

Conducting Empirical Legal Scholarship

Lee Epstein¹ Andrew D. Martin²

¹Northwestern University
School of Law & Department of Political Science

²Washington University in St. Louis
School of Law & Department of Political Science

The Course

- 1 Overview of Empirical Research
- 2 Designing Research
- 3 Collecting & Coding Data
- 4 Statistical Software
- 5 The Logic of Statistical Inference
- 6 Data Analysis

The Course

- 1 Overview of Empirical Research
 - Definition
 - The Basics
- 2 Designing Research
 - The Research Question
 - Theories and Their Observable Implications
 - Rival Hypotheses
 - Measurement
- 3 Collecting & Coding Data
 - From Where?
 - How Much?
 - Coding Data—> Statistical Software
- 4 Statistical Software

What is Empirical Research?

General Idea

Research that is based on observations of the world—in other words, on **data**, which is just a term for facts about the world. “Data” may be quantitative (numerical) or qualitative; neither is any more empirical than the other.

Distinct Feature

Many empirical projects start with hunches or “theories” about the way the world work but they do not end there; they attempt to determine whether those hunches or theories coincide with observations taken from the world.

How Would We Do Research if We Had **NO** Constraints?

How Do We Conduct Empirical Research?

Three Key Components

- Design your project
- Collect and code the data
- Analyze the data

The Course

- 1 Overview of Empirical Research
- 2 **Designing Research**
 - The Research Question
 - Theories and Their Observable Implications
 - Rival Hypotheses
 - Measurement

The Research Question

General Idea

Virtually all empirical research starts with a question you want to answer.

A Problem

But many questions are conceptual—that is, they are in a form that are difficult to answer empirically.

- Can we predict judges' votes in ideologically contested areas on the basis of the judges' ideology?
- Can we predict judges' votes in 13 areas of the law on the basis of the political party of the President who appointed them?

The Course

- 1 Overview of Empirical Research
 - Definition
 - The Basics
- 2 Designing Research
 - The Research Question
 - Theories and Their Observable Implications
 - Rival Hypotheses
 - Measurement

Definitions

Once you have selected a research question, begin **theorizing** about possible answers that you can use to generate **observable implications** (a/k/a hypotheses or expectations).

- Theorizing: developing a reasonable and precise speculation about the answer to a research question.
- Observable implications: things we would expect to observe in the real world if our theory is right.

Major Issues

- 1 How do you move from theory to observable implications?
- 2 What form do observable implications take?

Moving from Theory to Observable Implications

Example: Preference Theory

- Theory: Judges want the law to reflect their ideological preferences. And, because they lack an electoral connection, they are free to vote in accord with those preferences.
- Observable Implication: We should observe judges voting in accord with their ideological preferences, such that conservative judges cast conservative votes and liberals, liberal votes (**all else being equal**)

Example: Whistle-Blowing Theory

- Theory: Judges want the law to reflect their ideological preferences. But, because they worry about being overturned, they won't vote their preferences in the presence of a potential whistle blower.
- Observable implication: We should observe conservative judges casting conservative votes unless there's a liberal on the panel (**all else being equal**).

The Form of Observable Implications

Overview

Observable Implication: Preference Theory

If my theory is correct, we should observe judges voting in accord with their ideological preferences, such that conservative judges cast conservative votes and liberals, liberal votes.

Form of the Observable Implication

It is a claim about the relationship between (or among) variables that we can, at least in principle, observe.

- Variables: Characteristics that vary, such as ideology and votes.
- Variable types:
 - Dependent variables: what we are trying to explain (here, votes).
 - Independent variables: those that help account for what we are trying to explain (here, ideology).

The Form of Observable Implications

Examples

- If cities restrict the abortion right, the number of abortions will decrease.
- Women are more likely than men to favor environmental regulation.
- If a person is a strong political partisan, the likelihood that she will campaign for her party's candidate will increase.
- Crime rates are higher in states with strict gun control laws than in states without strict gun control laws.
- The lower the education level, the higher the odds of smoking.
- In comparing voters, those whose economic situations have gotten better will be more likely to vote for the incumbent candidate than are voters whose economic situations have gotten worse.

The Course

- 1 Overview of Empirical Research
 - Definition
 - The Basics
- 2 **Designing Research**
 - The Research Question
 - Theories and Their Observable Implications
 - **Rival Hypotheses**
 - Measurement

Overview

- You must consider alternative explanations—those that do not square with the theory you are offering (i.e., rival explanations or hypotheses).
- In all likelihood, you will have to incorporate these rival variables into your analysis (i.e., “control” for them).

Examples

Environmental Regulation

From various theories of gender & politics, we derive the following observable implication: Women are more likely than men to favor environmental regulation.

- What rival explanations should we consider?

Ideological Voting

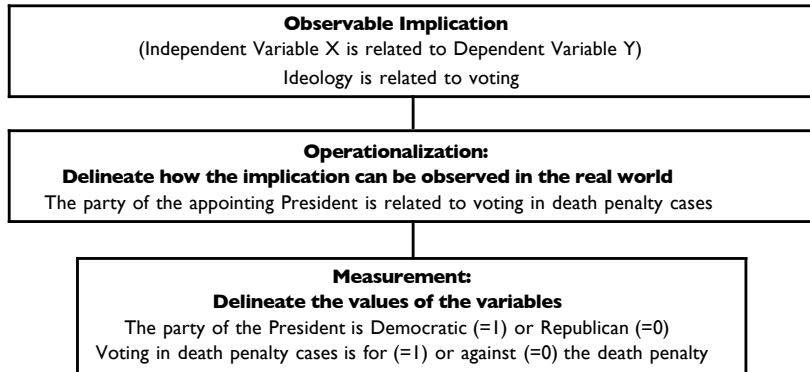
From preference-based theories, Sunstein et al. suggest that we should observe judges voting in accord with their ideological preferences, such that conservative judges cast conservative votes and liberals, liberal votes.

- What rival explanations should they consider?

The Course

- 1 Overview of Empirical Research
 - Definition
 - The Basics
- 2 **Designing Research**
 - The Research Question
 - Theories and Their Observable Implications
 - Rival Hypotheses
 - **Measurement**

Overview



Evaluating Measures/Measurement Methods

Reliability

The extent to which it is possible to replicate a measure, reproducing the same value (regardless of whether it is the right one) on the same standard for the same subject at the same time.

Validity

The extent to which a reliable measure reflects the underlying concept being measured.

- Facially valid: consistent with prior evidence or knowledge
- Approximately unbiased: right on average across repeated applications
- Efficiency: degree of reliability for unbiased measures

Example

Party of the Appointing President as a Measure of Ideology

Reliability

The extent to which it is possible to replicate a measure, reproducing the same value (regardless of whether it is the right one) on the same standard for the same subject at the same time.

Validity

The extent to which a reliable measure reflects the underlying concept being measured.

- Facially valid: consistent with prior evidence or knowledge

Example

Face Validity: Ideology of Presidents

President	Economic	Social
Roosevelt (D)	82.5	71.6
Truman (D)	74.6	74.3
Eisenhower (R)	38.8	36.9
Kennedy (D)	65.4	66.4
Johnson (D)	78.2	83.5
Nixon (R)	47.7	44.9
Ford (R)	38.8	39.3
Carter (D)	60.3	67.0
Reagan (R)	17.6	18.0
Bush (R)	33.1	32.8
Clinton (D)	63.1	72.0

Example

Party of the Appointing President as a Measure of Ideology

Validity

The extent to which a reliable measure reflects the underlying concept being measured.

- Facially valid: consistent with prior evidence or knowledge
- Approximately unbiased: right on average across repeated applications
- Efficiency: degree of reliability for unbiased measures

Solution

Consider/compare other measures

- NOMINATE Score of Appointing President
- Judicial Common Space Score of the Judge

The Course

- 1 Overview of Empirical Research
- 2 Designing Research
- 3 Collecting & Coding Data**
 - From Where?
 - How Much?
 - Coding Data—> Statistical Software
- 4 Statistical Software

From Where?

Countless Data-Generation Sources/Methods

Interviews, surveys, public records, archives, experiments, field research, and on and on

Some Tips

- Don't rely on surveys or interviews when "revealed preferences" are available.
- Without training, don't write your own surveys.
- Check out existing sources before spending lots of time and \$.

The Course

- 1 Overview of Empirical Research
 - Definition
 - The Basics
- 2 Designing Research
 - The Research Question
 - Theories and Their Observable Implications
 - Rival Hypotheses
 - Measurement
- 3 Collecting & Coding Data
 - From Where?
 - **How Much?**
 - Coding Data—> Statistical Software

How Much?

As much as you can!

Draw a random sample

Each member of the total population has a known probability of being selected.

- Equal probability sample
- Stratified random sample
- Cluster sample

The Course

- 1 Overview of Empirical Research
 - Definition
 - The Basics
- 2 Designing Research
 - The Research Question
 - Theories and Their Observable Implications
 - Rival Hypotheses
 - Measurement
- 3 Collecting & Coding Data
 - From Where?
 - How Much?
 - Coding Data—> Statistical Software

Overview

Definition

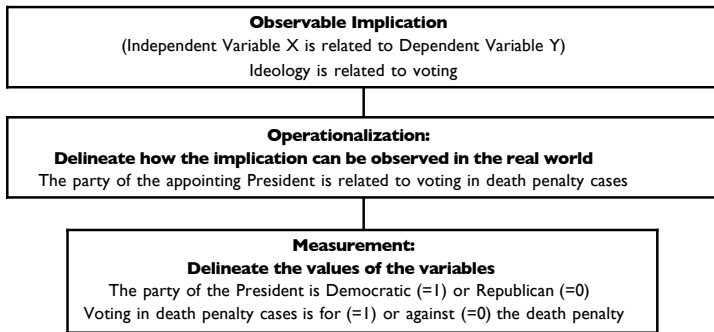
Coding data refers to the process of translating variables into a form you can systematically analyze.

Steps

- 1 Develop your measures.
- 2 Create your code book.
- 3 Enter your data.

Develop Your Measures

Measures should follow from your theory and observable implications.



Create Your Code Book

Basic Idea

- A codebook is a guide to the data. With the data and the code book, anyone should be able to replicate your analysis.
- At minimum, a code book should list all the variables and associated values.

Create Your Code Book

Example

Party of the Appointing President

0 = Republican

1 = Democrat

Note: Data on the party of the appointing president obtained from Smith (2004). N= 181. Cases identified by a LEXIS search on “capital punishment,” 01/01/95 - 12/31/02.

Vote on the Death Penalty

0 = Vote against the defendant (“conservative”)

1 = Vote in favor of the defendant (“liberal”)

Note: If a judge voted to grant the defendant any relief, then code the vote as for the defendant. N=181. Cases identified by a LEXIS search on “capital punishment,” 01/01/95 - 12/31/02.

Some Tips for Coding Data

- Ensure that the values of the variables are exhaustive
- Create more, rather than fewer, values but don't sacrifice reliability
- Establish that the values of the variables are mutually exclusive

The Course

- 1 Overview of Empirical Research
 - Definition
 - The Basics
- 2 Designing Research
 - The Research Question
 - Theories and Their Observable Implications
 - Rival Hypotheses
 - Measurement
- 3 Collecting & Coding Data
 - From Where?
 - How Much?
 - Coding Data—> Statistical Software
- 4 Statistical Software

Statistical Software

<STATA DEMONSTRATION>