COMPARATIVE POLITICAL ANALYSIS

I. Why Compare?

II. Four Basic Methods

III. The Case Study Method

IV. The Comparable Cases Strategy
FOUR REASONS WHY COMPARISON IS USEFUL

1. Description
2. Classification and Typology
<table>
<thead>
<tr>
<th>Rule For/Rule By</th>
<th>Common Good</th>
<th>Selfish Interests</th>
</tr>
</thead>
<tbody>
<tr>
<td>One</td>
<td>Monarchy</td>
<td>Tyranny</td>
</tr>
<tr>
<td>Few</td>
<td>Aristocracry</td>
<td>Oligarchy</td>
</tr>
<tr>
<td>Many</td>
<td>Polity</td>
<td>Democracy</td>
</tr>
</tbody>
</table>
FOUR REASONS WHY COMPARISON IS USEFUL (con’t)

3. Hypothesis-Testing

Hypothesis = A statement about a possible causal relationship between two or more concepts.
Theory = A definitive and logical statement(s) about how the world, or some key aspect of it, works. This is what Yin means by “propositions” that help spell out theoretical relations between variables.
Forming the Question

Case Studies testing theories can seem rather indeterminate when the research question is vague. Some studies will read like history or biographies, etc.
What is the unit?

Determining the units is also critical and aids in case selection and focus.

The issue of construct validity will be important. How does one define and measure variables in a case study? That is a fundamental question and related to reliability and internal validity.
Holistic vs. Embedded Case Studies

Holistic: One unit of analysis
- Cuban Missile Crisis explained by focusing only at Kennedy and Khrushchev.

Embedded: Multiple units of analysis.
- Cuban Missile Crisis explained by focusing not only leaders but bureaucracies within their governments such as US Air Force and Soviet Missile Forces.
A future Study could be done on California and Texas

Tort Reform

Lower Medical Bills

Malpractice Insurance Caps

Specifying the propositions that test or illustrate theoretical relationships
FOUR REASONS WHY COMPARISON IS USEFUL (con’t)

4. Prediction

"You will grow richer and richer. Then one day you’ll discover money can’t buy happiness—and you’ll give it all away."
TYPES OF CASE STUDIES

1. Descriptive (Configurative-Ideographic)
2. Plausibility Probes
3. Most Likely Case Studies (MLCS)
4. Least Likely Case Studies (LLCS)
5. Deviant Case Studies (Yin: critical cases?)
1. Most Similar Systems Design (MSSD)

Key is to identify what factor leads to dissimilar outcomes of Y when the cases appear rather similar in most regards.
Using MSSD to Explain Why Turnout Rates Vary in Three Countries

<table>
<thead>
<tr>
<th>Features</th>
<th>Country 1</th>
<th>Country 2</th>
<th>Country 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Culture</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>GNP/Capita</td>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>Pres v. Parl</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>Regis. Rules</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>Key Exp. Factor: Party System</td>
<td>Multi-party</td>
<td>Multi-party</td>
<td>Two-party</td>
</tr>
<tr>
<td>Outcome: Turnout</td>
<td>T.O. &gt; 75%</td>
<td>T.O. &gt; 75%</td>
<td>T.O. &lt; 55%</td>
</tr>
</tbody>
</table>
Using MSSD to explain the emergence of guerrilla movements

<table>
<thead>
<tr>
<th>Case</th>
<th>Cuba</th>
<th>Venz</th>
<th>Guate</th>
<th>Colo</th>
<th>Peru</th>
<th>Boli</th>
</tr>
</thead>
<tbody>
<tr>
<td>Key Peasant Groups</td>
<td>Squatters</td>
<td>Share-cropper</td>
<td>Tenant</td>
<td>Share-cropper</td>
<td>Serfdom</td>
<td>Smallholders</td>
</tr>
<tr>
<td>Outcome: Guerrilla Support</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Chi</th>
<th>Arg</th>
<th>Urg</th>
<th>Bra</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographic pressures</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>International pressures</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>Mobilization of Proponents</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>Mobilization of Opponents</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>Union Density</td>
<td>Low</td>
<td>High</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>Institutional Access for Veto Players</td>
<td>Limited</td>
<td>Medium</td>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td>Policy Outcome</td>
<td>Extensive reform</td>
<td>Moderate reform</td>
<td>Moderate reform</td>
<td>Limited reform</td>
</tr>
</tbody>
</table>
TYPES OF COMPARABLE CASE STRATEGIES (con’t)

2. Most Different Systems Design (MDSD)
Most Different System Designs

The key to this type of design is to understand that very different units/cases have the same outcome (Y variable). The search is then for a key explanatory variable common to the cases that all appear very different from each other.

Using MDSD to explain high levels of charitable giving.

<table>
<thead>
<tr>
<th>Features</th>
<th>Israel</th>
<th>Iran</th>
<th>U.S.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major Religion</td>
<td>A</td>
<td>D</td>
<td>G</td>
</tr>
<tr>
<td>GNP/capita</td>
<td>B</td>
<td>E</td>
<td>H</td>
</tr>
<tr>
<td>Tax codes</td>
<td>C</td>
<td>F</td>
<td>I</td>
</tr>
<tr>
<td>Key explanatory factor: religiosity</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Outcome to be explained: level of giving</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
</tbody>
</table>
Single-Case Studies

- Illustrations – often complementary with statistical methodologies
- Plausibility Probes – similar to pilot projects to see if more extensive/intensive studies would be useful and feasible. Yin calls these “exploratory.”
- Deviant Cases – Seek explanation for anomalies, which may lead to new theory development.
- Critical cases or other TESTS -- controversial
Case Study Methodologies

Problems to Avoid
Note that the issues to be discussed henceforth in this lecture are most applicable to case studies that are used as evidence to support hypotheses (explanatory studies) and not necessarily case studies as description, pilot projects, or illustrations.
Constructing Causal Theories: 5 Rules

Rule #1: Theories should be internally consistent. The logic must not be contradictory. If so, then the explanatory power may be limited. This is a fundamental point made by Yin.

“good description is better than bad explanation.” (King, Keohane, & Verba 1994)
Rule #2 Constructing Causal Theories

- The hypotheses should be falsifiable.
  - If one cannot think of any way for your arguments to be wrong, you have a problem.
  - What type of evidence will falsify your theory? Nonfalsifiable theories, like Marxism, because it makes several normative claims, cannot be empirically verified fully.
Rule #3 Constructing Causal Theories

Select Dependent Variables Carefully:
- Make sure Y variable is truly dependent and not endogenous.
- The Y variable should vary – do not truncate or limit through other aspects of research design.
Rule #4 Constructing Causal Theories

Maximize Concreteness:

- Define/operationalize concepts like culture or national interest – (narrow the gap between concept and indicator). This is the issue of construct validity in Yin and Trochin.

- Make sure your writing regarding such issues is clear.
Rule #5 Constructing Causal Theories

State theories to encompass as much as possible. Theories should:

- Generalize!!! -- applicable to many cases.
  - There may be a problem of External Validity.
- Maximize Leverage -- can account for phenomena in own study plus that of other competing theories.
Indeterminate Research Designs

This problem arises when there:

- Lack of variation in Y (mentioned above)
- More inferences than observable implications (over-specified)
- When competing X variables are highly correlated. (multicollinearity). This part will make more sense in the coming weeks.
Lack of variation in Y

If our dependent variables do not vary, then how can we fully attribute their value to variation in the X vars.

For example, if we have 2 cases where deregulation (X) resulted in less pollution but ignore the possibilities where pollution seemed to increase, our inference may be biased – there may be no relationship between regulation and pollution but will not see this unless we let Y vary.
Lack of variation in Y

This is the “Non-barking dog Problem” (Sherlock Holmes books)

- We need to explain why nonevents happen as well in regard to our X vars. For example, studies that look at how X vars are related to war without also looking at situations where wars do not occur are biased.
Over-specification

Here you have more X variables than cases (N).

In quantitative analysis, this is a degrees of freedom problem and a model will not run.

Essentially, how can you tell the difference between the effects of several X variables in a few cases when all the X’s are present. You need variation through additional cases, or fewer X’s.
The biggest problem: Case Selection

Random selection is optimal but not feasible in case studies.

However, some case studies select on the dependent variable in a biased manner.

- Do not select cases that are known to support the hypotheses. The cases that inspired the theory should also not be used as evidence (tautology).
- Again, letting $Y$ vary partially avoids this problem.
Case Selection and Randomness

Note that an inference is based on a sample from a greater population. The more cases the more accurate our inferences.

Hence, case studies are more error prone in making inferences, although they may do a better job telling us how processes work. This must be kept in mind. External and Internal validity are harder to obtain.
Case Selection and Randomness

Statistical analyses, however, are better at showing the overall significance of variables and measuring the magnitudes of their effects (precision).

Every study must be sober regarding its own limits and claims considering methodological trade-offs.
Since case studies can contain a limited amount of variation (because of low N), it is important that any control variables included are essential. You may omit when:

- If control X does not affect Y.
- If X control is uncorrelated with theoretical X. (but include if forecasting/predicting if control X affects Y, or else error occurs).
Seeking Variation through multiple cases

One can always gain more explanatory power by expanding observations across space or time.

- For space, add more countries, states, cities, etc.
- For time, expand into panel or time series format across one or multiple cases.
Single Case studies and Variation

How can a single case study be used for inferences? If it remains a single observation ($N=1$), it can not.

The trick is to produce multiple $N$ by breaking into subunits or time units.

- Studying Mexico states.
- Studying policy changes across time.
Crucial (Critical) Case Studies?

Eckstein (1975) A crucial case is one where something occurs where/when it should be least likely to occur, and thus can be enough evidence to support a theory.

No single case could ever proof a theory.
- Alternative Explanations? Spurious?
- Measurement Error?
- Determinism? The world is probabilistic and cannot rule out stochastic chance.
Multicollinearity

Here, the problem is that X variables can be used to predict other X variables. You need to cut the X’s but also avoid omitted variable bias.

For example, in attempting to explain poverty, we would not want to include both a city’s literacy rate and average property tax if they are highly correlated (higher property tax may mean more education money).
Biased Data

Also beware of biases that occur naturally in data, or lack of data.

- For example, if one wanted to study how economic performance leads to, or is affected, by civil war, the results may be biased by not including states that are most severely affected and for which little data is available (Congo, Cambodia, etc.)
Nonsystematic Error

Data also may have nonsystematic error that does not bias the results but makes predictions less efficient. All cases may be measured too high or low.

In other words, theoretical relationships may be verified between variables but the magnitude of the effects too low or high.