Class 2: Conceptualization, Operationalization, and Measurement

LAW 7675: Legal Analytics 2 (Spring 2021)

Charlotte Alexander & Susan Navarro Smelcer

Georgia State University
College of Law

January 19, 2021

Agenda

1. Conceptualization and Operationalization

2. Measurement

Agenda

1. Conceptualization and Operationalization

2. Measurement

Conceptualization

- Our hypotheses make predictions about the concepts we think are important
- But most concepts worth studying are latent—they cannot be directly observed
  - Justice, power, ideology, legitimacy, democracy, inequality, etc.
- Conceptualization is the process of translating an abstract idea into a specific (i.e., concrete and measurable) variable
- Need to create a **clear** and **explicit** definition of the concept

Operationalization

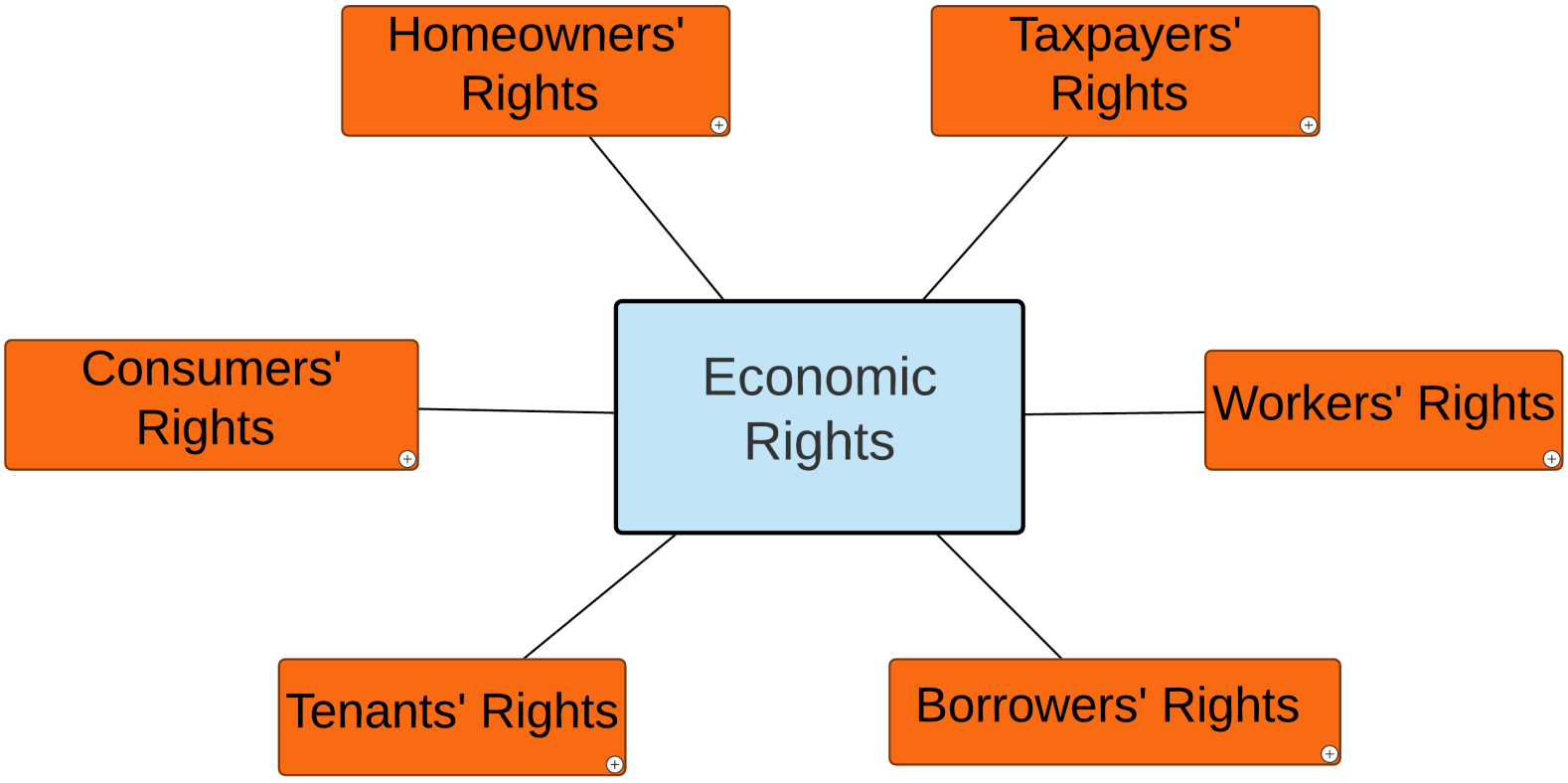
- Once we have a definition of our concept, we need to figure out how to measure it
- Defined as "[t]he process of specifying the operations that will indicate the value of cases on a variable"
- Variety of ways to operationalize a variable in social research
  - Surveys (single answer, index, scales)
  - Observational (available) data
  - Direct versus indirect observation

Operationalization

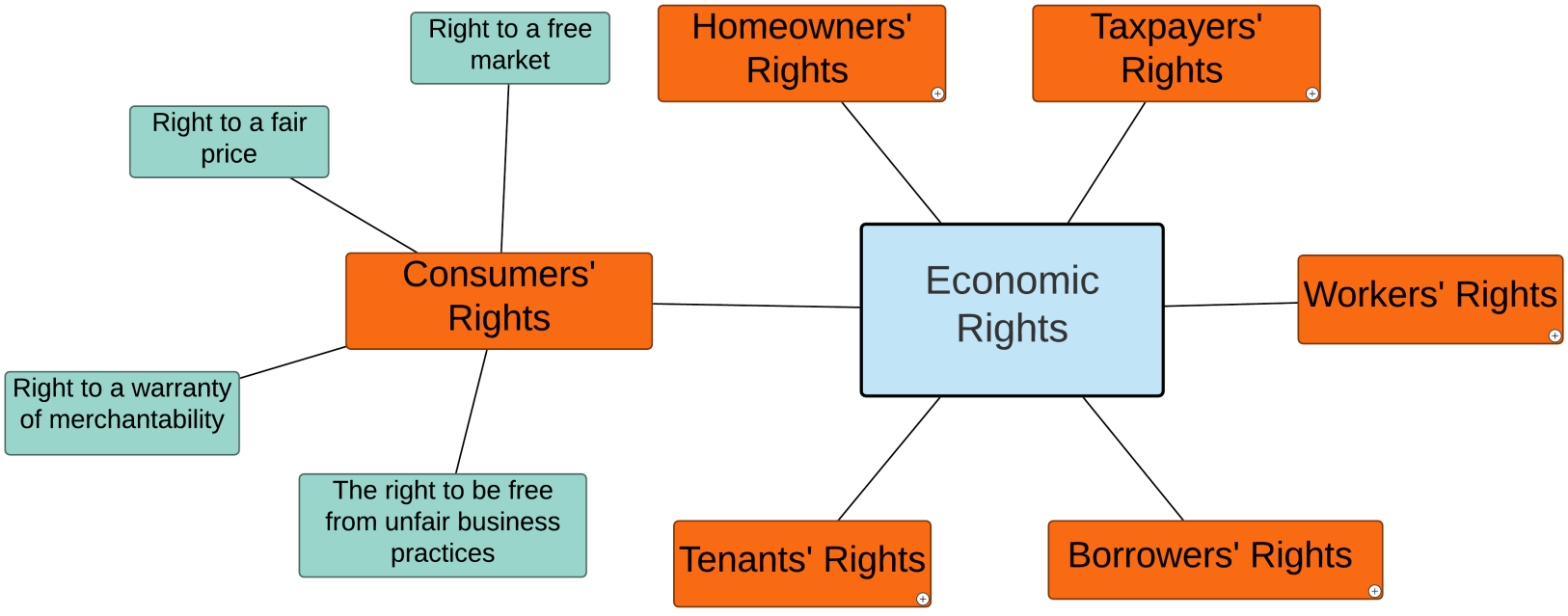
- We will operationalize our concepts with text-based measures
  - Subject-specific dictionaries
  - Word frequency
  - Sentiment
  - Topic modeling
  - Etc.

Our Latent Concept: Economic Rights

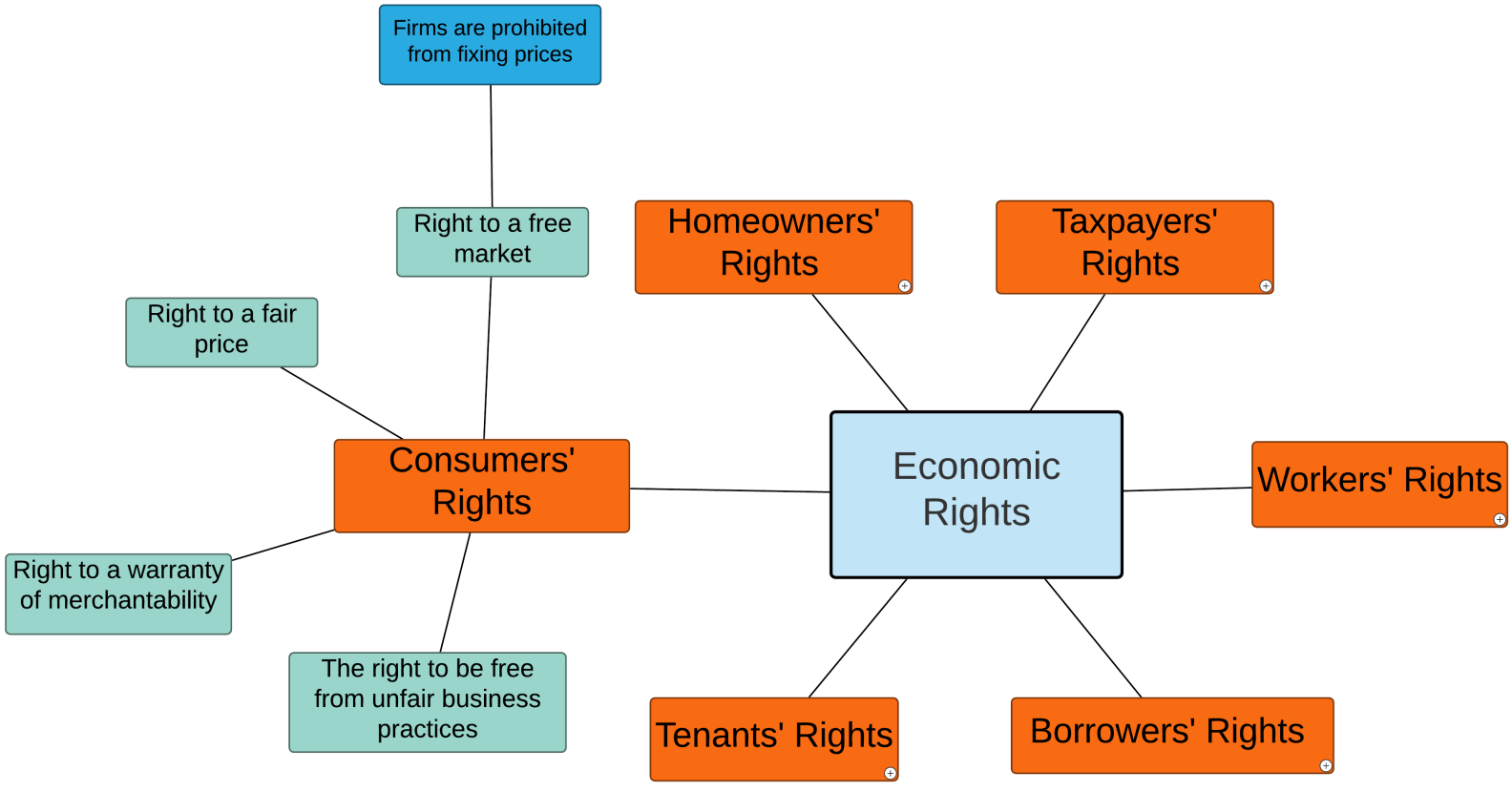
Defining Our Concept



Defining Our Concept

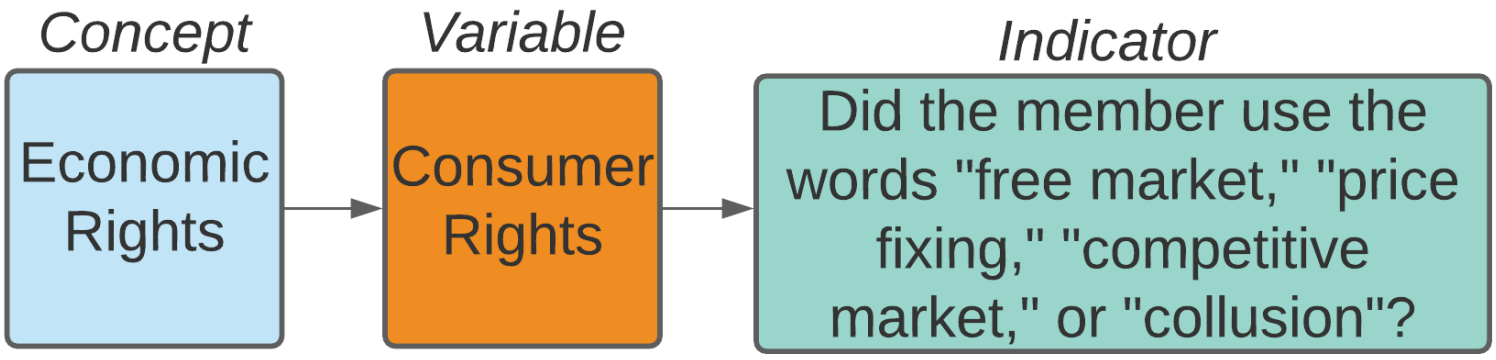


Choosing an Indicator



Operationalizing Speech

Our indicators will be focused on the *words* uttered by members of Congress



Agenda

1. Conceptualization and Operationalization

2. Measurement

Levels of Measurement

- We can design a variety of measures to capture members' discussion of our conceptualization of economic rights
- Measures may be stand-alone, additive, or set-theoretic (i.e., necessary and sufficient conditions)

Levels of Measurement

Categorical	Numeric
NOMINAL (unordered categories): Categorizing remarks as mentioning "free market" or "competition" or not	**RATIO (interval data with a true zero):** Number of words uttered by a member of Congress during a single speech
ORDINAL (ordered categories but distance between the units is not meaningful): Categorizing a remarks as being "strongly," "moderately," or "not at all" concerned with consumer rights	**INTERVAL (numeric data where distance between the units is meaningful):** Number of words describing workers' rights minus number of words describing consumer rights

Reliability

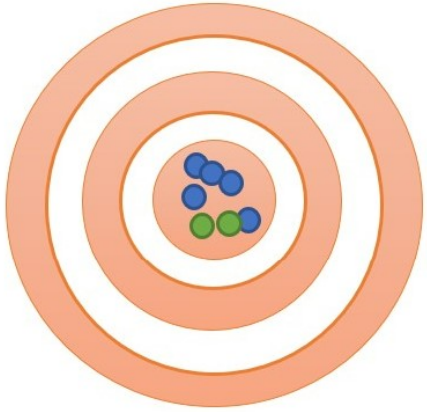
If you measured a concept over and over again (or it was measured by two different people), would you receive the same result?

- Analogous to the idea of precision
- EXAMPLE: If you asked a member of Congress how often they talked about economic competition or antitrust law, the answers you recieve would be tained by poor memory or self-censoring

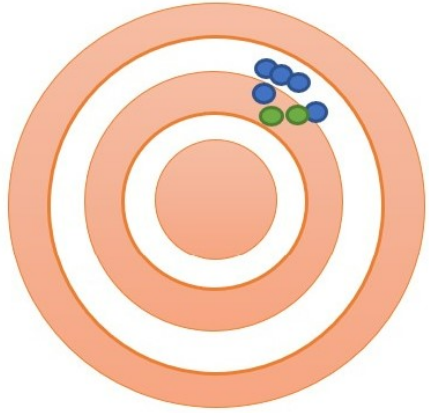
Validity

- *Construct validity*: Does the measure actually capture the concept?
  - EXAMPLE: If your word list contains the words "freedom," "economic," "liberty," and "rights," will you actually be capturing your conceptualization of economic rights or something else?
- *Measurement validity*: Does the measure actually capture what it claims to measure?
  - EXAMPLE: Does the number of minutes talking about economic rights on the floor *actually* capture how much a member cares about the issue, or are you capturing a different variable, like seniority?

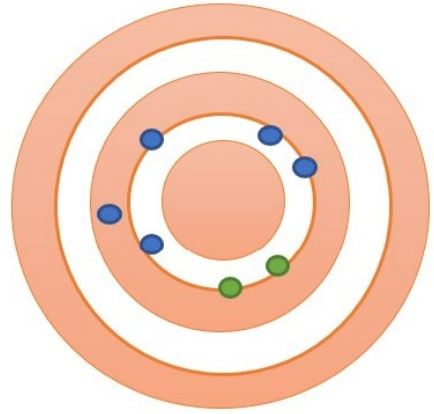
Comparing Reliability and Validity



Valid, reliable



Invalid, reliable



Valid, unreliable