POLS 205
Political Science as a Social Science

Examples of Theory-Building in Political Science

Christopher Adolph

University of Washington, Seattle

April 7, 2010
What is the research question?

What is the dependent variable?

What is the unit of analysis?

What are the key independent variables?

What are the hypotheses and alternatives?
Ecological fallacy

Correlations among aggregates do not imply the same correlations among individuals
Ecological fallacy

Correlations among aggregates do not imply the same correlations among individuals

Example:

Aggregate correlation: In the early 20th century, US states with many immigrants also had more college educated residents. In symbols, $\text{corr}(\text{Immigrants}_{\text{State}}, \text{CollegeEdu}_{\text{State}}) > 0$

But

Individual correlation: Did immigrants have more years of education than natives? No. They had less: $\text{corr}(\text{Immigrant}_i, \text{CollegeEdu}_i) < 0$
Ecological fallacy

Correlations among aggregates do not imply the same correlations among individuals

Example:

Aggregate correlation: In the early 20th century, US states with many immigrants also had more college educated residents. In symbols, \( \text{corr}(\text{Immigrants}_{\text{State}}, \text{CollegeEdu}_{\text{State}}) > 0 \)

But

Individual correlation: Did immigrants have more years of education than natives? \textit{No}. They had less: \( \text{corr}(\text{Immigrant}_i, \text{CollegeEdu}_i) < 0 \)
Ecological fallacy

Correlations among aggregates do not imply the same correlations among individuals

Example:

Aggregate correlation: In the early 20th century, US states with many immigrants also had more college educated residents. In symbols, $\text{corr(Immigrants}_{\text{State}}, \text{CollegeEdu}_{\text{State}}) > 0$

But

Individual correlation: Did immigrants have more years of education than natives? No. They had less: $\text{corr(Immigrant}_{i}, \text{CollegeEdu}_{i}) < 0$

What is going on?
Ecological fallacy

Assuming aggregate correlations tell us about individual correlations is a fallacy, or frequently misleading practice.

Individual level correlations can be very different from the aggregate ("ecological") correlations.

Note: this is the original meaning of ecology, a collection of individuals, not the modern environmental usage.

Recognition of the ecological fallacy in the early 1950s changed course of social science towards individual level surveys.
What’s the matter with Kansas?

Republican Party advocates & implements policies which benefit high income: low marginal tax rates, low taxes on capital, a small welfare state.
What’s the matter with Kansas?

Republican Party advocates & implements policies which benefit high income: low marginal tax rates, low taxes on capital, a small welfare state

So rational, materialist voters should split on income: high income voters → Republicans; low income voters → Democrats
What’s the matter with Kansas?

Republican Party advocates & implements policies which benefit high income: low marginal tax rates, low taxes on capital, a small welfare state.

So rational, materialist voters should split on income: high income voters $\rightarrow$ Republicans; low income voters $\rightarrow$ Democrats.

Thomas Frank and others note that Kansas has a lower average income than most US states, but a higher Republican vote.
What's the matter with Kansas?

Republican Party advocates & implements policies which benefit high income: low marginal tax rates, low taxes on capital, a small welfare state.

So rational, materialist voters should split on income: high income voters → Republicans; low income voters → Democrats.

Thomas Frank and others note that Kansas has a lower average income than most US states, but a higher Republican vote.

Does this finding contradict rational materialist voting?
What’s the matter with Kansas? Nothing!

Frank has committed the ecological fallacy. To show that Kansans vote irrationally, we need individual level data.
Republican vote by state in 2004

Vote share for George Bush

Average income within state

30%
50%
70%

$20,000
$30,000
Bush vote in 2004 by income

Vote share for George Bush

Family income

30%

50%

70%

0

$100,000

$200,000
The graph illustrates the probability of voting for Bush among different income groups in three states: Mississippi, Ohio, and Connecticut. The x-axis represents three income groups: poor voters, middle-income voters, and rich voters. The y-axis represents the probability of voting for Bush, ranging from 25% to 75%. The graph shows a clear trend where the probability of voting for Bush increases with income for each state.
Republican vote for regular church attenders minus Republican vote for nonattenders

Difference in Republican vote

20%

0

1960 1980 2000
Bush vote in 2004 by income and religious attendance

- if you attend church more than once/week
- if you attend once or twice/month
- if you never attend church

Probability of voting for Bush

poor  middle-income  rich
Rich-state, poor-state gap in Republican vote among poor, middle-income, and rich voters

- High-income voters
- Middle-income voters
- Low-income voters

Republican vote in poor states, minus Republican vote in rich states

Whites only: Rich-state, poor-state gap in Republican vote among poor, middle-income, and rich voters

- High-income white voters
- Middle-income white voters
- Low-income white voters

What is the research question?

What is the dependent variable?

What is the unit of analysis?

What are the key independent variables?

What are the hypotheses and alternatives?