
- Cross-section design
- Case study design
- Taken all together
Data of the dependent variable both before and after the observation of the independent variable have to be available to us.

Example:
New speed laws → the number of car accidents in a city
• Times series designs work best when the independent variable occurs at a particular moment or during a fairly brief period of time.

• It is important to confirm that no other major changes before and after the introduction of the independent variable.

• No control over the introduction of the independent variable.

• Does this design help us meet three conditions of causal inference?
Example: Effect of a White House media campaign on citizens’ attitude
Cross-Sectional Design

- Measurements of the independent and dependent variables are taken at approximately same time
- No control over the introduction of the independent variable.
- Example: New speed laws → the number of car accidents in a city
Speed Laws and Car Accidents in Ten Cities

The Number of Car Accidents per Month vs. Speed Limit in Urban Areas

The graph shows a positive correlation between the speed limit in urban areas and the number of car accidents per month. As the speed limit increases, the number of accidents also tends to increase.
Cross-Sectional Design

- Example: “Those with more formal education earn more income”
- We collect the data (from surveys) about individual’s education and income

<table>
<thead>
<tr>
<th>Group</th>
<th>Income</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment (High Education)</td>
<td>$80,000</td>
</tr>
<tr>
<td>Control (Low Education)</td>
<td>$30,000</td>
</tr>
</tbody>
</table>

- Question: Can we conclude that education determined income?
- Cross-sectional designs improve *external validity* at the expense of *internal validity*. Why?
Case Study Design

- Examines one of a few cases of a phenomenon in detail
- Typically uses personal interviews, document analysis, and observation
- Example:
  - How policymakers responded differently to the Cuban Missile Crisis?
  - Democratization in Spain in 1974
  - Success and failure of social revolution
Case Study Design

- Descriptive case studies
- Explanatory case studies
- Single case studies
- Comparative case studies
- Advantages and disadvantages
  - Useful for describing the process
  - Difficult to generalize findings
## Summary

<table>
<thead>
<tr>
<th>Types of Design</th>
<th>Establish that X and Y Covary</th>
<th>Establish that X Precedes Y</th>
<th>Establish Non-spuriousness</th>
<th>Establish Generalizability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>maybe</td>
</tr>
<tr>
<td>Time-series</td>
<td>yes</td>
<td>yes</td>
<td>maybe</td>
<td>maybe</td>
</tr>
<tr>
<td>Cross-sectional</td>
<td>yes</td>
<td>usually not</td>
<td>maybe</td>
<td>maybe</td>
</tr>
<tr>
<td>Descriptive case study</td>
<td>maybe</td>
<td>maybe</td>
<td>maybe</td>
<td>no</td>
</tr>
<tr>
<td>Exploratory case study</td>
<td>maybe</td>
<td>maybe</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Explanatory case study</td>
<td>yes</td>
<td>yes</td>
<td>maybe</td>
<td>maybe</td>
</tr>
</tbody>
</table>