Theory Development and Hypothesis Generation

Department of Government
London School of Economics and Political Science

1 Finish Measurement

2 Theory

3 Generating Hypotheses

4 Hypothesis Testing

5 Preview

1 Finish Measurement

2 Theory

3 Generating Hypotheses

4 Hypothesis Testing

5 Preview

Assessing Measurement Quality

1. Conceptual clarity
2. Construct validity
   - Convergent validity
   - Divergent validity
3. Accuracy and precision

Assessing Measures I

- Conceptual clarity is about knowing what we want to measure
- Sloppy concepts make for bad measures
  - Ambiguity
  - Vagueness

Assessing Measures I

- Conceptual clarity is about knowing what we want to measure
- Sloppy concepts make for bad measures
  - Ambiguity
  - Vagueness
- Revise concept definition as needed

Assessing Measures II

- Construct validity is the degree to which a variable measures a concept[1]

[1]Note: Kellstedt and Whitten call this "content validity". They use "construct validity" to mean whether a measure has predictive validity (i.e., that the measure is related to measures of other concepts that are theorized to be related).

Assessing Measures II

- Construct validity is the degree to which a variable measures a concept[1]
- Construct validity is **high** if a variable is a measure of the concept we care about

[1] Note: Kellstedt and Whitten call this "content validity". They use "construct validity" to mean whether a measure has predictive validity (i.e., that the measure is related to measures of other concepts that are theorized to be related).

Assessing Measures II

- Construct validity is the degree to which a variable measures a concept[1]
- Construct validity is **high** if a variable is a measure of the concept we care about
- Construct validity is **low** if a variable is actually a measure of something else

[1] Note: Kellstedt and Whitten call this "content validity". They use "construct validity" to mean whether a measure has predictive validity (i.e., that the measure is related to measures of other concepts that are theorized to be related).

Assessing Construct Validity

- Multiple measures!
- Convergence (Convergent validity)
- Discrimination (Discriminant validity)

1 Finish Measurement

2 Theory

3 Generating Hypotheses

4 Hypothesis Testing

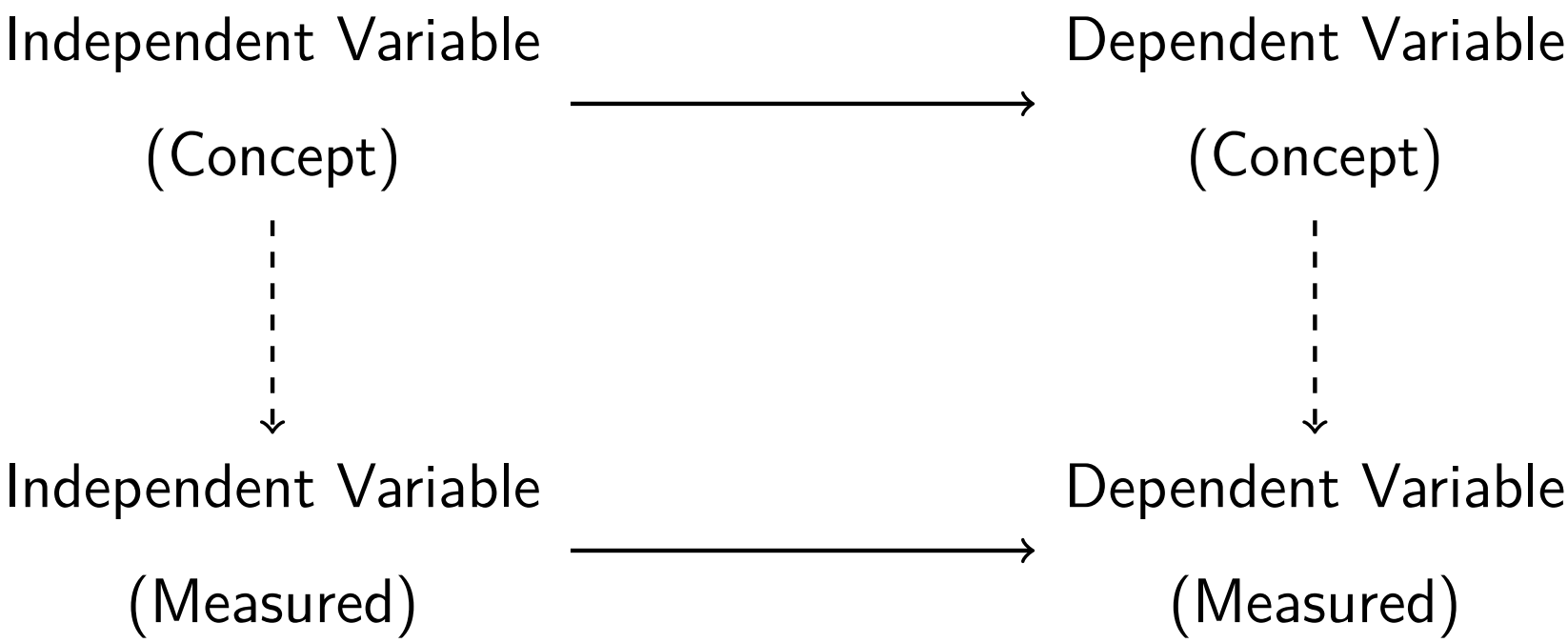
5 Preview

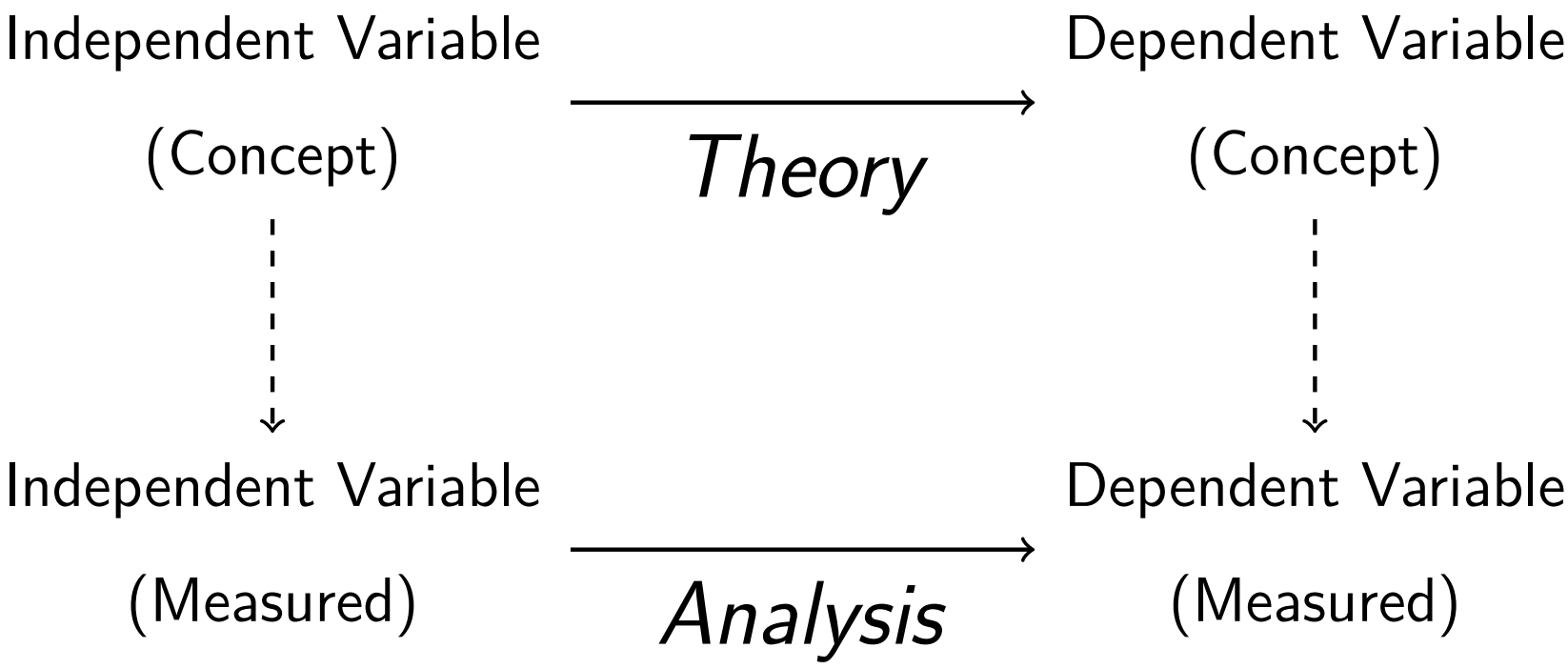
Scientific method

1. Research question(s)
2. Clarify the core concepts
3. **Develop theory**
4. Derive specific, testable hypotheses
5. Plan data collection
6. Gather data/evidence
7. Analyze data
8. Draw inferences

Key Points from Last Week

1. Theory is about concepts
2. Analysis is about measured variables
3. So our task as scientists is to:
   - Find observable implications of theory
   - Draw theoretical implications from measures


Independent Variable
(Concept)
Dependent Variable
(Concept)
Independent Variable
(Measured)
Dependent Variable
(Measured)


Independent Variable
(Concept)
Theory
Dependent Variable
(Concept)
Independent Variable
(Measured)
Analysis
Dependent Variable
(Measured)

What is a theory?

- Kellstedt and Whitten's definition:[2]
  A tentative conjecture about the causes of some phenomenon of interest

[2] Kellstedt and Whitten, p.3

What is a theory?

- Kellstedt and Whitten's definition:[2]
  A tentative conjecture about the causes of some phenomenon of interest
- Another way of saying this:
  An argument that attempts to explain how concepts are causally related

[2] Kellstedt and Whitten, p.3

What is a theory?

- Kellstedt and Whitten's definition:[2]
  A tentative conjecture about the causes of some phenomenon of interest

- Another way of saying this:
  An **argument** that attempts to **explain** how concepts are causally related

[2] Kellstedt and Whitten, p.3

Generating Theory I

- One way to theorize is to reason *inductively*
- Induction works by drawing generalities from specific observations
- Sometimes called "bottom-up" theorizing

Generating Theory II

- An alternative way of developing theory is through *deduction*
- Deduction begins from general, assumed principles/axioms to reach more specific observable realities

Generating Theory II

- An alternative way of developing theory is through *deduction*
- Deduction begins from general, assumed principles/axioms to reach more specific observable realities
- Common example: Rational choice theory

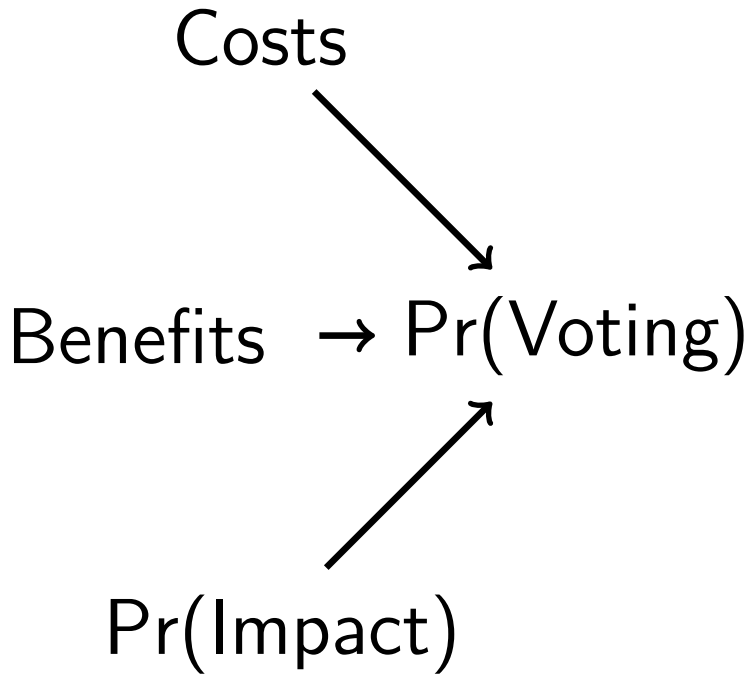
Generating Theory III

- "The Calculus of Voting" is a *rational choice* theory
  - Assumes utility maximization is the driver of all behaviour
  - Understanding phenomena is a matter of figuring out utility structures, especially those created by institutions

The Calculus of Voting

Theory: Voting is explained by 3 factors

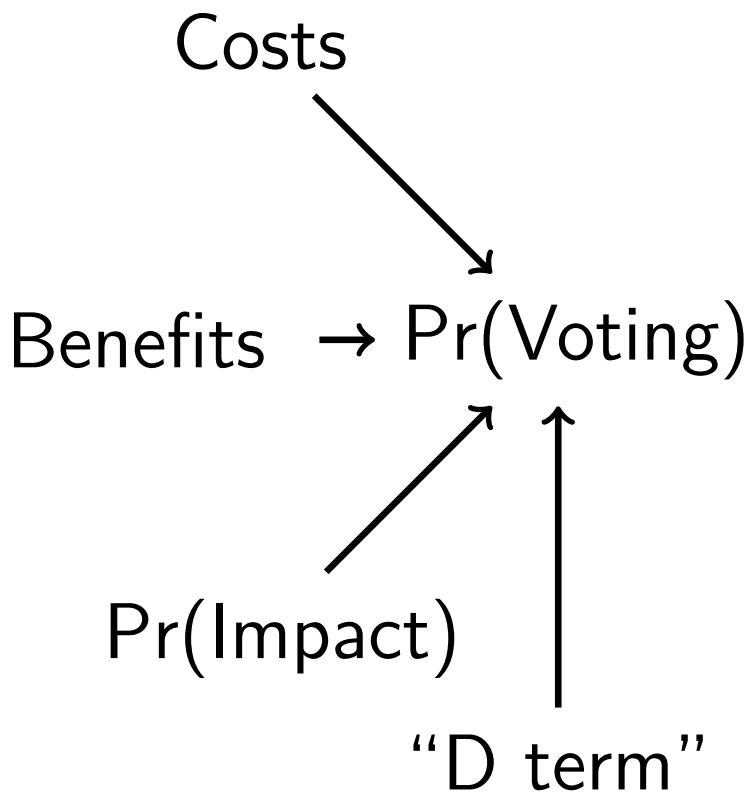
- Costs of voting
- Benefits from preferred alternative winning
- Probability of impacting result



The Calculus of Voting

Theory: Voting is explained by 4 factors

- Costs of voting
- Benefits from preferred alternative winning
- Probability of impacting result
- Benefits from voting *per se*



Aside: Assumptions

If a theory require assumptions, is that theory credible?

Generating Theory III

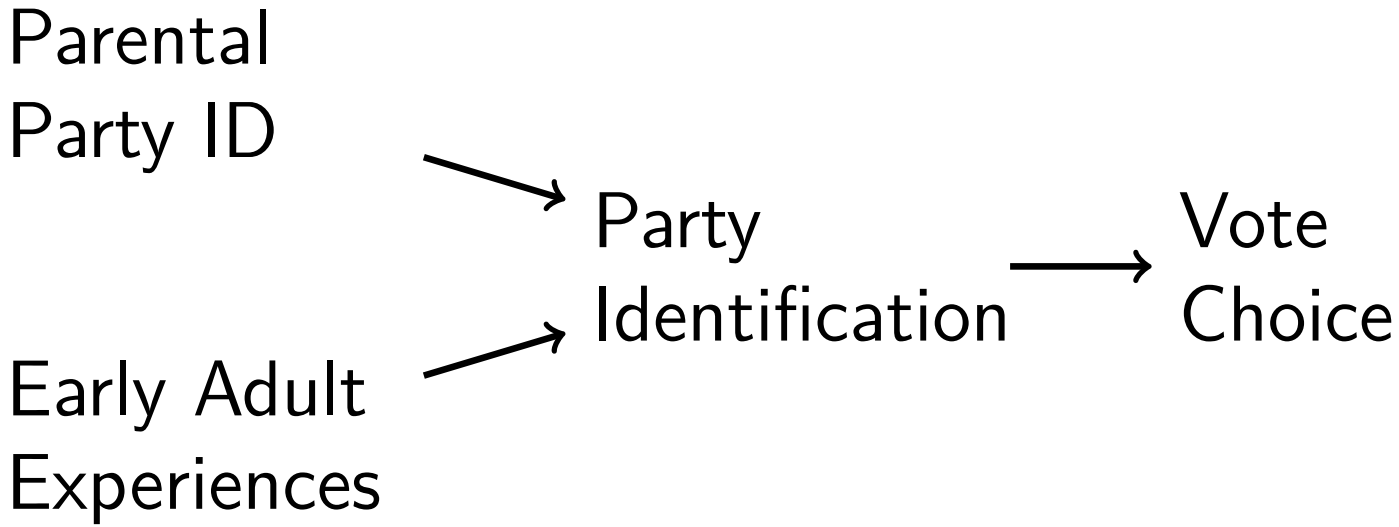
- "The Calculus of Voting" is a *rational choice* theory
  - Assumes utility maximization is the driver of all behaviour
  - Understanding phenomena is a matter of figuring out utility structures, especially those created by institutions

Generating Theory III

- "The Calculus of Voting" is a *rational choice* theory
  - Assumes utility maximization is the driver of all behaviour
  - Understanding phenomena is a matter of figuring out utility structures, especially those created by institutions
- Not the only broad theoretical paradigm

The Michigan Model

Theory: Vote choice is explained by long-standing partisan identification, which is in turn shaped by early socialization.



Induction vs. Deduction?

- Induction and deduction are both integral to science
- Theory testing and theory building both require observation

Theory Generation in Practice

As you theorize an explanation for some phenomenon, you will draw on:

- General principles
- Extant theory
- Specific evidence

What makes for a good theory?

- Truth
- Falsifiability
- Relevance
- Coherence
- Generality
- Parsimony

Generality & Parsimony

Think for 90 seconds about each of these principles:

- Generality: Theories that can explain more are preferred over theories that can explain less
- Parsimony: Simple theories are preferred over complex theories

Are these principles defensible?
Are they any good?

1 Finish Measurement

2 Theory

3 Generating Hypotheses

4 Hypothesis Testing

5 Preview

Hypotheses

[3] Kellstedt and Whitten (p.4)

Hypotheses

- Definition: a theory-based statement about a relationship that we expect to observe.[3]

[3] Kellstedt and Whitten (p.4)

Hypotheses

- Definition: a theory-based statement about a relationship that we expect to observe.[3]
- Features
  - Derived from theory
  - Specific
  - Empirical/observable

[3] Kellstedt and Whitten (p.4)

How do we generate hypotheses?

- Think about *observable implications*
- What would evidence consistent with this theory be?
- What would evidence inconsistent with this theory be?

How do we generate hypotheses?

- Think about *observable implications*
- What would evidence consistent with this theory be?
- What would evidence inconsistent with this theory be?
  - This is *falsifiability*

Example: Broad Street Cholera

- 1854 outbreak of cholera in London
  - Around Broad Street (Soho)
  - 616 eventual deaths

Example: Broad Street Cholera

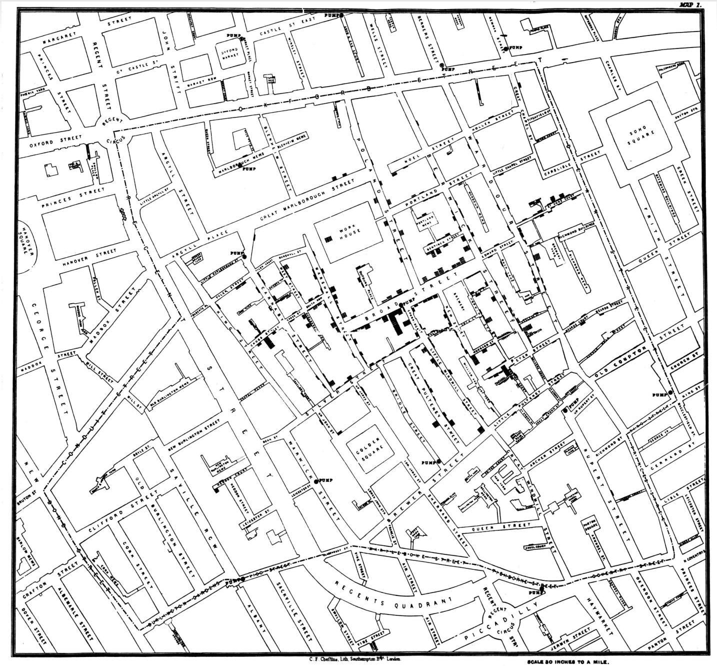
- 1854 outbreak of cholera in London
  - Around Broad Street (Soho)
  - 616 eventual deaths
- What causes transmission of cholera?

Example: Broad Street Cholera

- 1854 outbreak of cholera in London
  - Around Broad Street (Soho)
  - 616 eventual deaths
- What causes transmission of cholera?
- Dominant theory at time: “miasma”

Example: Broad Street Cholera

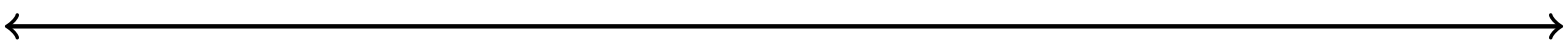
- 1854 outbreak of cholera in London
  - Around Broad Street (Soho)
  - 616 eventual deaths
- What causes transmission of cholera?
- Dominant theory at time: "miasma"
- Hypotheses:
  - Clean up garbage $\rightarrow$ $\downarrow$ cholera
  - Open windows $\rightarrow$ $\downarrow$ cholera


MAP 1.
OXFORD STREET
REGENT CIRCUS
PRINCES STREET
HANOVER SQUARE
HANOVER STREET
GREAT MARLBOROUGH STREET
ARGYLL PLACE
WORK HOUSE
BROAD STREET
PORTLAND
SOHO SQUARE
GOLDEN SQUARE
OLD COMPTON STREET
QUEEN STREET
REGENTS QUADRANT
PICCADILLY
REGENT CIRCUS
BREWER STREET
NEW BURLINGTON STREET
SAVILLE ROW
CLIFFORD STREET
OLD BURLINGTON STREET
CORK STREET
ALBEMARLE STREET
DOVER STREET
SACKVILLE STREET
AIR STREET
HAYMARKET
RUPERT STREET
WARDOUR STREET
POLAND STREET
DEAN STREET
FRITH STREET
GREEK STREET
CHARLES ST
GEORGE STREET
MADDOX STREET
NEW BOND STREET
GRAFTON STREET
GERRARD ST
LISLE STREET
CHURCH ST
PUMP
C.F. Cheffins, Lith, Southampton Bldgs. London.
SCALE 30 INCHES TO A MILE.

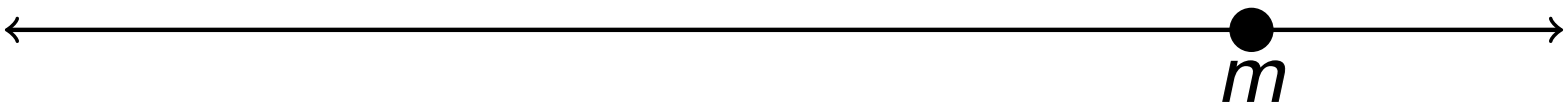
Observational Equivalence

- Definition: All hypotheses for two (or more) theories are identical
- What to do?
  - Generate more specific expectations
  - Move outside scope conditions
  - Settle for lack of explanation

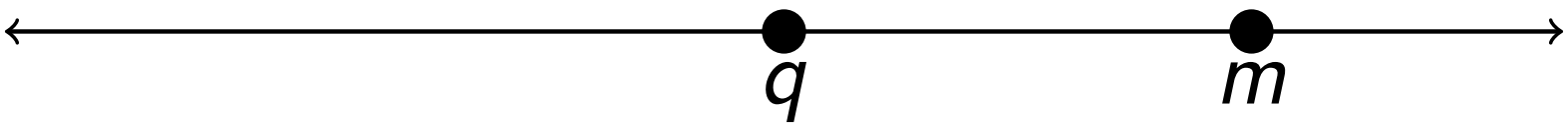
Median Voter Theory of Legislatures



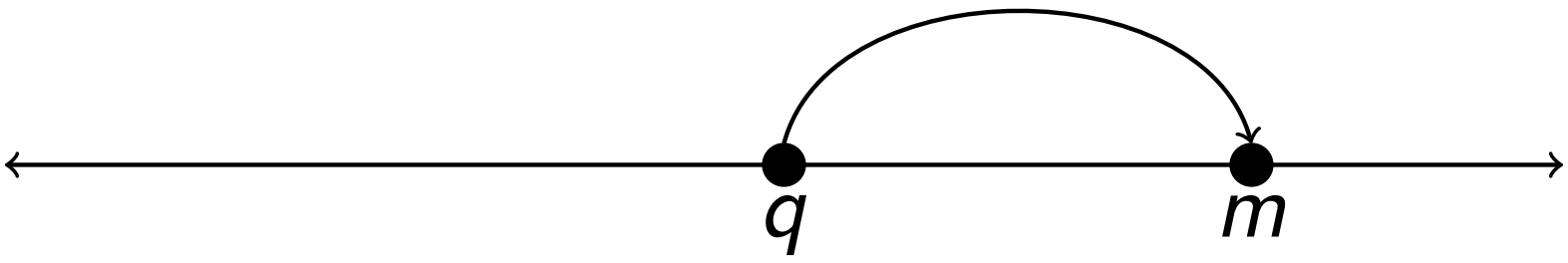
Median Voter Theory of Legislatures



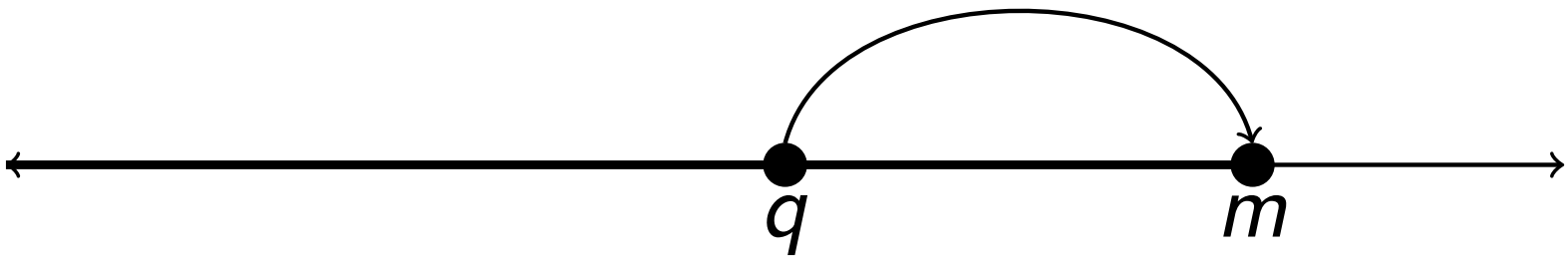
Median Voter Theory of Legislatures



Median Voter Theory of Legislatures



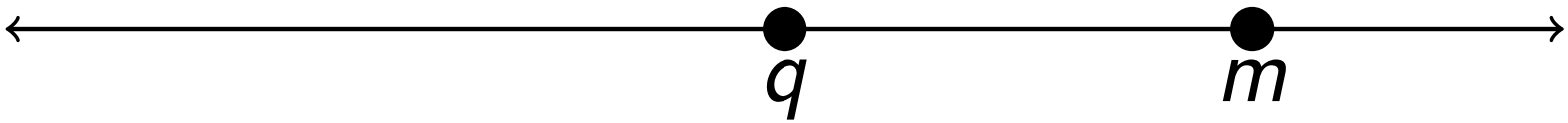
Median Voter Theory of Legislatures



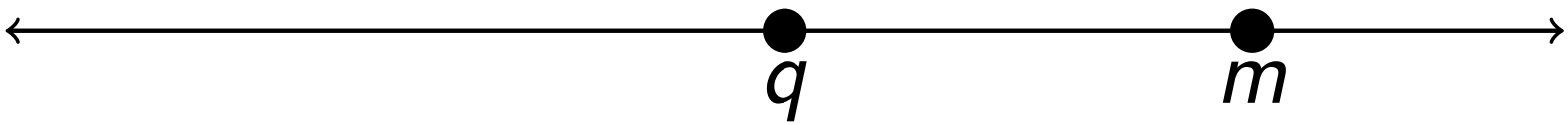
If this is true, why do we sometimes see policies left of m in the U.S. House?

Three Competing Theories

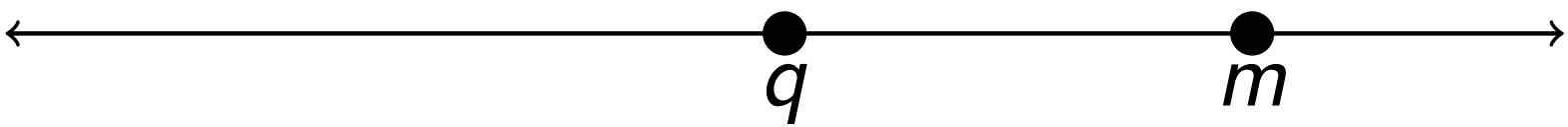
Committee Gatekeeping



Party Cartels



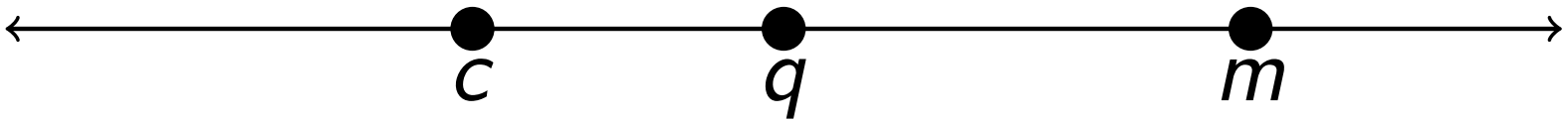
Pivotal Politics



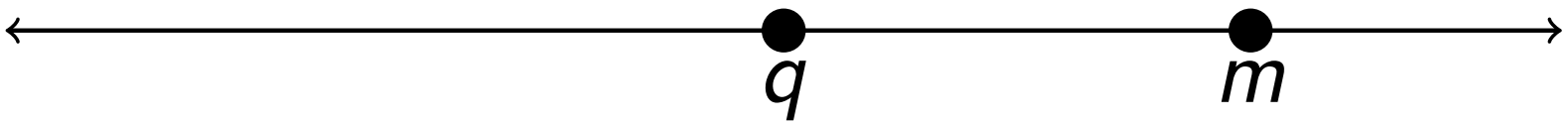
Adapted From: Krehbiel, Keith. 1999. "Paradoxes of Parties in Congress." *Legislative Studies Quarterly* 24(1): 31–64.

Three Competing Theories

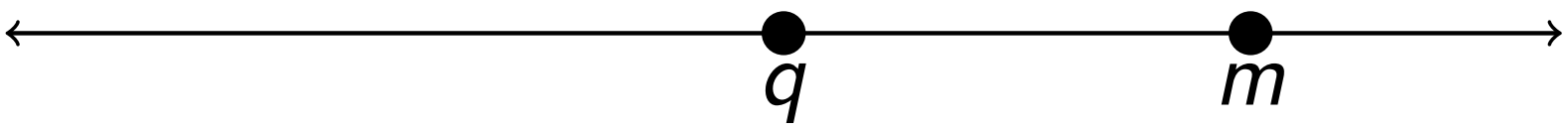
Committee Gatekeeping



Party Cartels



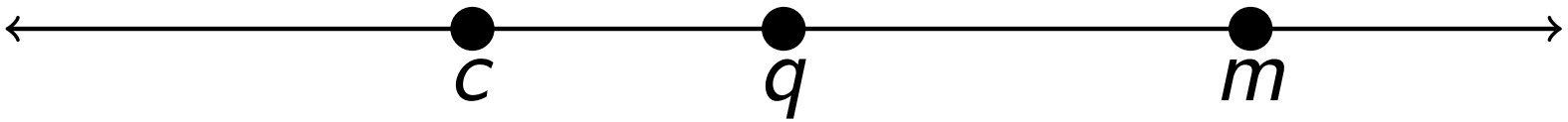
Pivotal Politics



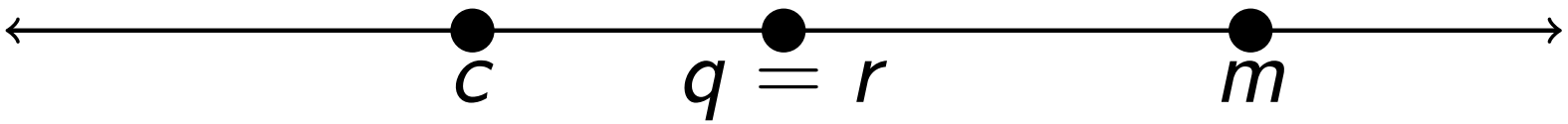
Adapted From: Krehbiel, Keith. 1999. "Paradoxes of Parties in Congress." *Legislative Studies Quarterly* 24(1): 31–64.

Three Competing Theories

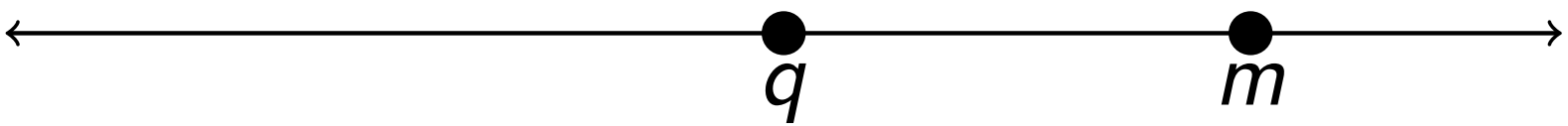
Committee Gatekeeping



Party Cartels



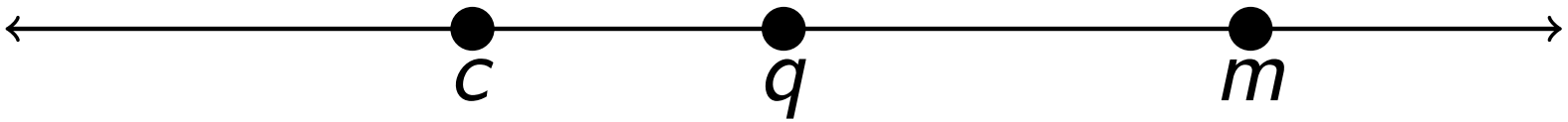
Pivotal Politics



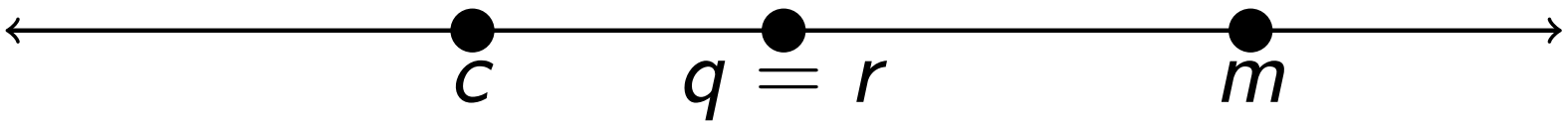
Adapted From: Krehbiel, Keith. 1999. "Paradoxes of Parties in Congress." *Legislative Studies Quarterly* 24(1): 31–64.

Three Competing Theories

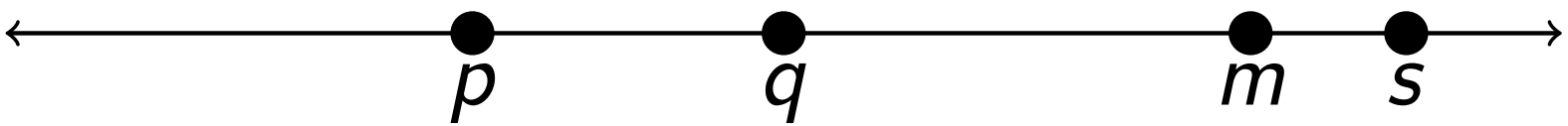
Committee Gatekeeping



Party Cartels



Pivotal Politics



Adapted From: Krehbiel, Keith. 1999. "Paradoxes of Parties in Congress." *Legislative Studies Quarterly* 24(1): 31–64.

1 Finish Measurement

2 Theory

3 Generating Hypotheses

4 Hypothesis Testing

5 Preview

Hypothesis Testing

- Multiple schools of thought
- History is conflictual and murky
- Two strands of literature
  - Philosophy of science
  - Statistics

Principle of Hypothesis Testing

1. Identify and collect data

Principle of Hypothesis Testing

1. Identify and collect data
2. Data should include:
   - Independent variable(s)
   - Dependent variable(s)

Principle of Hypothesis Testing

1. Identify and collect data
2. Data should include:
   - Independent variable(s)
   - Dependent variable(s)
3. Need variation on both

Principle of Hypothesis Testing

1. Identify and collect data
2. Data should include:
   - Independent variable(s)
   - Dependent variable(s)
3. Need variation on both
4. Test difference between outcomes when (possibly) causal variable differs

Forms of Hypothesis Testing

Null hypothesis

Begin with *null hypothesis*

Your hypothesis expects an alternative state of the world

c/o Ronald Fisher

Alternative hypotheses

Begin with 2(+) alternative hypotheses

Accept hypothesis consistent with observation

c/o Jerzy Neyman and Egon Pearson

Fearon's Counterfactuals

- Sometimes we cannot test our hypothesis with actual observations
- What does Fearon suggest we do?

A Good Test

- Correct level of analysis
- Within scope conditions of theory
- Well-defined concepts
- Measures of high construct validity, accuracy, and precision
- Possible to observe any correlation between potential cause and outcome
- Consistent with or an improvement upon past methods
- Test using different data than data used to generate theory

Some Testing Challenges

1. Deterministic and probabilistic causality
2. Effect heterogeneity
3. Multiple causation
4. Equifinality
5. Confirmation or disconfirmation bias

1 Finish Measurement

2 Theory

3 Generating Hypotheses

4 Hypothesis Testing

5 Preview

Preview of Next Week

- What is a case?
- What are case studies?
- How do we use case studies to test and/or build theories?