Survey Research and Design

Nonresponse and Survey Weighting

William Marble

September 28, 2023

The Goal of a Survey

- When we conduct a survey we're not interested in the answers just of the people we talk to
- Instead, want to use the survey to *generalize* to a larger population
- Raises the question: how similar is our sample to the population?

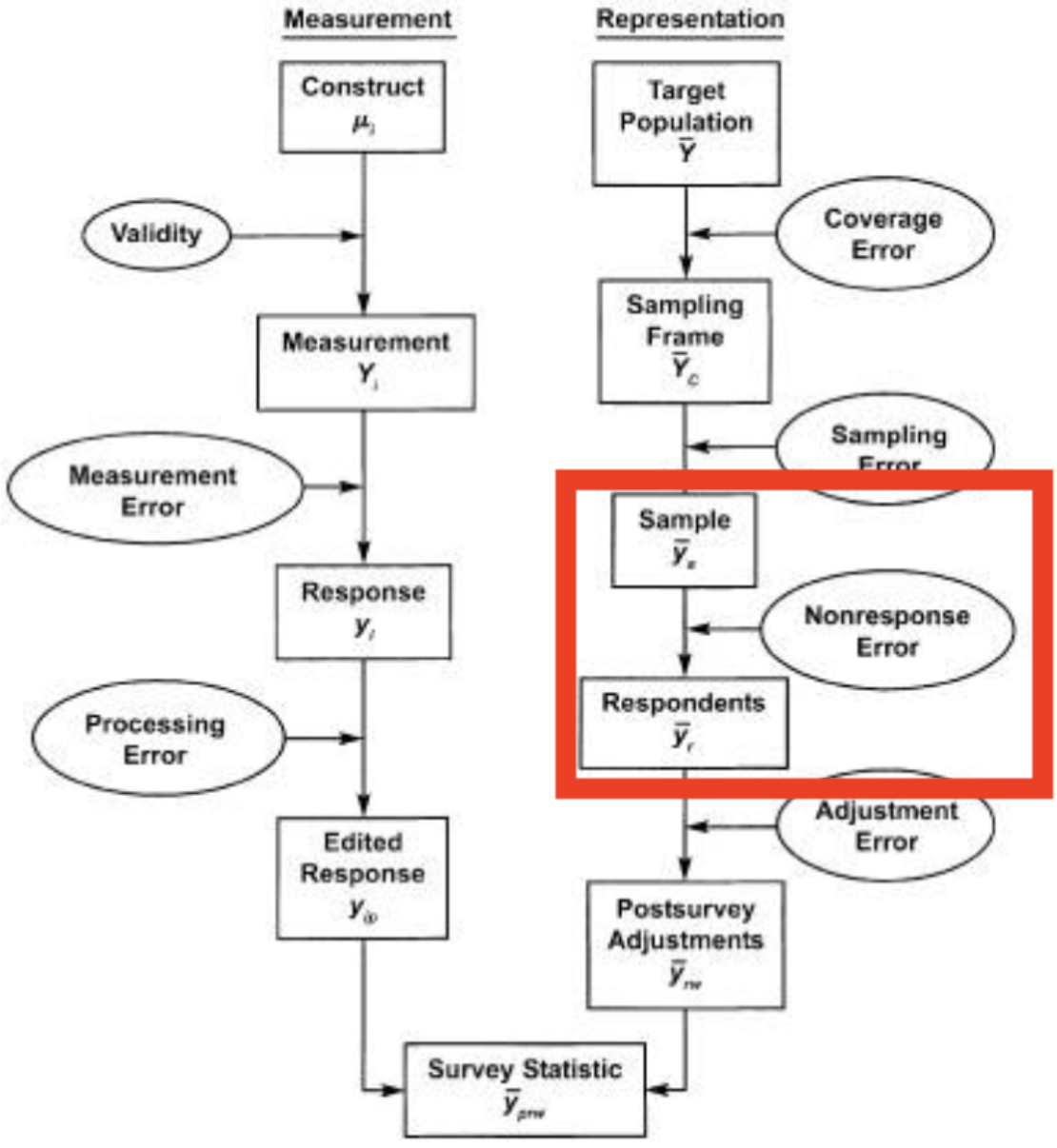


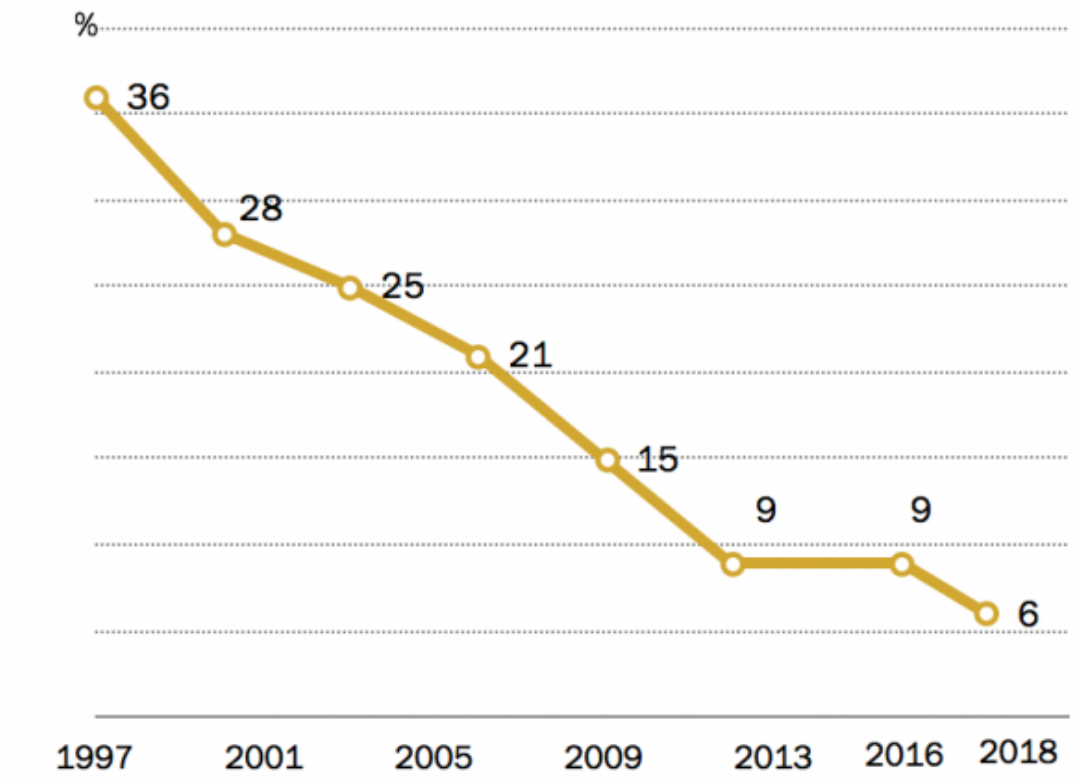
Figure 2.5 Survey life cycle from a quality perspective.

Differences Between Population and Sample

- There are many causes of discrepancy between our sample and the population
- Sampling error is pretty innocuous: it's random, so unrelated to the attributes we want to measure
- Coverage error and nonresponse error can be more pernicious: may be correlated with attributes of interest
- Requires careful analysis and relies on (often untestable) assumptions

After brief plateau, telephone survey response rates have fallen again

Response rate by year (%)



Note: Response rate is AAPOR RR3. Only landlines sampled 1997-2006. Rates are typical for surveys conducted in each year.
Source: Pew Research Center telephone surveys conducted 1997-2018.

PEW RESEARCH CENTER

Nonresponse Overview

Nonresponse is the failure to collect data from a unit we intend to include in the survey.

Nonresponse Overview

Nonresponse is the failure to collect data from a unit we intend to include in the survey.

Two forms of nonresponse:

- **Unit nonresponse** is failure to collect *any* data from a unit
- **Item nonresponse** is failure to collect *only some* data from a unit (i.e. they skip a question)

The Problem with Nonresponse

- The values given by our respondents may differ from the values in the population as a whole
- If this difference is systematic — occurs over many hypothetical "runs" of the survey — then it is called **nonresponse bias**

Reasoning About Nonresponse

We need some notation...

Y is the attribute we are trying to measure

- E.g., partisanship, attitudes, or income

Reasoning About Nonresponse

We need some notation...

Y is the attribute we are trying to measure

- E.g., partisanship, attitudes, or income

P is the response propensity

- The probability that an individual will respond to the survey

Reasoning About Nonresponse

We need some notation...

Y is the attribute we are trying to measure

- E.g., partisanship, attitudes, or income

P is the response propensity

- The probability that an individual will respond to the survey

Z and X are auxiliary variables

- E.g., age, race, or education

Three Models of Nonresponse

How is nonresponse related to the attribute of interest, Y?

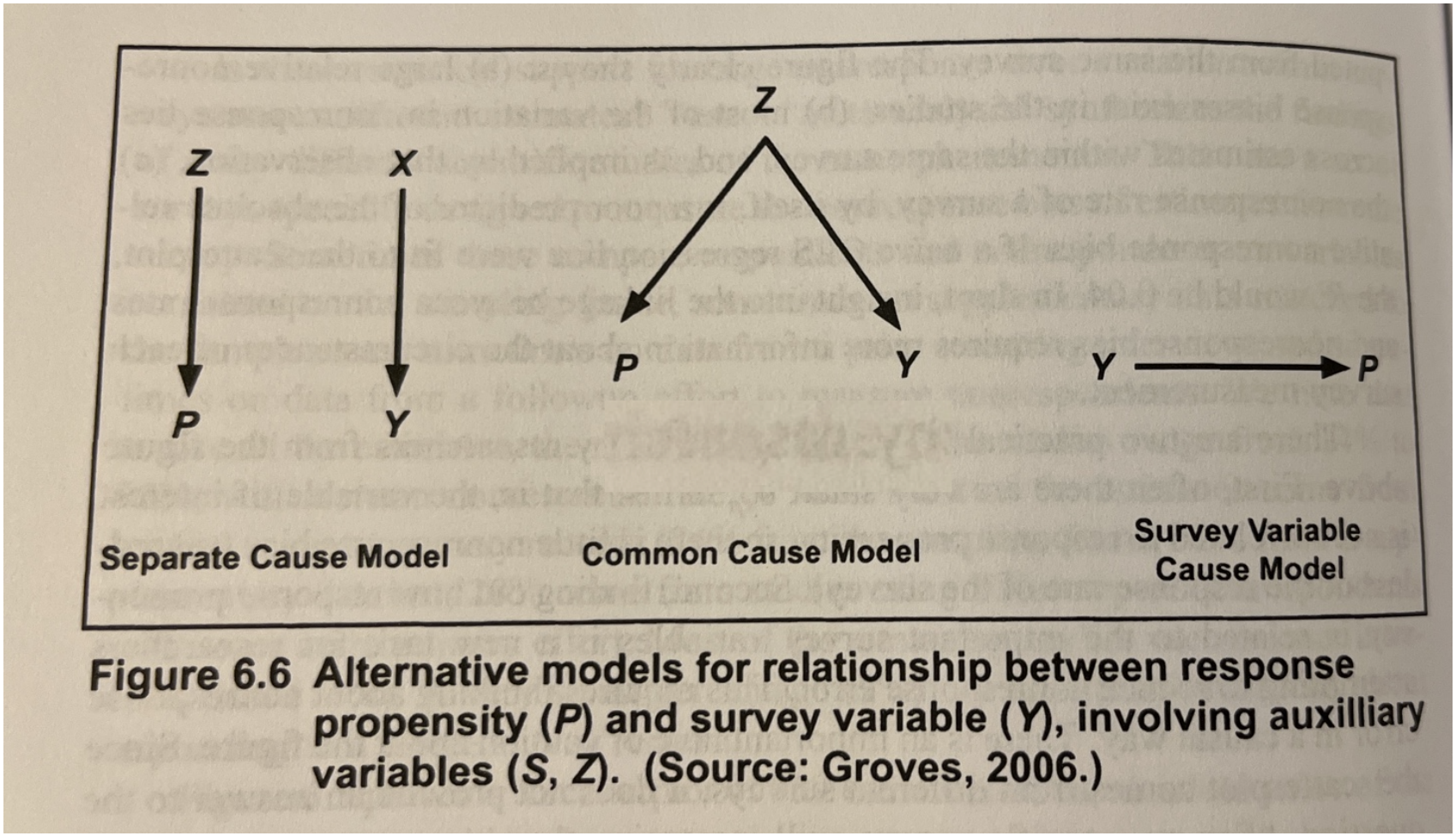


Figure 6.6 Alternative models for relationship between response propensity (P) and survey variable (Y), involving auxilliary variables (S, Z). (Source: Groves, 2006.)

Models of Nonresponse

Separate Cause Model: $Z \to P$, $X \to Y$

- Auxiliary variable Z affects response propensity P but not attribute of interest Y
- Auxiliary variable X affects attribute Y, and X is uncorrelated with Z
- In this case, nonresponse is **ignorable**
- Nonresponse does not cause bias in our estimate of Y

Models of Nonresponse

Separate Cause Model: $Z \rightarrow P$, $X \rightarrow Y$

- Auxiliary variable Z affects response propensity P but not attribute of interest Y
- Auxiliary variable X affects attribute Y, and X is uncorrelated with Z
- In this case, nonresponse is **ignorable**
- Nonresponse does not cause bias in our estimate of Y

Example: internet use increases response propensity, but isn't correlated with partisanship. We don't need to adjust for nonresponse.

Models of Nonresponse

Common Cause Model: $Z \to P$, $Z \to Y$

- Auxiliary variable Z affects response propensity P **and** affects Y
- If we can "adjust" for Z, then we can get a good estimate of Y
- We might know the distribution of Z in the population that we can use to adjust our survey estimates
- In this case, nonresponse is **conditionally ignorable**: if we have information about Z, then nonresponse does not cause problems
- Intuition: among groups defined by Z, respondents are representative of the population

Models of Nonresponse

Common Cause Model: $Z \rightarrow P$, $Z \rightarrow Y$

- Auxiliary variable Z affects response propensity P **and** affects Y
- If we can "adjust" for Z, then we can get a good estimate of Y
- We might know the distribution of Z in the population that we can use to adjust our survey estimates
- In this case, nonresponse is **conditionally ignorable**: if we have information about Z, then nonresponse does not cause problems
- Intuition: among groups defined by Z, respondents are representative of the population

Example: Women are more likely to take surveys, and they are more Democratic. We can adjust for overrepresentation of women in our survey.

Models of Nonresponse

Survey Variable Cause Model: $Y \rightarrow P$

- The value of the attribute of interest Y directly affects response propensity P
- Nonresponse is **nonignorable**
- Because we can't observe Y among non-responders, this presents a problem with no straightforward solution

Models of Nonresponse

Survey Variable Cause Model: $Y \rightarrow P$

- The value of the attribute of interest Y directly affects response propensity P
- Nonresponse is **nonignorable**
- Because we can't observe Y among non-responders, this presents a problem with no straightforward solution

Example: People who support Trump are less likely to trust pollsters, so they respond at a lower rate (even after adjusting for demographics).

Reasoning about Nonresponse

- Response rates can be low; whether this is a problem depends on the cause of nonresponse

Reasoning about Nonresponse

- Response rates can be low; whether this is a problem depends on the cause of nonresponse
- Always think through likely reasons for nonresponse
- Are those reasons correlated with Y?
  - If not, nonresponse isn't a problem
  - If so, it depends. If it's a common-cause situation, we might be able to adjust for the common cause of P and Y

Reasoning about Nonresponse

- Response rates can be low; whether this is a problem depends on the cause of nonresponse
- Always think through likely reasons for nonresponse
- Are those reasons correlated with Y?
  - If not, nonresponse isn't a problem
  - If so, it depends. If it's a common-cause situation, we might be able to adjust for the common cause of P and Y
- These are substantive questions that the data cannot answer! We need to rely on **assumptions**
- Ideally, we make the assumptions as plausible as possible

Reasons for Nonresponse

Four sets of factors:

1. Social environmental factors
2. Person-level factors
3. Interviewer factors
4. Survey design factors

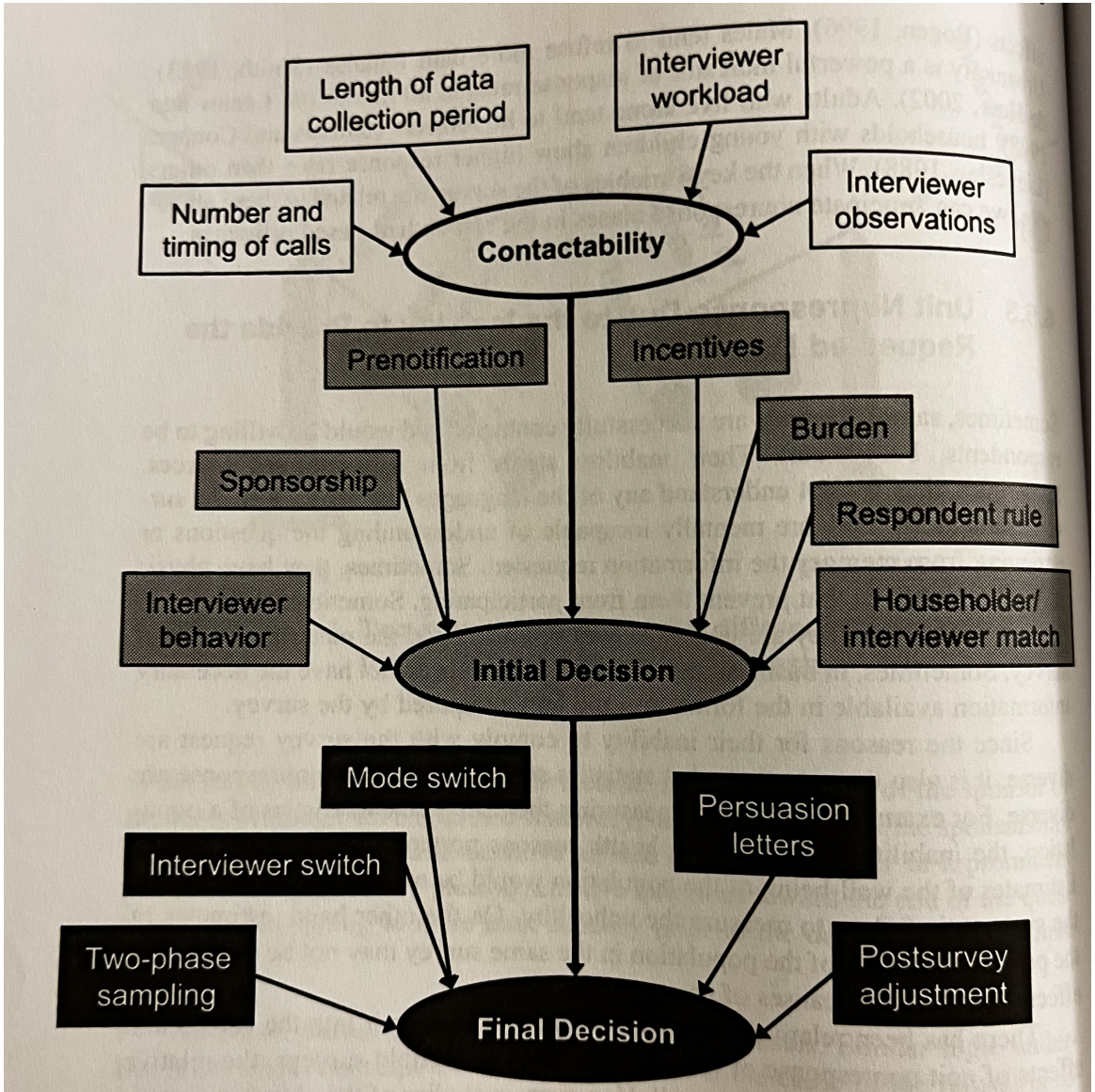


Figure 6.12 Tools for reducing unit nonresponse rates.

Survey Weighting: Overview

- Survey weighting allows us to correct for imbalances between our sample and the population on **observable attributes**
- Suppose our survey respondents are 65% women and 35% men
  - We know that the acutal distribution in the population is (around) 50-50
  - We can develop weights to make our sample "look" like it has a 50-50 gender split
  - We will *upweight* men and *downweight* women
- Whether this gives us better estimates for attributes we want to study depends on which model of nonresponse is at play

What Proportion of the Population Are Democrats?

Hypothetical Data

Gender	Democrat	Republican	Pop. Proportion	Sample Proportion
Men	40%	60%	50%	35%
Women	60%	40%	50%	65%

What Proportion of the Population Are Democrats?

Hypothetical Data

Gender	Democrat	Republican	Pop. Proportion	Sample Proportion
Men	40%	60%	50%	35%
Women	60%	40%	50%	65%

What is the proportion of Democrats in the sample?

What Proportion of the Population Are Democrats?

Hypothetical Data

Gender	Democrat	Republican	Pop. Proportion	Sample Proportion
Men	40%	60%	50%	35%
Women	60%	40%	50%	65%

What is the proportion of Democrats in the sample?

$$\begin{aligned} &\%\ \text{Dem. Men} \times \text{Sample Prop. Men} \quad + \quad \%\ \text{Dem. Women} \times \text{Sample Prop. Women} \\ &40\% \times 35\% \quad + \quad 60\% \times 65\% = 53\% \end{aligned}$$

What is the proportion of Democrats in the population?

What Proportion of the Population Are Democrats?

Hypothetical Data

Gender	Democrat	Republican	Pop. Proportion	Sample Proportion
Men	40%	60%	50%	35%
Women	60%	40%	50%	65%

What is the proportion of Democrats in the sample?

$$\%\ \text{Dem. Men} \times \text{Sample Prop. Men} \quad + \quad \%\ \text{Dem. Women} \times \text{Sample Prop. Women}$$

$$40\% \times 35\% \quad + \quad 60\% \times 65\% = 53\%$$

What is the proportion of Democrats in the population?

$$\%\ \text{Dem. Men} \times \text{Population Prop. Men} \quad + \quad \%\ \text{Dem. Women} \times \text{Population Prop. Women}$$

$$40\% \times 50\% \quad + \quad 60\% \times 50\% = 50\%$$

What Proportion of the Population Are Democrats?

Hypothetical Data

Gender	Democrat	Republican	Pop. Proportion	Sample Proportion
Men	40%	60%	50%	35%
Women	60%	40%	50%	65%

To get the right answer we can reweight:

$$\%\ \text{Dem. Men} \times \text{Sample Prop. Men} \times \overbrace{\frac{\text{Population Prop. Men}}{\text{Sample Prop. Men}}}^{\text{weight for men}}$$

$$+\ \%\ \text{Dem. Women} \times \text{Sample Prop. Women} \times \underbrace{\frac{\text{Population Prop. Women}}{\text{Sample Prop. Women}}}_{\text{weight for women}}$$

$$= 40\% \times 35\% \times \frac{50\%}{35\%} + 60\% \times 65\% \times \frac{50\%}{65\%} = 50\%$$

What Proportion of the Population Are Democrats?

Hypothetical Data

Gender	Democrat	Republican	Pop. Proportion	Sample Proportion
Men	40%	60%	50%	35%
Women	60%	40%	50%	65%

To get the right answer we can reweight:

$$\%\ \text{Dem. Men} \times \text{Sample Prop. Men} \times \overbrace{\frac{\text{Population Prop. Men}}{\text{Sample Prop. Men}}}^{\text{weight for men}}$$

$$+\ \%\ \text{Dem. Women} \times \text{Sample Prop. Women} \times \underbrace{\frac{\text{Population Prop. Women}}{\text{Sample Prop. Women}}}_{\text{weight for women}}$$

$$= 40\% \times 35\% \times \frac{50\%}{35\%} + 60\% \times 65\% \times \frac{50\%}{65\%} = 50\%$$

When Does Weighting Work?

- The intuition of weighting is to make our sample "look like" the population
- But there are many different variables we could consider when assessing similarity between the sample and population
  - online sample probably use Twitter more than the population
  - people who answer phone surveys use the phone more
  - survey respondents might be more extraverted
  - maybe survey respondents are more likely to be pet owners than the population (??)
- Which differences should we adjust for, and which can we safely ignore?

What Do We Need to Generate Weights?

1. **Assumption**: the attribute of interest Y is ignorable conditional on the weighting variables Z
2. We have measured Z for units in our survey
3. We know what the population distribution of Z is

Population Distribution of Z

We need to measure Z in our survey and in the population if we want to generate weights. **This requirement restricts the set of variables we can weight on.**

Population Distribution of Z

We need to measure Z in our survey and in the population if we want to generate weights. **This requirement restricts the set of variables we can weight on.**

- Z may be a vector of variables
- For most sociodemographic variables, we have good population targets from the Census
- Common to weight on these variables because many political attitudes are also related to these variables (i.e. the conditional ignorability assumption might be plausible)

Population Distribution of Z

Ideally we would know the full **joint distribution** of Z:

- E.g., we would know what proportion of Americans fall into each category of age $\times$ race $\times$ sex $\times$ education
- The we could reweight each "cell"
- This weighting method is called **poststratification**
- Problem: sometimes we don't know the joint distribution, or there are empty cells in our survey
  - why would this cause problems?

Population Distribution of Z

Often we'll just weight to the **marginal distribution** of each variable

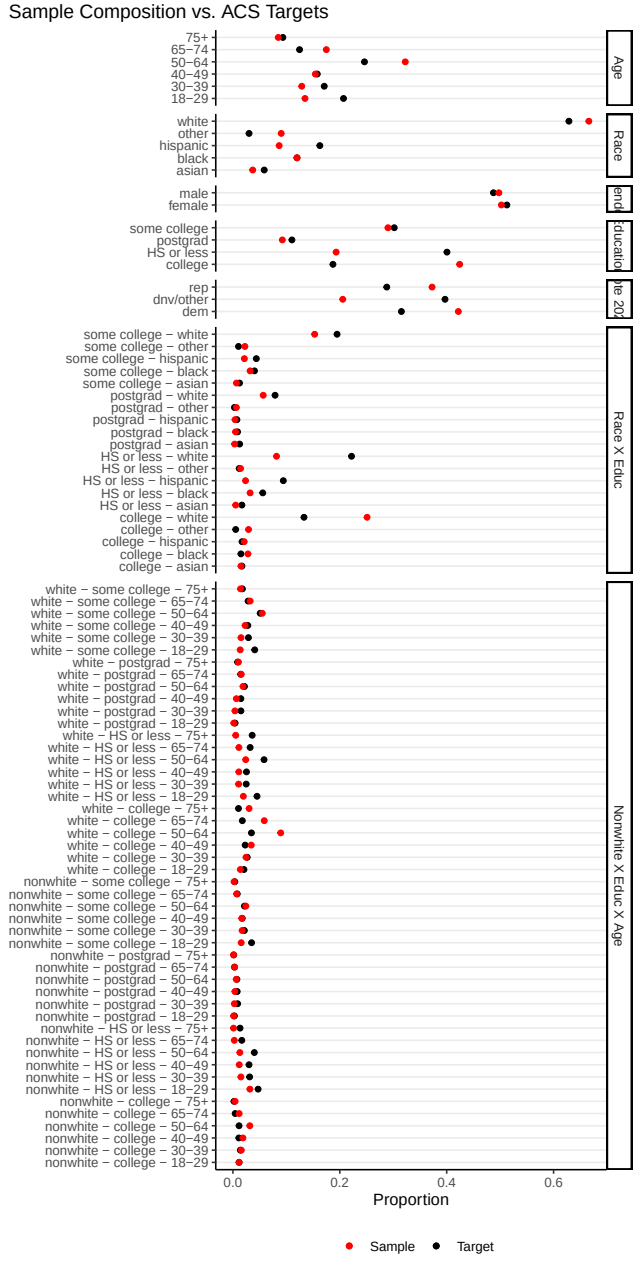
- E.g., we know how many people are each age, we know how many people are each race, each sex, and have each education level
- But we don't know the "crosstab"
- In this case, we can use "raking" to match the population margins without knowledge of the full joint distribution
- **Assumption**: interactions aren't all that important
  - E.g., differences in opinion across age groups are the same for men and women

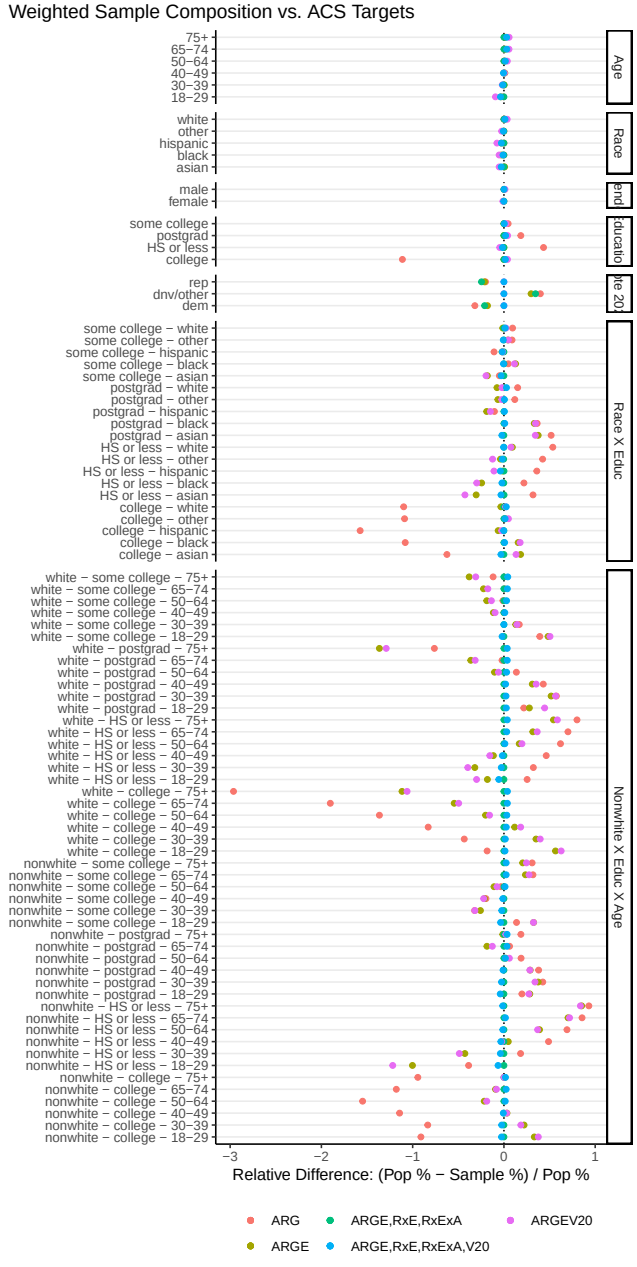
Weighting: General Process

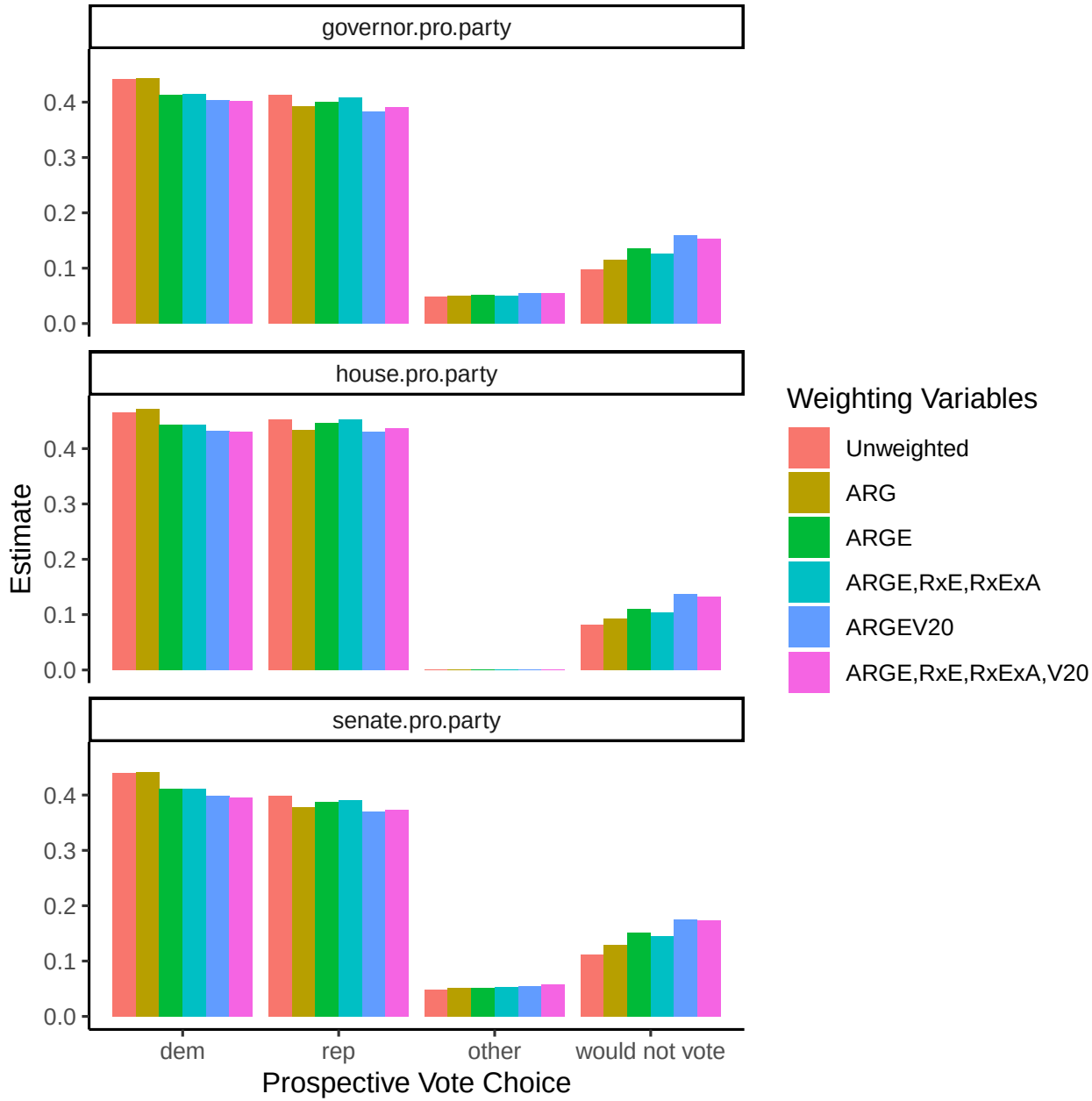
- Define the target population
- Determing the variables you want to use in weighting
- Gather data on the distribution of weighting variables in the population
- Conduct the survey (being sure to collect weighting variables)
- Use one of many weighting techniques to calculate weights for each person in the survey

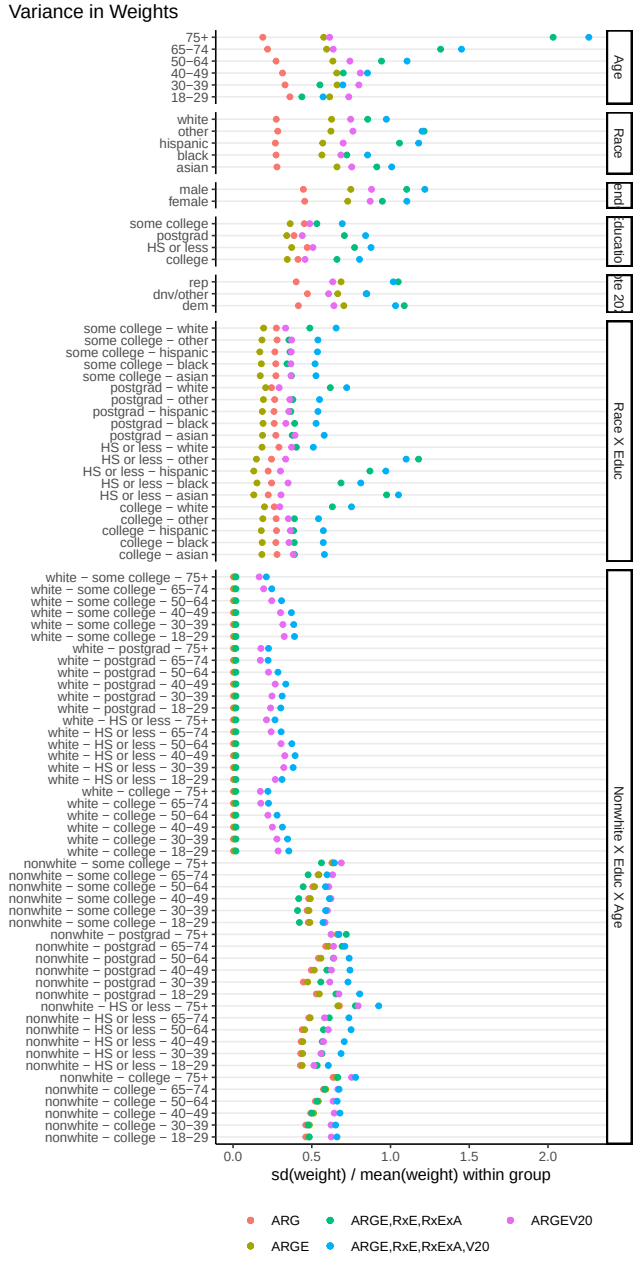
Acquiring Weighting Targets

- Demographics are most common
- Need to gather data on population distribution:
  - Census data finder website
  - NHGIS
  - Current Population Survey (March Supplement)
  - Voter files
  - Other government surveys


Sample Composition vs. ACS Targets
75+
65−74
50−64
40−49
30−39
18−29
Age
white
other
hispanic
black
asian
Race
male
female
end
some college
postgrad
HS or less
college
ducatio
rep
dnv/other
dem
te 20
some college − white
some college − other
some college − hispanic
some college − black
some college − asian
postgrad − white
postgrad − other
postgrad − hispanic
postgrad − black
postgrad − asian
HS or less − white
HS or less − other
HS or less − hispanic
HS or less − black
HS or less − asian
college − white
college − other
college − hispanic
college − black
college − asian
Race X Educ
white − some college − 75+
white − some college − 65−74
white − some college − 50−64
white − some college − 40−49
white − some college − 30−39
white − some college − 18−29
white − postgrad − 75+
white − postgrad − 65−74
white − postgrad − 50−64
white − postgrad − 40−49
white − postgrad − 30−39
white − postgrad − 18−29
white − HS or less − 75+
white − HS or less − 65−74
white − HS or less − 50−64
white − HS or less − 40−49
white − HS or less − 30−39
white − HS or less − 18−29
white − college − 75+
white − college − 65−74
white − college − 50−64
white − college − 40−49
white − college − 30−39
white − college − 18−29
nonwhite − some college − 75+
nonwhite − some college − 65−74
nonwhite − some college − 50−64
nonwhite − some college − 40−49
nonwhite − some college − 30−39
nonwhite − some college − 18−29
nonwhite − postgrad − 75+
nonwhite − postgrad − 65−74
nonwhite − postgrad − 50−64
nonwhite − postgrad − 40−49
nonwhite − postgrad − 30−39
nonwhite − postgrad − 18−29
nonwhite − HS or less − 75+
nonwhite − HS or less − 65−74
nonwhite − HS or less − 50−64
nonwhite − HS or less − 40−49
nonwhite − HS or less − 30−39
nonwhite − HS or less − 18−29
nonwhite − college − 75+
nonwhite − college − 65−74
nonwhite − college − 50−64
nonwhite − college − 40−49
nonwhite − college − 30−39
nonwhite − college − 18−29
Nonwhite X Educ X Age
0.0
0.2
0.4
0.6
Proportion
Sample
Target


Weighted Sample Composition vs. ACS Targets
75+
65−74
50−64
40−49
30−39
18−29
Age
white
other
hispanic
black
asian
Race
male
female
end
some college
postgrad
HS or less
college
ducatio
rep
dnv/other
dem
ie 20
some college − white
some college − other
some college − hispanic
some college − black
some college − asian
postgrad − white
postgrad − other
postgrad − hispanic
postgrad − black
postgrad − asian
HS or less − white
HS or less − other
HS or less − hispanic
HS or less − black
HS or less − asian
college − white
college − other
college − hispanic
college − black
college − asian
Race X Educ
white − some college − 75+
white − some college − 65−74
white − some college − 50−64
white − some college − 40−49
white − some college − 30−39
white − some college − 18−29
white − postgrad − 75+
white − postgrad − 65−74
white − postgrad − 50−64
white − postgrad − 40−49
white − postgrad − 30−39
white − postgrad − 18−29
white − HS or less − 75+
white − HS or less − 65−74
white − HS or less − 50−64
white − HS or less − 40−49
white − HS or less − 30−39
white − HS or less − 18−29
white − college − 75+
white − college − 65−74
white − college − 50−64
white − college − 40−49
white − college − 30−39
white − college − 18−29
nonwhite − some college − 75+
nonwhite − some college − 65−74
nonwhite − some college − 50−64
nonwhite − some college − 40−49
nonwhite − some college − 30−39
nonwhite − some college − 18−29
nonwhite − postgrad − 75+
nonwhite − postgrad − 65−74
nonwhite − postgrad − 50−64
nonwhite − postgrad − 40−49
nonwhite − postgrad − 30−39
nonwhite − postgrad − 18−29
nonwhite − HS or less − 75+
nonwhite − HS or less − 65−74
nonwhite − HS or less − 50−64
nonwhite − HS or less − 40−49
nonwhite − HS or less − 30−39
nonwhite − HS or less − 18−29
nonwhite − college − 75+
nonwhite − college − 65−74
nonwhite − college − 50−64
nonwhite − college − 40−49
nonwhite − college − 30−39
nonwhite − college − 18−29
Nonwhite X Educ X Age
−3
−2
−1
0
1
Relative Difference: (Pop % − Sample %) / Pop %
ARG
ARGE
ARGE,RxE,RxExA
ARGE,RxE,RxExA,V20
ARGEV20


governor.pro.party
house.pro.party
senate.pro.party
Estimate
0.0
0.1
0.2
0.3
0.4
dem
rep
other
would not vote
Prospective Vote Choice
Weighting Variables
Unweighted
ARG
ARGE
ARGE,RxE,RxExA
ARGEV20
ARGE,RxE,RxExA,V20


Variance in Weights
75+
65−74
50−64
40−49
30−39
18−29
Age
white
other
hispanic
black
asian
Race
male
female
end
some college
postgrad
HS or less
college
educatio
rep
dnv/other
dem
ite 20
some college − white
some college − other
some college − hispanic
some college − black
some college − asian
postgrad − white
postgrad − other
postgrad − hispanic
postgrad − black
postgrad − asian
HS or less − white
HS or less − other
HS or less − hispanic
HS or less − black
HS or less − asian
college − white
college − other
college − hispanic
college − black
college − asian
Race X Educ
white − some college − 75+
white − some college − 65−74
white − some college − 50−64
white − some college − 40−49
white − some college − 30−39
white − some college − 18−29
white − postgrad − 75+
white − postgrad − 65−74
white − postgrad − 50−64
white − postgrad − 40−49
white − postgrad − 30−39
white − postgrad − 18−29
white − HS or less − 75+
white − HS or less − 65−74
white − HS or less − 50−64
white − HS or less − 40−49
white − HS or less − 30−39
white − HS or less − 18−29
white − college − 75+
white − college − 65−74
white − college − 50−64
white − college − 40−49
white − college − 30−39
white − college − 18−29
nonwhite − some college − 75+
nonwhite − some college − 65−74
nonwhite − some college − 50−64
nonwhite − some college − 40−49
nonwhite − some college − 30−39
nonwhite − some college − 18−29
nonwhite − postgrad − 75+
nonwhite − postgrad − 65−74
nonwhite − postgrad − 50−64
nonwhite − postgrad − 40−49
nonwhite − postgrad − 30−39
nonwhite − postgrad − 18−29
nonwhite − HS or less − 75+
nonwhite − HS or less − 65−74
nonwhite − HS or less − 50−64
nonwhite − HS or less − 40−49
nonwhite − HS or less − 30−39
nonwhite − HS or less − 18−29
nonwhite − college − 75+
nonwhite − college − 65−74
nonwhite − college − 50−64
nonwhite − college − 40−49
nonwhite − college − 30−39
nonwhite − college − 18−29
Nonwhite X Educ X Age
0.0
0.5
1.0
1.5
2.0
sd(weight) / mean(weight) within group
ARG
ARGE
ARGE,RxE,RxExA
ARGE,RxE,RxExA,V20
ARGEV20

Actual Weighting Process

....next time

The Washington Post
Democracy Dies in Darkness

New research shows just how badly a citizenship question would hurt the 2020 Census

It could lead to a huge undercount, particularly of Latinos and immigrants

Analysis by Matt Barreto, Chris Warshaw, Matthew A. Baum,
Bryce J. Dietrich, Rebecca Goldstein and Maya Sen

April 22, 2019 at 7:45 a.m. EDT

Nonresponse on the Census

1. Latinos and immigrants fear citizenship information wouldn't be protected.
    - "Nationally, only 35 percent of immigrants and 31 percent of Latinos trusted the Trump administration to protect this information and not share it with other federal agencies"

Nonresponse on the Census

1. Latinos and immigrants fear citizenship information wouldn't be protected.
    - "Nationally, only 35 percent of immigrants and 31 percent of Latinos trusted the Trump administration to protect this information and not share it with other federal agencies"
2. A citizenship question experiment lowered willingness to respond to the census.
    - "All respondents were asked whether they would complete the census, but a random half of respondents were told that it would include a citizenship question."
    - "caused a drop of more than two percentage points among all respondents. It caused six-point drop among Latinos and an 11-point drop among those who are foreign born."

Nonresponse on the Census

1. Latinos and immigrants fear citizenship information wouldn't be protected.
   - "Nationally, only 35 percent of immigrants and 31 percent of Latinos trusted the Trump administration to protect this information and not share it with other federal agencies"
2. A citizenship question experiment lowered willingness to respond to the census.
   - "All respondents were asked whether they would complete the census, but a random half of respondents were told that it would include a citizenship question."
   - "caused a drop of more than two percentage points among all respondents. It caused six-point drop among Latinos and an 11-point drop among those who are foreign born."
3. Adding a citizenship question would hurt the quality of information that Americans provide.
   - "The survey mimicked the census short-form questionnaire and included an experiment in which half of respondents received a questionnaire with a citizenship question and half did not."
   - "Latinos skipped 4 percent more questions when they received the citizenship question. Latinos born in Mexico or Central America skipped 11 percent more questions."