Lecture 2: Sampling and measurement

Jacob M. Montgomery

Quantitative Political Methodology

Lecture 2

Class business

Getting slides

Last time:

`https://jmontgomery.github.io/PS363Slides/01Introduction.pdf`

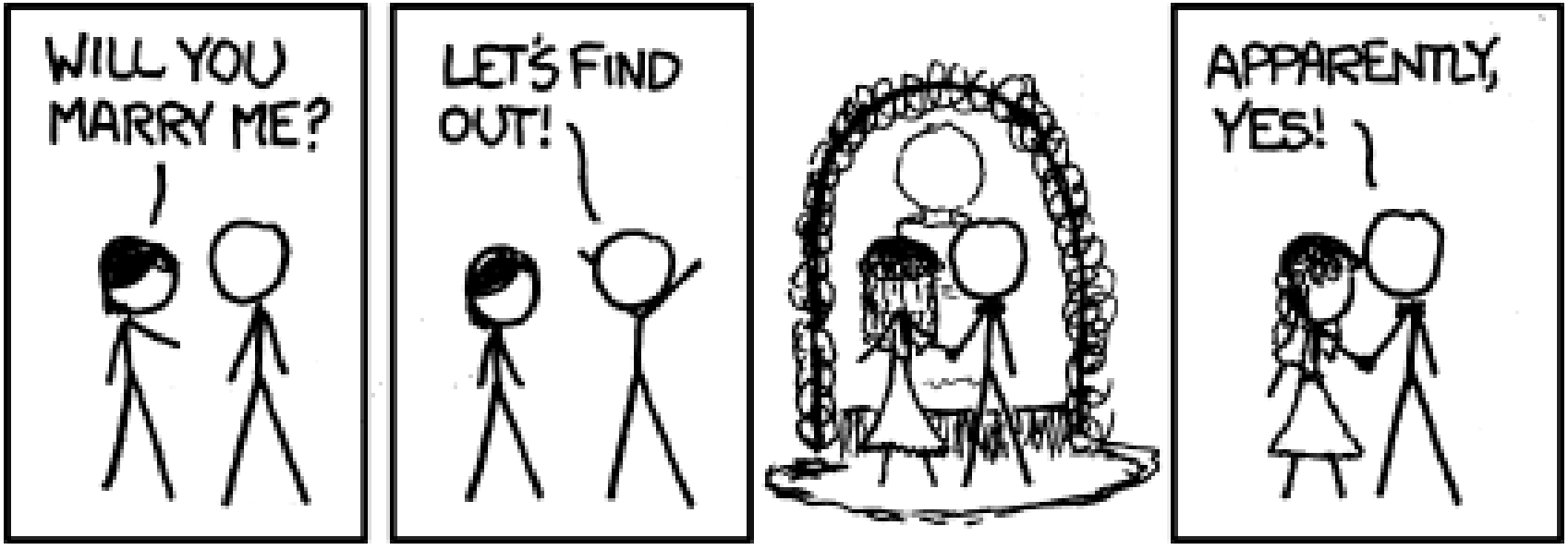
This time:

`https://jmontgomery.github.io/PS363Slides/02SamplingMeasurement.pdf`

Facebook and survey

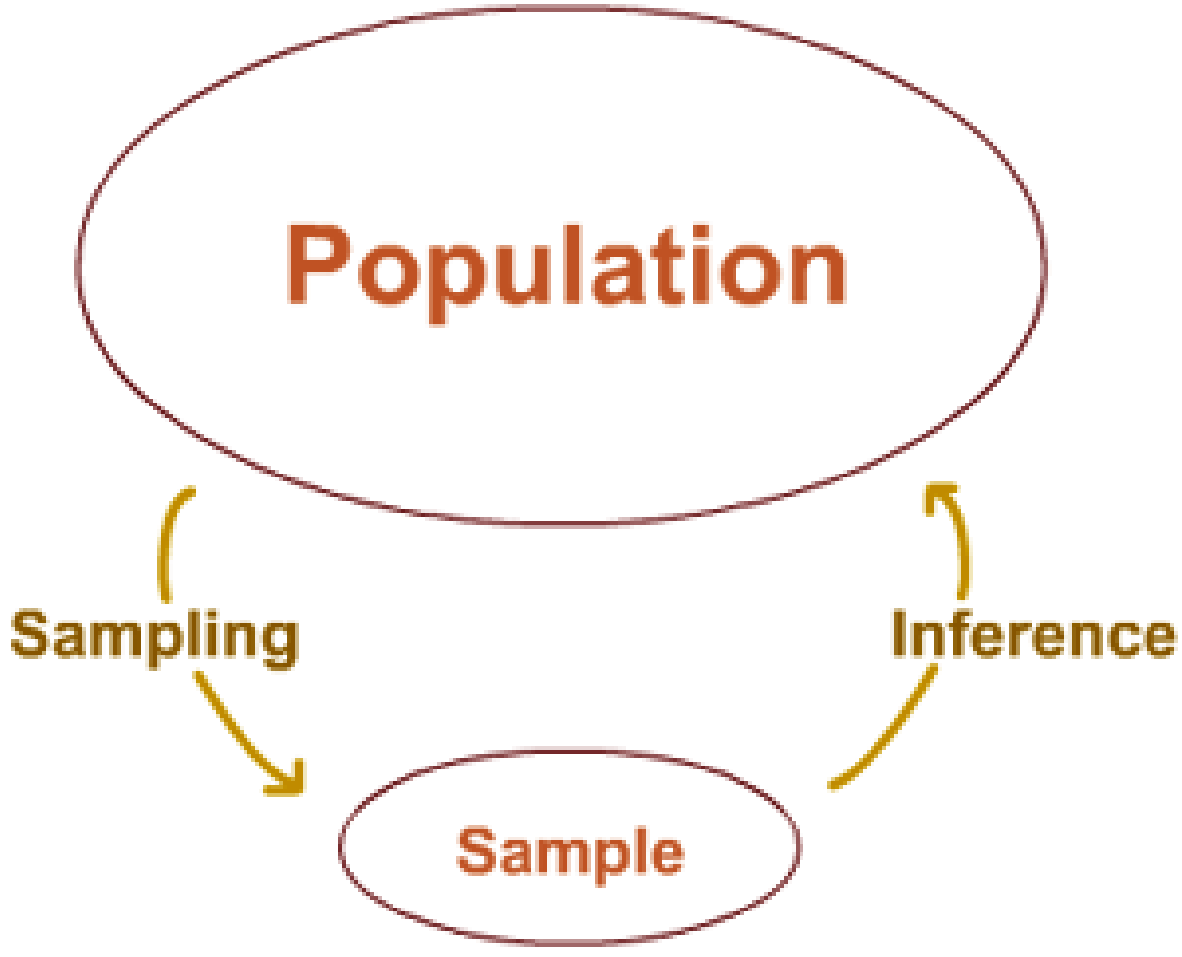
- Sign up for our Facebook group:
  `https://www.facebook.com/groups/1071702902960687/`
- Take the class survey! Can't assign teams until you all do.

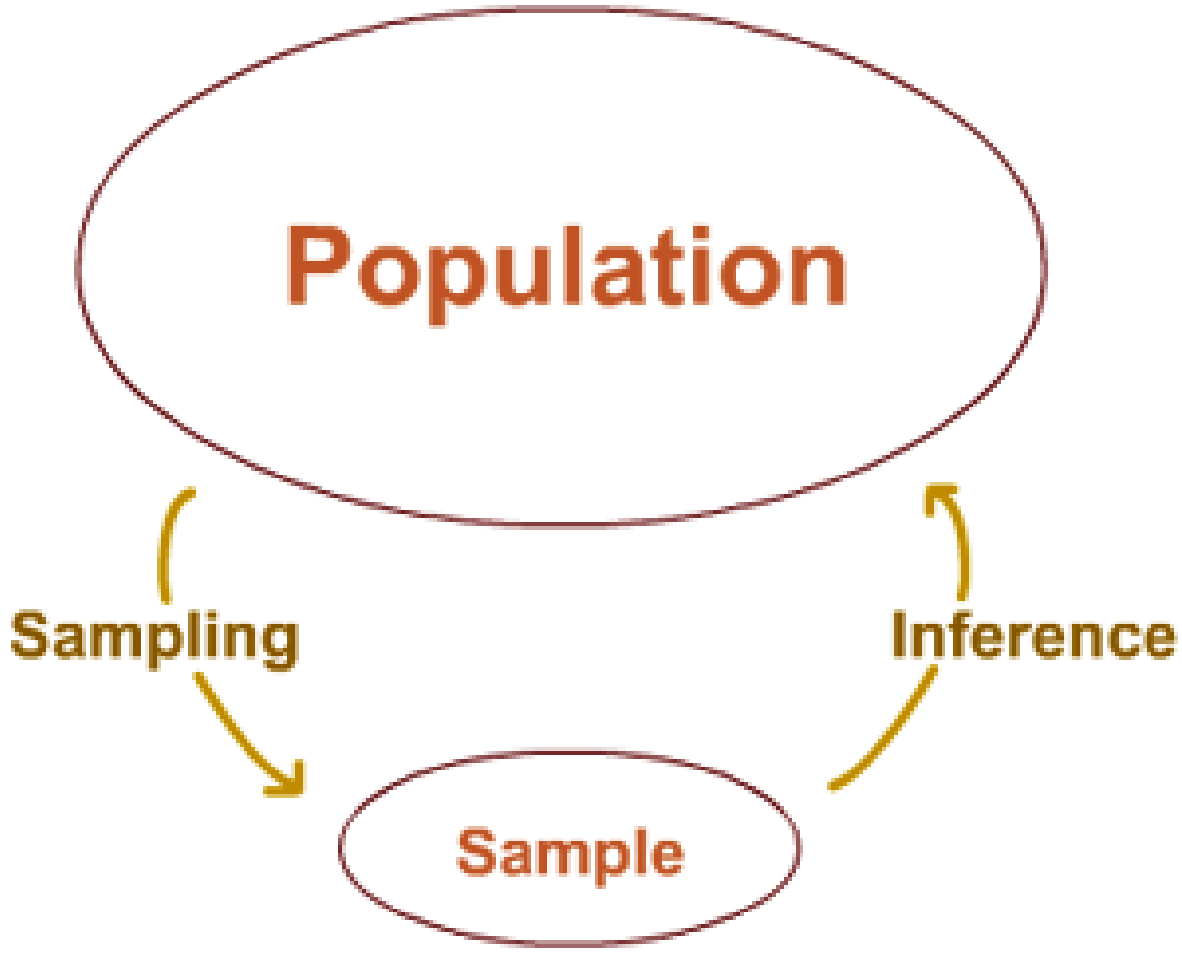
Statistical methods, measures, sampling, and error



Topics for today

- A (very) broad view of statistical methods in science
- Brief lecture and group exercise on measures, sampling, and error.


Population
Sampling
Inference
Sample



Example: What percent of this class approves of President Trump?

Using that information, what is the percent of WashU undergraduates that approve?

Some big concepts

How it all works:

- Populations

Some big concepts

How it all works:

- Populations have parameters

Some big concepts

How it all works:

- Populations have parameters (often denoted μ)
  - Example?

Some big concepts

How it all works:

- Populations have parameters (often denoted μ)
  - Example?
  - Why can't we know it?
- Choose a sample (randomly or otherwise)

Some big concepts

How it all works:

- Populations have parameters (often denoted μ)
  - Example?
  - Why can't we know it?
- Choose a sample (randomly or otherwise)
- Gather data from this sample (e.g., ?)

Some big concepts

How it all works:

- Populations have parameters (often denoted μ)
  - Example?
  - Why can't we know it?
- Choose a sample (randomly or otherwise)
- Gather data from this sample (e.g., ?)
- Calculate statistics

Some big concepts

How it all works:

- Populations have parameters (often denoted μ)
  - Example?
  - Why can't we know it?
- Choose a sample (randomly or otherwise)
- Gather data from this sample (e.g., ?)
- Calculate statistics (e.g., $\bar{x} = \frac{\sum x_i}{n}$)

Some big concepts

How it all works:

- Populations have parameters (often denoted μ)
  - Example?
  - Why can't we know it?
- Choose a sample (randomly or otherwise)
- Gather data from this sample (e.g., ?)
- Calculate statistics (e.g., $\bar{x} = \frac{\sum x_i}{n}$)
  - Is this descriptive or inferential?

Some big concepts

How it all works:

- Populations have parameters (often denoted μ)
  - Example?
  - Why can't we know it?
- Choose a sample (randomly or otherwise)
- Gather data from this sample (e.g., ?)
- Calculate statistics (e.g., $\bar{x} = \frac{\sum x_i}{n}$)
  - Is this descriptive or inferential?
- Use statistics to make a statement about the real world

Some big concepts

How it all works:

- Populations have parameters (often denoted μ)
  - Example?
  - Why can't we know it?
- Choose a sample (randomly or otherwise)
- Gather data from this sample (e.g., ?)
- Calculate statistics (e.g., $\bar{x} = \frac{\sum x_i}{n}$)
  - Is this descriptive or inferential?
- Use statistics to make a statement about the real world
  - Is this descriptive or inferential?

Sampling

Key concepts:

- Sampling methods
- Sampling bias

How to sample: Simple random sample (SRS)

- All individuals in population have equal chance of being sampled $(1/n)$

How to sample: Simple random sample (SRS)

- All individuals in population have equal chance of being sampled $(1/n)$
- Sampling error decreases with sample size

How to sample: Simple random sample (SRS)

- All individuals in population have equal chance of being sampled $(1/n)$
- Sampling error decreases with sample size
- Often difficult/impossible to achieve in practice

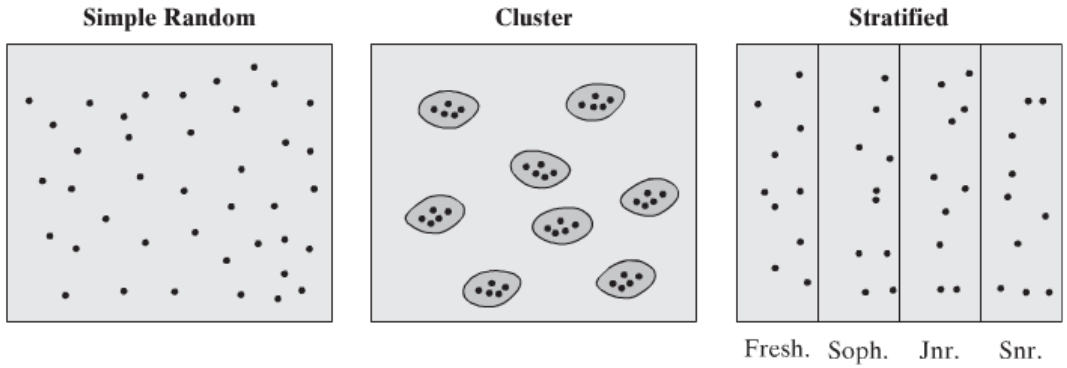
How to sample: Simple random sample (SRS)

- All individuals in population have equal chance of being sampled $(1/n)$
- Sampling error decreases with sample size
- Often difficult/impossible to achieve in practice

"If you don't believe in random sampling, the next time you have a blood test, tell the doctor to take it all."

Building on the SRS

- **Stratified** random sample: SRS *within* pre-specified groups
- **Cluster** sample: survey within SRS of population groupings (cities, blocks, etc.)
- **Multistage** sample—SRS at multiple levels:
  - Example: Randomly selecting states, then randomly selecting counties, then randomly selecting households



Sampling

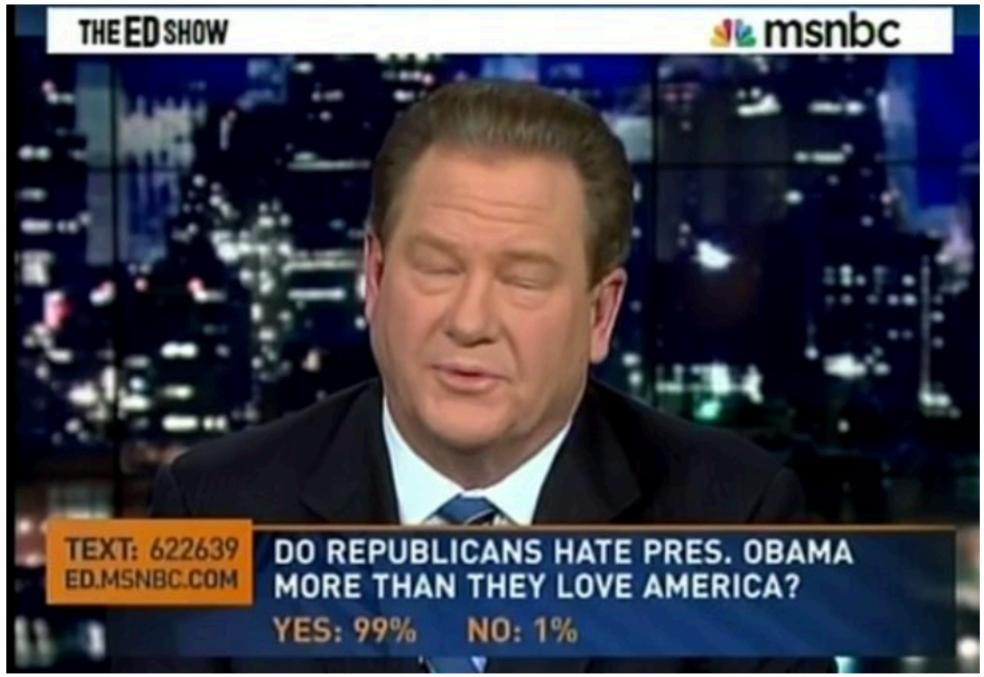
- Sample size (Discuss later)
- Sampling bias
  - Convenience vs. random samples
  - Undercoverage
  - Nonresponse
- Response bias
- Wording effects

Convenience samples

Maureen Dowd (2/9/03):

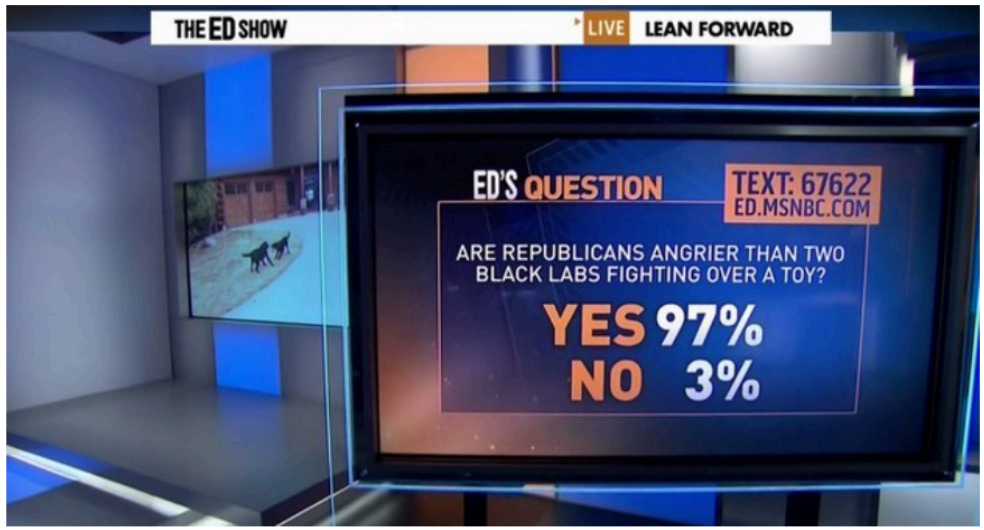
> *Most Americans are willing to give Mr. Bush his war even though they are dubious that it will curb terror.*
>
> *A CNN online poll shows that 82 percent think going to war with Iraq will provoke another attack on the U.S., as opposed to 13 percent who think it will prevent one.*

1.



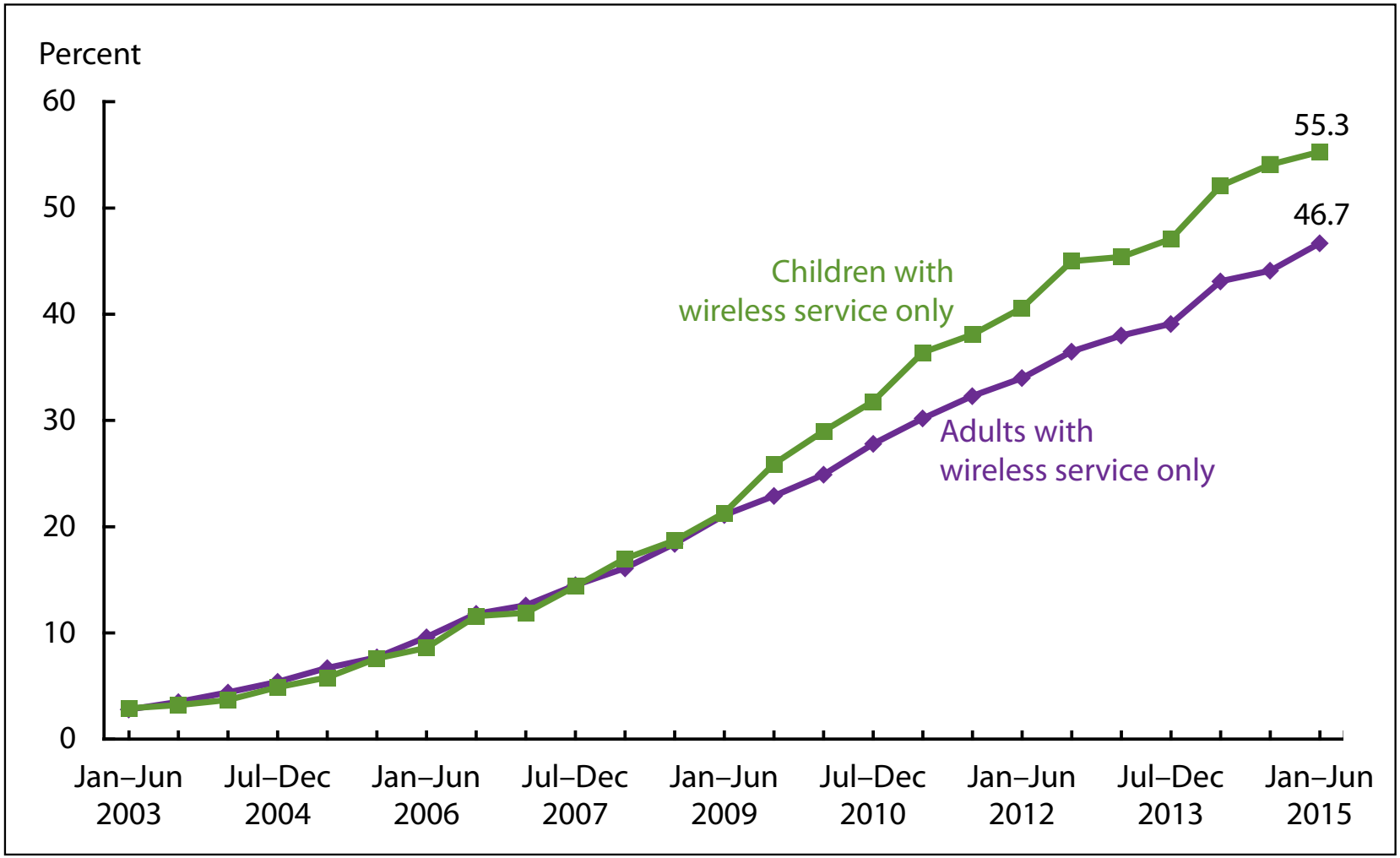
Ed Show/MSNBC

2.



Ed Show/MSNBC

Figure. Percentages of adults and children living in households with only wireless telephone service: United States, 2003–2015



NOTE: Adults are aged 18 and over; children are under age 18.
DATA SOURCE: CDC/NCHS, National Health Interview Survey.

Cellphone-only respondents are demographically distinct

Unweighted demographic profiles of respondents with only a cellphone versus all other respondents

	Cellphone only	All other respondents
	%	%
Male	58*	49
Female	42*	51
White, non-Hispanic	62*	76
Black, non-Hispanic	10	9
Hispanic	18*	8
18-29	28*	8
30-49	37*	21
50-64	24*	33
65+	10*	36
High school grad or less	31*	27
Some college/Associate degree	29	29
Bachelor's degree or more	39*	44
Urban	43*	30
Suburban	40*	52
Rural	16*	18
Family income		
Less than $30,000	32*	22
$30,000-$74,999	33*	31
$75,000 or more	29*	37

Surveys Face Growing Difficulty Reaching, Persuading Potential Respondents

	1997	2000	2003	2006	2009	2012
	%	%	%	%	%	%
Contact rate (percent of households in which an adult was reached)	90	77	79	73	72	62
Cooperation rate (percent of households contacted that yielded an interview)	43	40	34	31	21	14
Response rate (percent of households sampled that yielded an interview)	36	28	25	21	15	9

PEW RESEARCH CENTER 2012 Methodology Study. Rates computed according to American Association for Public Opinion Research (AAPOR) standard definitions for CON2, COOP3 and RR3. Rates are typical for surveys conducted in each year.

Response bias

"Recalling the last ten traffic lights you drove through, how many of them were red when you entered the intersections?"

Response bias

Q: “Did you vote in the elections this November?”

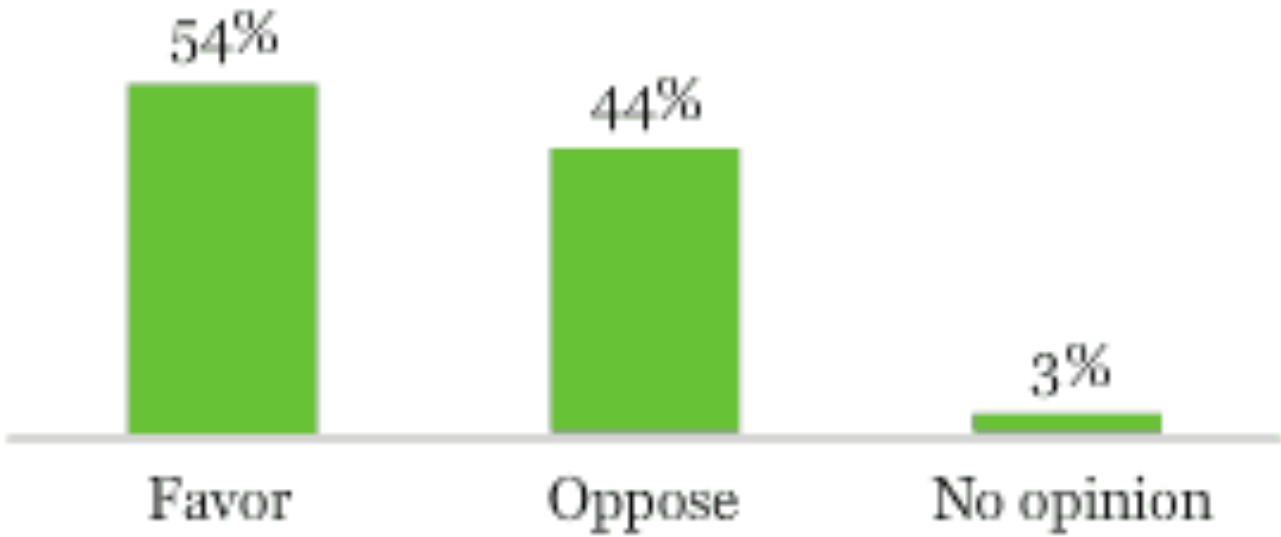
Response bias

Q: "Did you vote in the elections this November?"

Q: "In talking to people about elections, we often find that a lot of people were not able to vote because they weren't registered, they were sick, or they just didn't have time. How about you–did you vote in the elections this November?"

Question wording

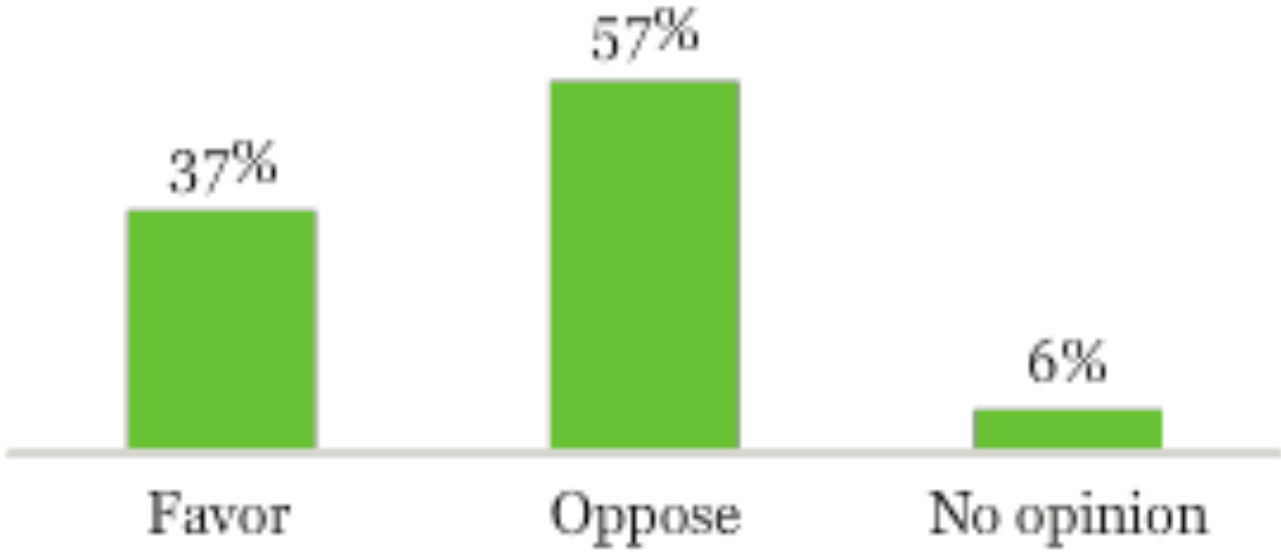
Do you favor or oppose the federal government temporarily taking over major U.S. banks in danger of failing in an attempt to stabilize them?



USA Today/Gallup, Feb. 20-22, 2009

Question wording: Nationalize

Do you favor or oppose the federal government temporarily nationalizing major U.S. banks in danger of failing in an attempt to stabilize them?



USA Today/Gallup, Feb. 20-22, 2009

Big Ideas

- Inferences to populations depend on having a representative sample
- Wording and context shape survey responses
- People are sometimes unwilling to share true feelings or beliefs