Pew Research Center

Morris H. Hansen Memorial Lecture

# Exploring the Assumption That Online Opt-in Respondents Are Answering in Good Faith

**Courtney Kennedy**

*Director of Survey Research*

# Pew Research Center's Lens on Nonprobability

# Overview of Our Nonprobability Work

**Main collaborators**: Andrew Mercer*, Arnold Lau, Nick Hatley, Dorene Asare-Marfo, Scott Keeter, Nick Bertoni

# A well-known finding: estimates from opt-in surveys tend to be less accurate

| Found online opt-in surveys were less accurate | Found online opt-in surveys were just as accurate |
|---|---|
| Malhotra and Krosnick (2007) | Vavrek and Rivers (2008) |
| Chang and Krosnick (2008) | Ansolabehere and Schaffner (2014) |
| Yeager et al. (2011) | |
| Szolnoki and Hoffmann (2013) | |
| Erens et al. (2014) | |
| Sturgis et al. (2016) | |
| Dutwin and Buskirk (2017) | |
| MacInnis et al. (2018) | |
| Pennay et al. (2018) | |
| Silver (2018) | |
| Mercer et al. (forthcoming) | |

# This holds even when relatively advanced statistical approaches are used

## Dutwin and Buskirk (2017)

“advanced techniques such as propensity weighting and sample matching did not improve these measures, and in some cases made matters worse”

## Mercer et al. (2018)

“even the most effective adjustment strategy [sample matching, propensity modeling, and raking] was only able to remove about 30% of the original bias.”

# Why?

# Why?

**Are our adjustment models not sophisticated enough?**

**Or is the data not genuine and, thus, immune to modeling fixes?**

# Prior research on nonprobability respondents being not genuine or fraudulent

**Fraudulent interviews from outside the U.S.**

Kennedy et al. 2018; Moss 2018; Ahler et al. 2019; Kennedy et al. 2021

**Click farms**

Pasternak 2019

**Trolling**

Lopez and Hillygus 2018

**Cases failing various data quality checks**

Fan et al. 2006; Oppenheimer et al. 2009; Walker et al. 2009; Hopper 2012; Baxter 2016; Vannette 2017; Liu and Wronski 2018; Shanahan 2018; McDowell 2019; Puleston 2019; Kennedy et al. 2021

# Advancing the conversation

- Bogus interviews are not randomly distributed across opt-in samples. They are concentrated in certain subgroups.

- For some subgroups, sizeable shares of cases may not actually belong to that subgroup.

- Statistical models assume respondent demographic info is accurate, but in some cases that may be untrue.

- Whether a hybrid design reduces or increases MSE may depend on the subgroup.

# Research design of 2021 benchmarking study

**In 2021 we fielded six samples**

- Opt-in panel 1 n = 4,912
- Opt-in panel 2 n = 4,931
- Opt-in panel 3 n = 4,955
- ABS panel 1 n = 5,027
- ABS panel 2 n = 5,147
- ABS panel 3 n = 4,965

**We used a common questionnaire with 27 benchmark questions**

**We weighted each sample independently on 12 variables**

- core demographic interactions (e.g., race/ethnicity x education, age x sex)
- civic, political participation (volunteerism, registered voter status)
- other known online panel biases (internet frequency, religion, partisanship)

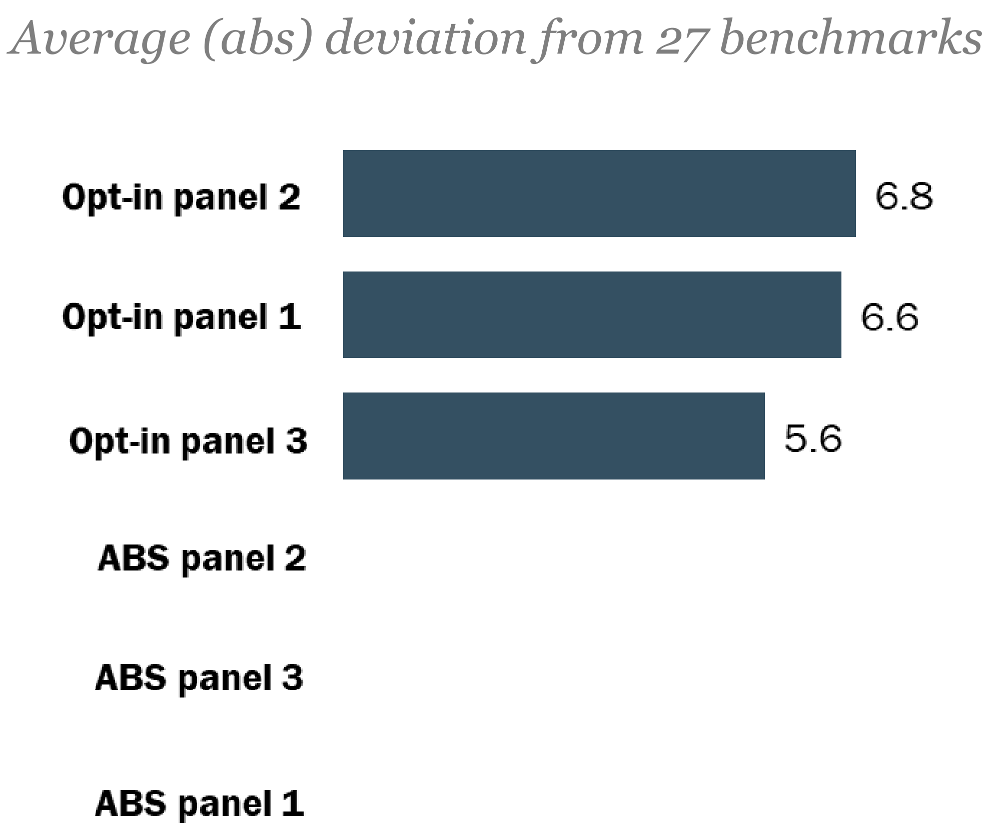
Pew Research Center, Mercer et al. forthcoming

# Consistent with other studies, opt-in sample estimates contained larger errors on average

*Average (abs) deviation from 27 benchmarks*

Pew Research Center, Mercer et al. forthcoming

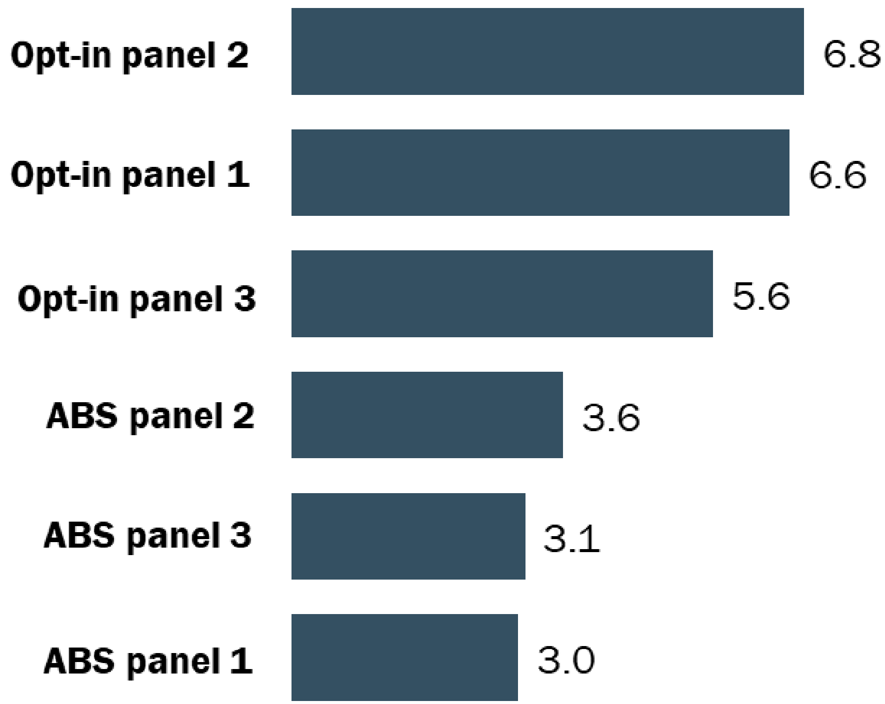
# Consistent with other studies, opt-in sample estimates contained larger errors on average




Pew Research Center, Mercer et al. forthcoming

# Consistent with other studies, opt-in sample estimates contained larger errors on average

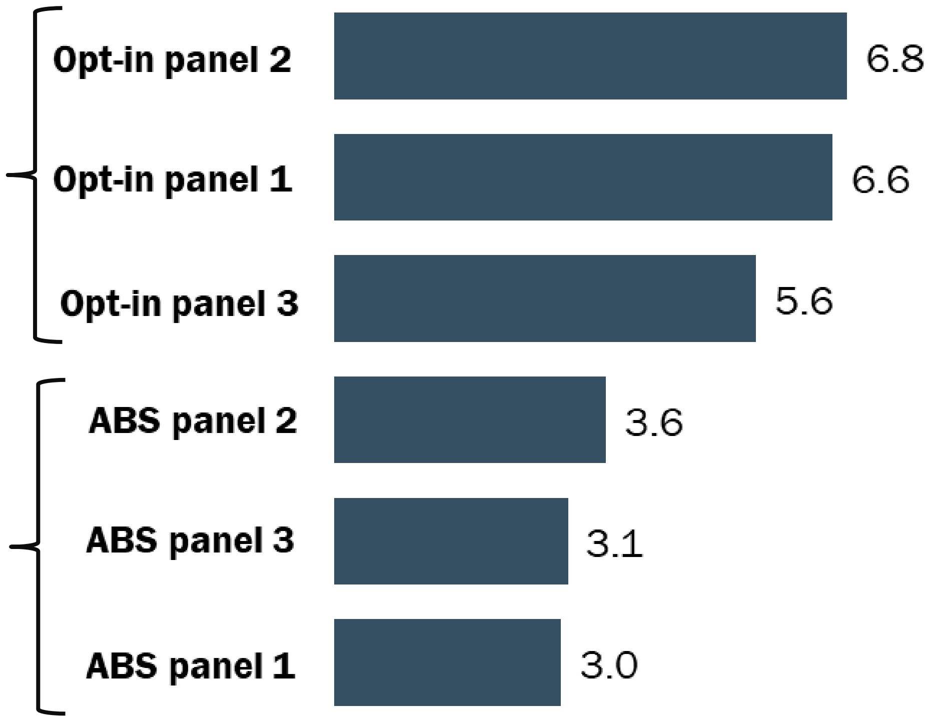
*Average (abs) deviation from 27 benchmarks*




Pew Research Center, Mercer et al. forthcoming

# Consistent with other studies, opt-in sample estimates contained larger errors on average

*Average (abs) deviation from 27 benchmarks*




Subsequent slides show the average for the opt-in panels and the average for the ABS panels

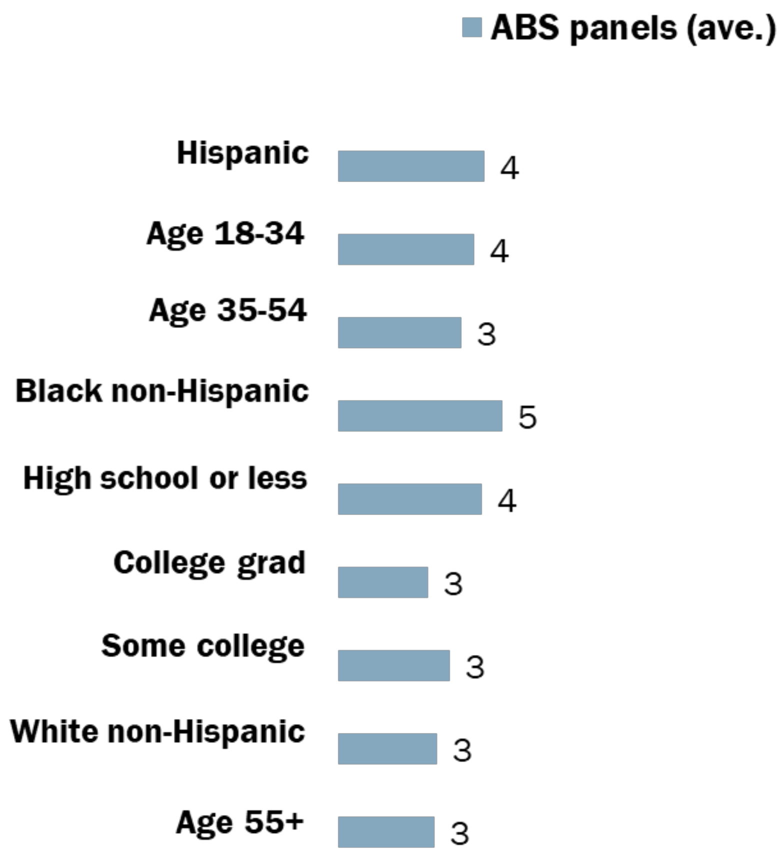
Pew Research Center, Mercer et al. forthcoming

# Where are the large opt-in sample errors coming from?

*Average (abs) deviation from benchmarks, **by subgroup***

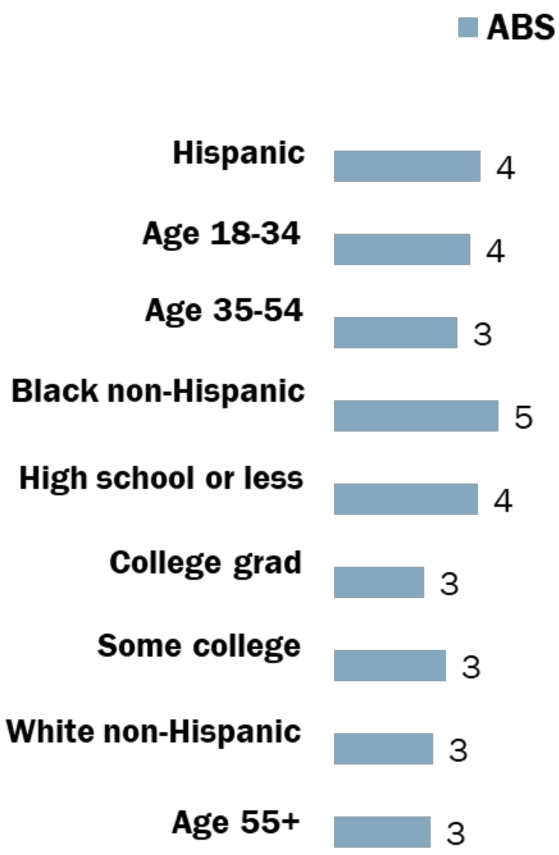
# Where are the large opt-in sample errors coming from?

*Average (abs) deviation from benchmarks, **by subgroup***



# Where are the large opt-in sample errors coming from?

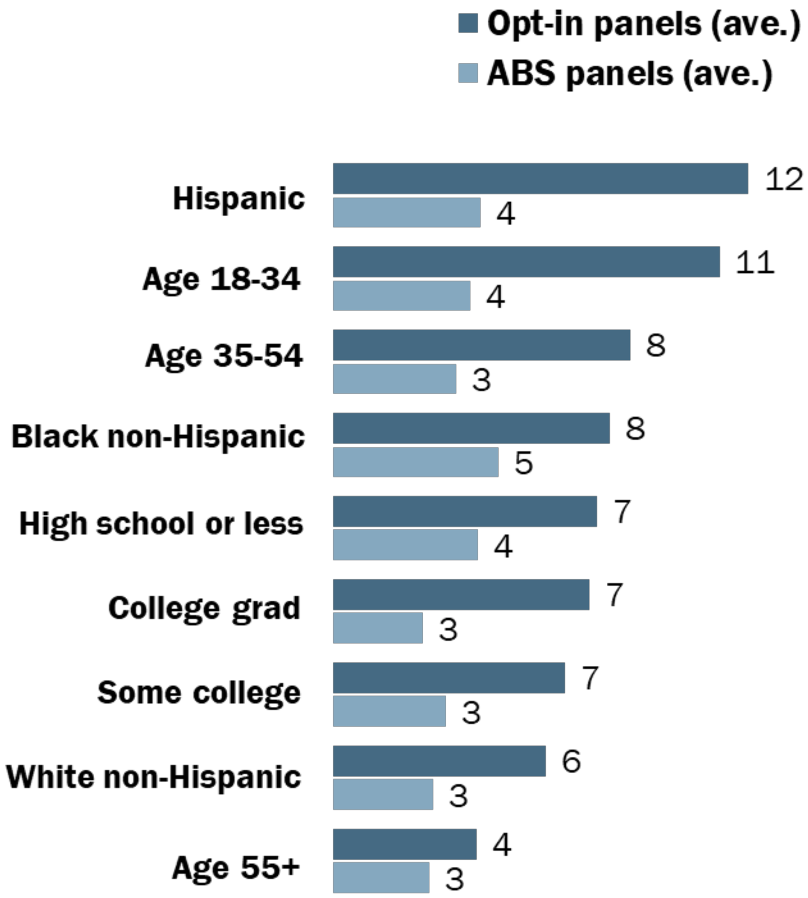
*Average (abs) deviation from benchmarks,* ***by subgroup***




The accuracy of ABS sample estimates does not vary much by subgroup

# Where are the large opt-in sample errors coming from?

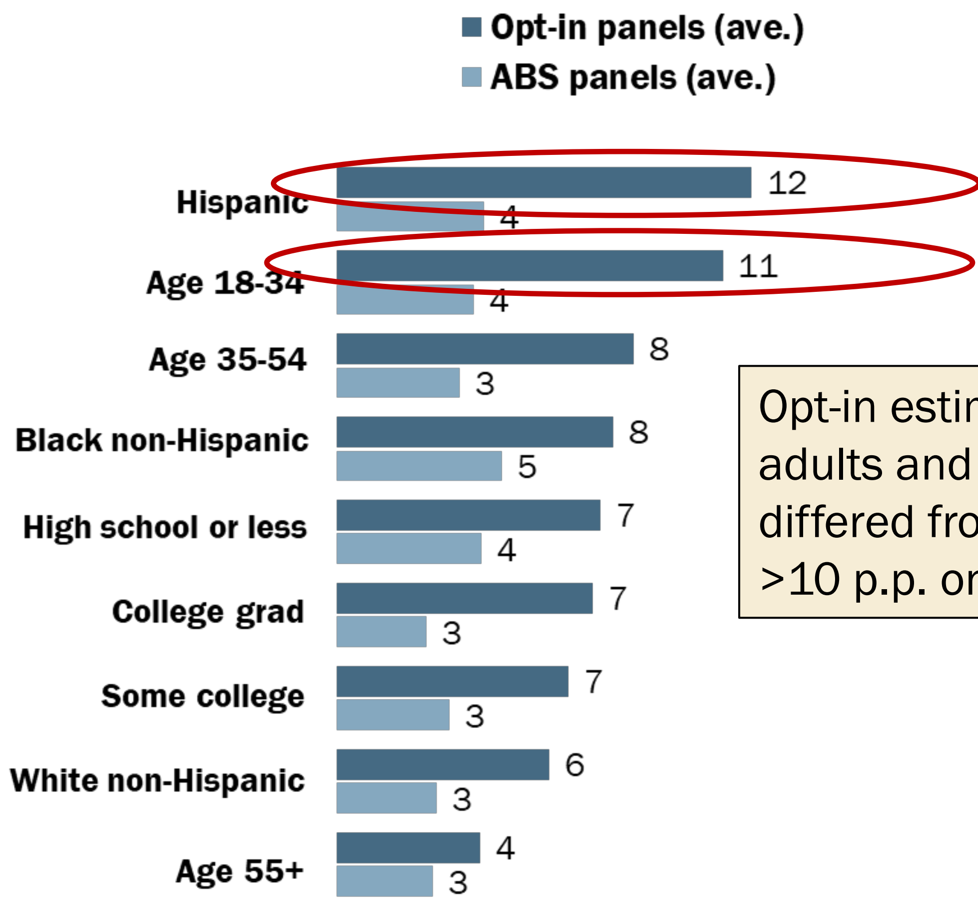
*Average (abs) deviation from benchmarks,* ***by subgroup***




But the accuracy of opt-in estimates varies greatly by subgroup

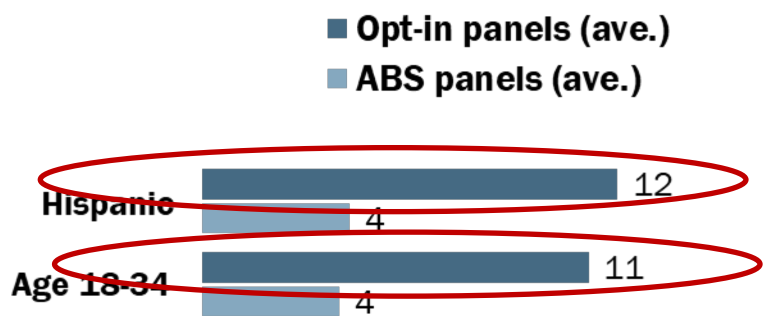
# Where are the large opt-in sample errors coming from?

*Average (abs) deviation from benchmarks, **by subgroup***




Opt-in estimates for Hispanic adults and young adults differed from benchmarks by >10 p.p. on average

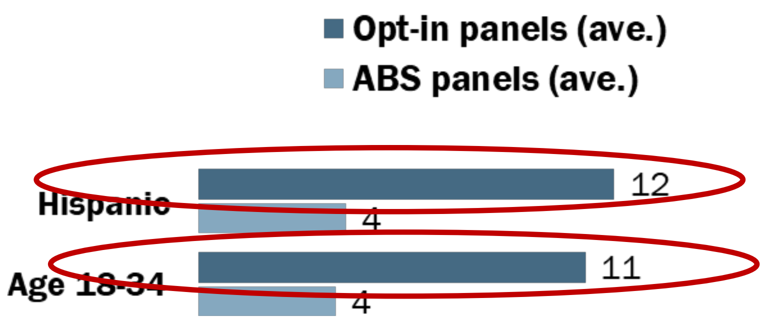
# *Average (abs) deviation from benchmarks, **by subgroup***




Should we attribute this to:

**Nonresponse/Noncoverage error**
These opt-in respondents are genuine (answering in good faith) and just adorably unique or quirky

# *Average (abs) deviation from benchmarks, **by subgroup***



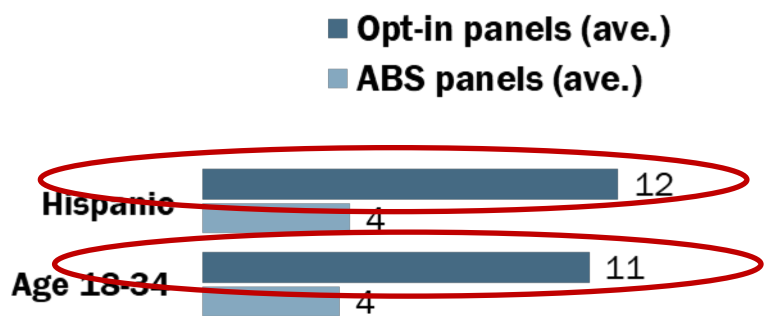

Should we attribute this to:

**Nonresponse/Noncoverage error**
These opt-in respondents are genuine (answering in good faith) and just adorably unique or quirky

OR

**Bogus answers/Bad faith responding**
These opt-in respondents are not who they pretend do be (answering in bad faith)

## *Average (abs) deviation from benchmarks, **by subgroup***




Should we attribute this to:

**Nonresponse/Noncoverage error**
These opt-in respondents are genuine (answering in good faith) and just adorably unique or quirky

OR

**Bogus answers/Bad faith responding**
These opt-in respondents are not who they pretend do be (answering in bad faith)

OR

A combination of the two

# Why does the mechanism matter?

# Why does the mechanism matter?

**Nonresponse/Noncoverage error**
These opt-in respondents are genuine (answering in good faith) and just adorably unique or quirky

This can be fixed with more accurate modeling

# Why does the mechanism matter?

**Nonresponse/Noncoverage error**
These opt-in respondents are genuine (answering in good faith) and just adorably unique or quirky

This can be fixed with more accurate modeling

**Bogus answers/Bad faith responding**
These opt-in respondents are not who they pretend do be (answering in bad faith)

This cannot

# How can we determine which mechanism?

# How can we determine which mechanism?

**Nonresponse/Noncoverage error**
These opt-in respondents are genuine (answering in good faith) and just adorably unique or quirky

Answers would be plausible, but reflective of a biased sample (e.g., too few 1st generation)

# How can we determine which mechanism?

**Nonresponse/Noncoverage error**
These opt-in respondents are genuine (answering in good faith) and just adorably unique or quirky

Answers would be plausible, but reflective of a biased sample (e.g., too few 1st generation)

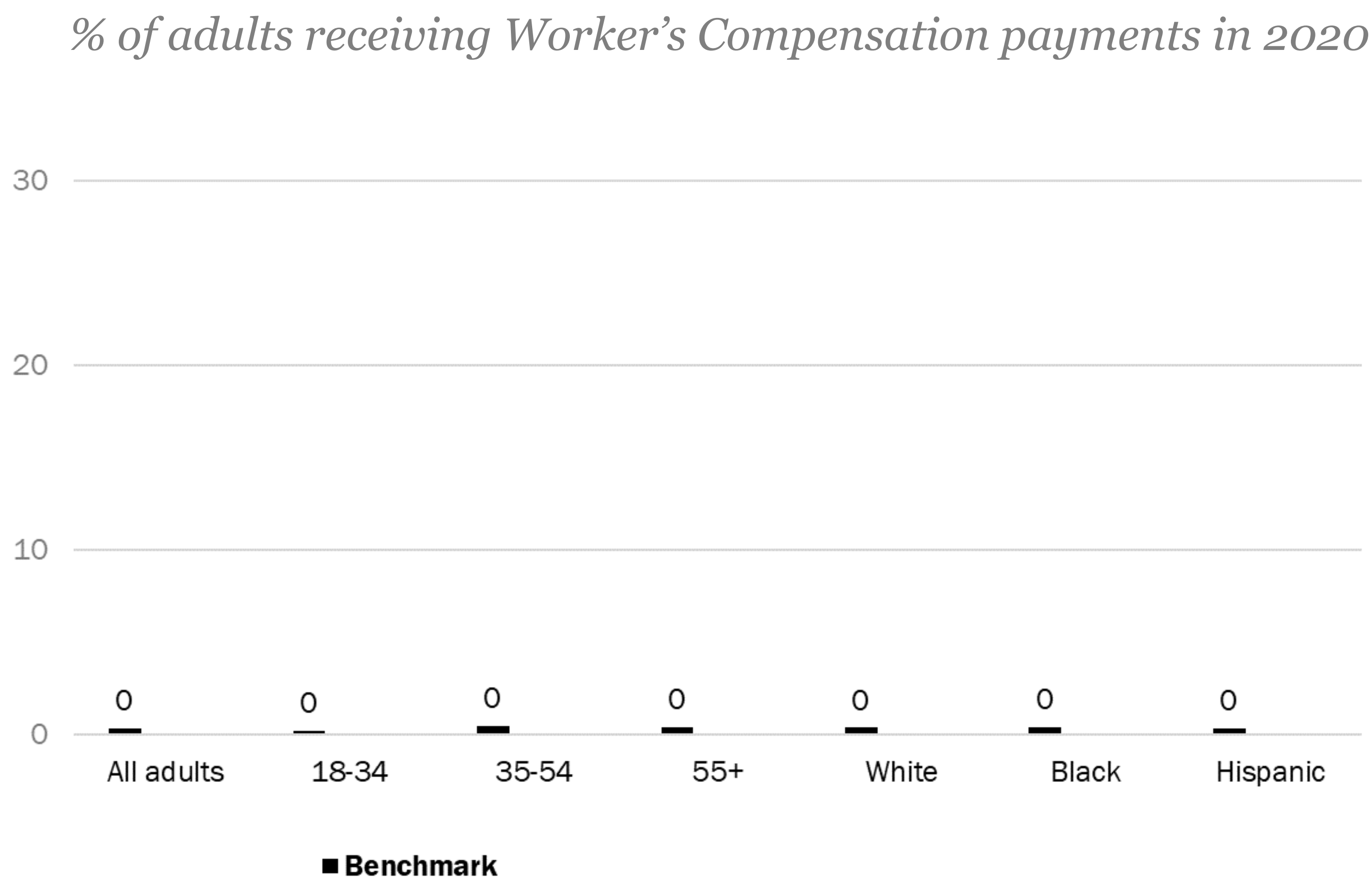
**Bogus answers/Bad faith responding**
These opt-in respondents are not who they pretend do be (answering in bad faith)

Answers would not be plausible, reflective of bad faith

# Do opt-in errors stem from respondents being adorably unique or bogus?

*% of adults receiving Worker's Compensation payments in 2020*

# Do opt-in errors stem from respondents being adorably unique or bogus?

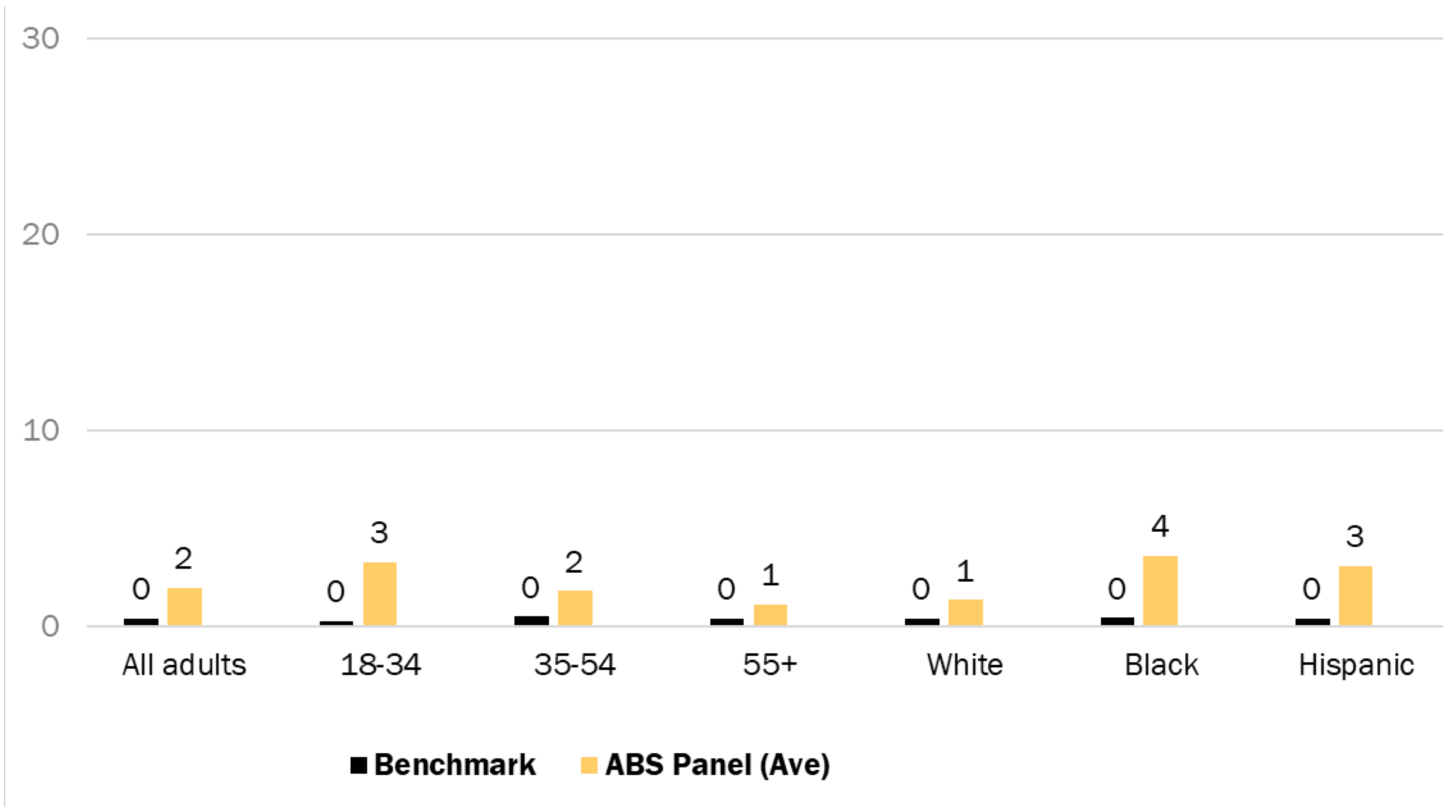





Pew Research Center

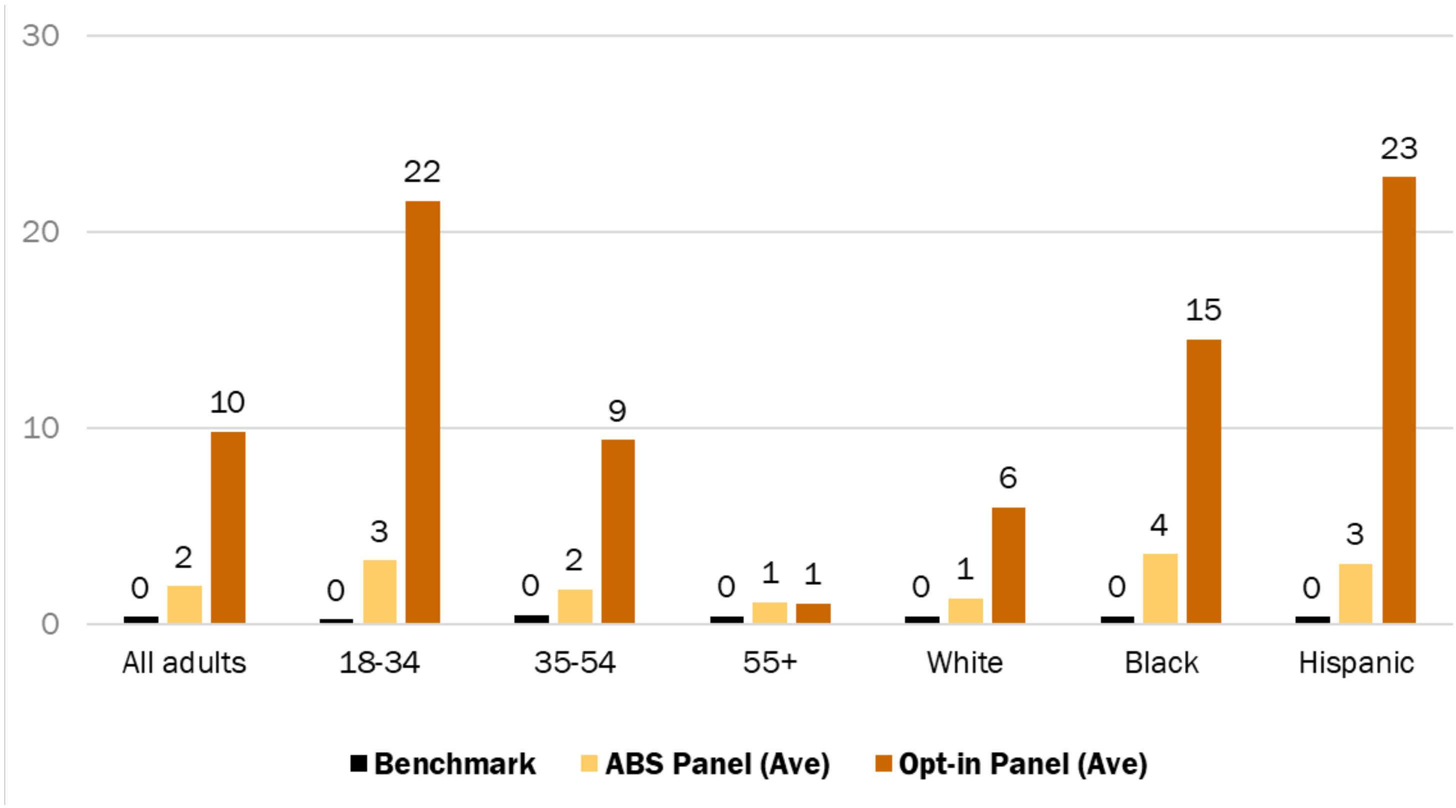
# Do opt-in errors stem from respondents being adorably unique or bogus?

*% of adults receiving Worker's Compensation payments in 2020*



# Do opt-in errors stem from respondents being adorably unique or bogus?

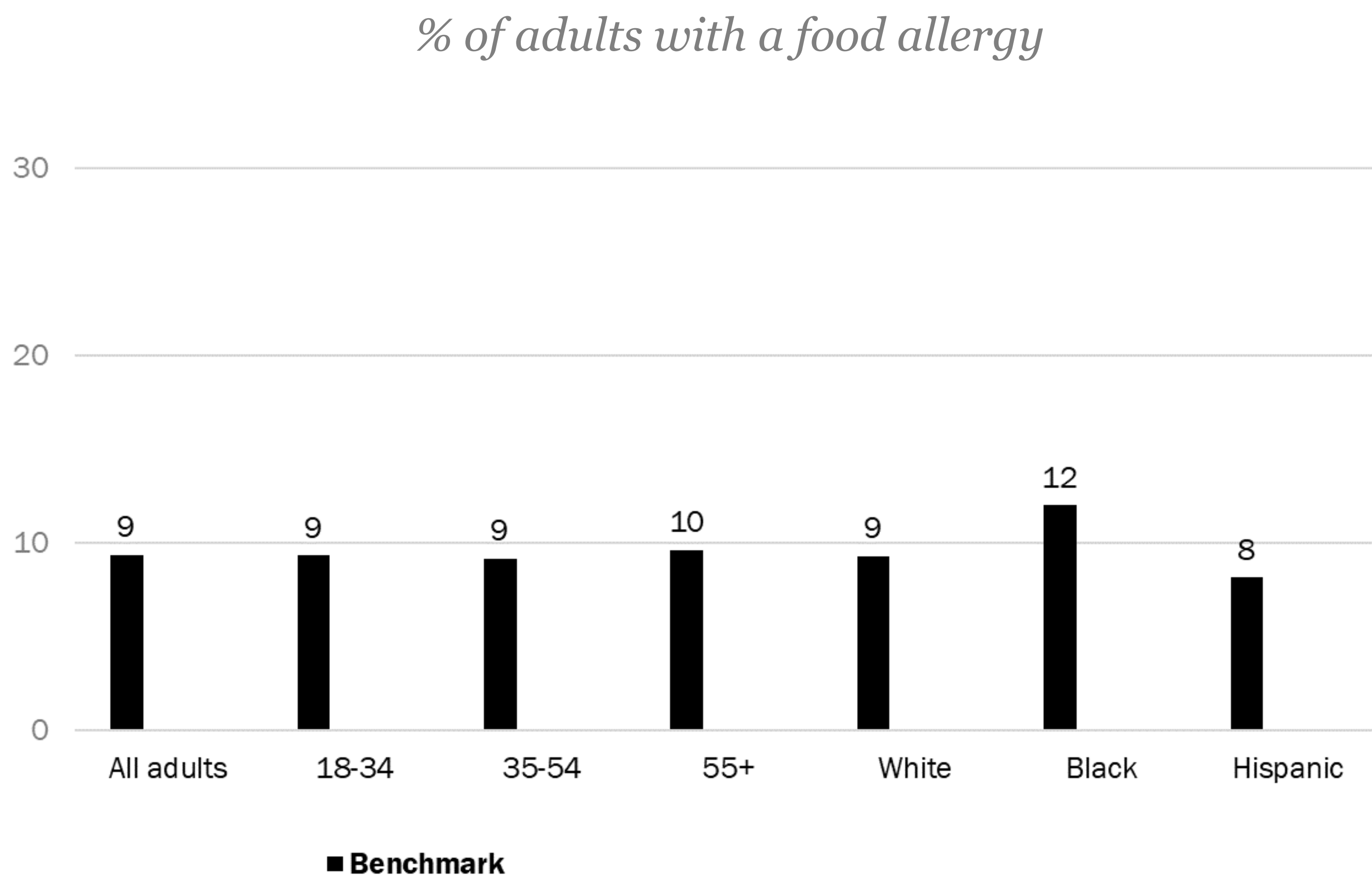
*% of adults receiving Worker's Compensation payments in 2020*



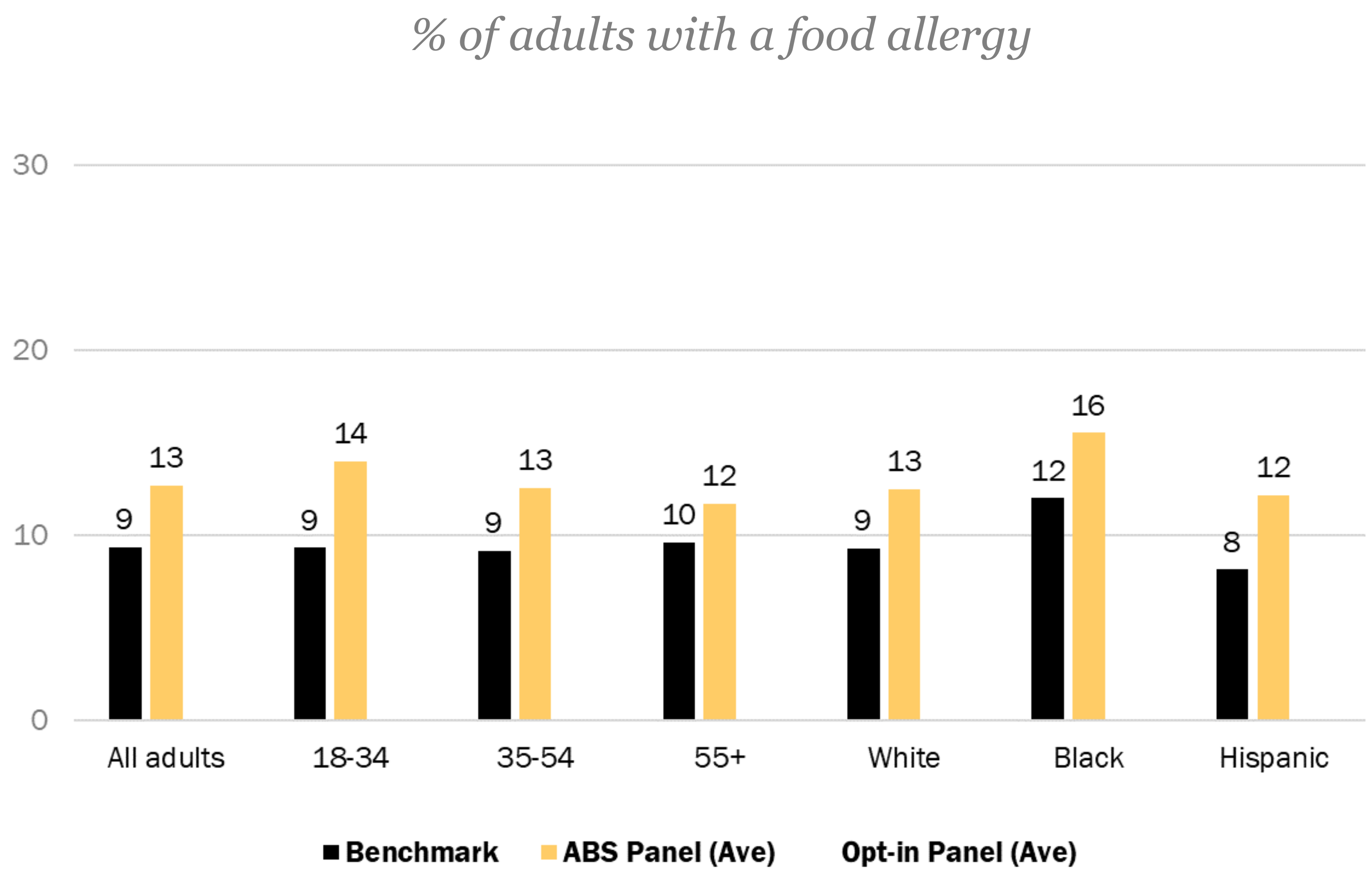
# Do opt-in errors stem from respondents being adorably unique or bogus?

*% of adults with a food allergy*

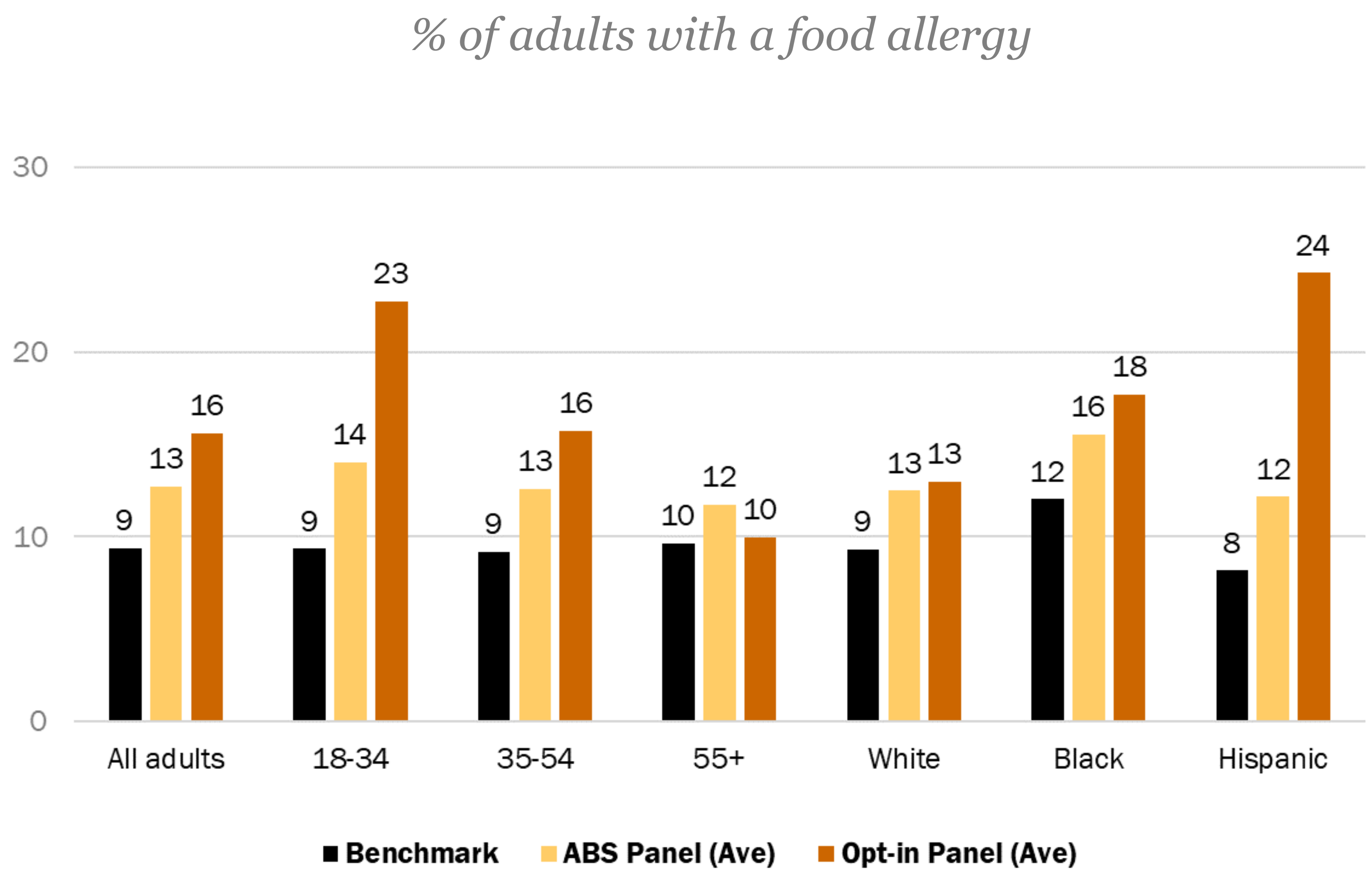
# Do opt-in errors stem from respondents being adorably unique or bogus?



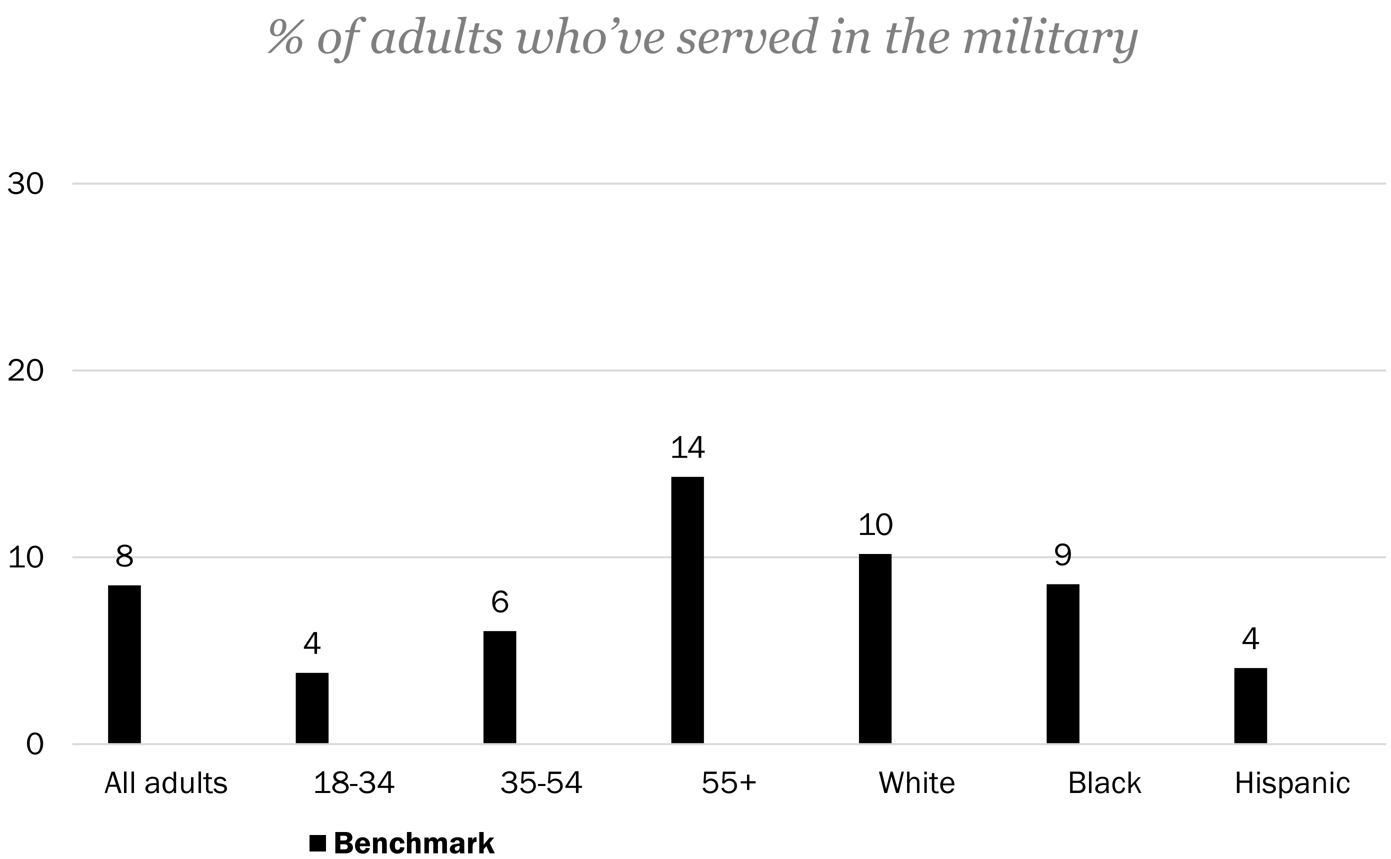
# Do opt-in errors stem from respondents being adorably unique or bogus?



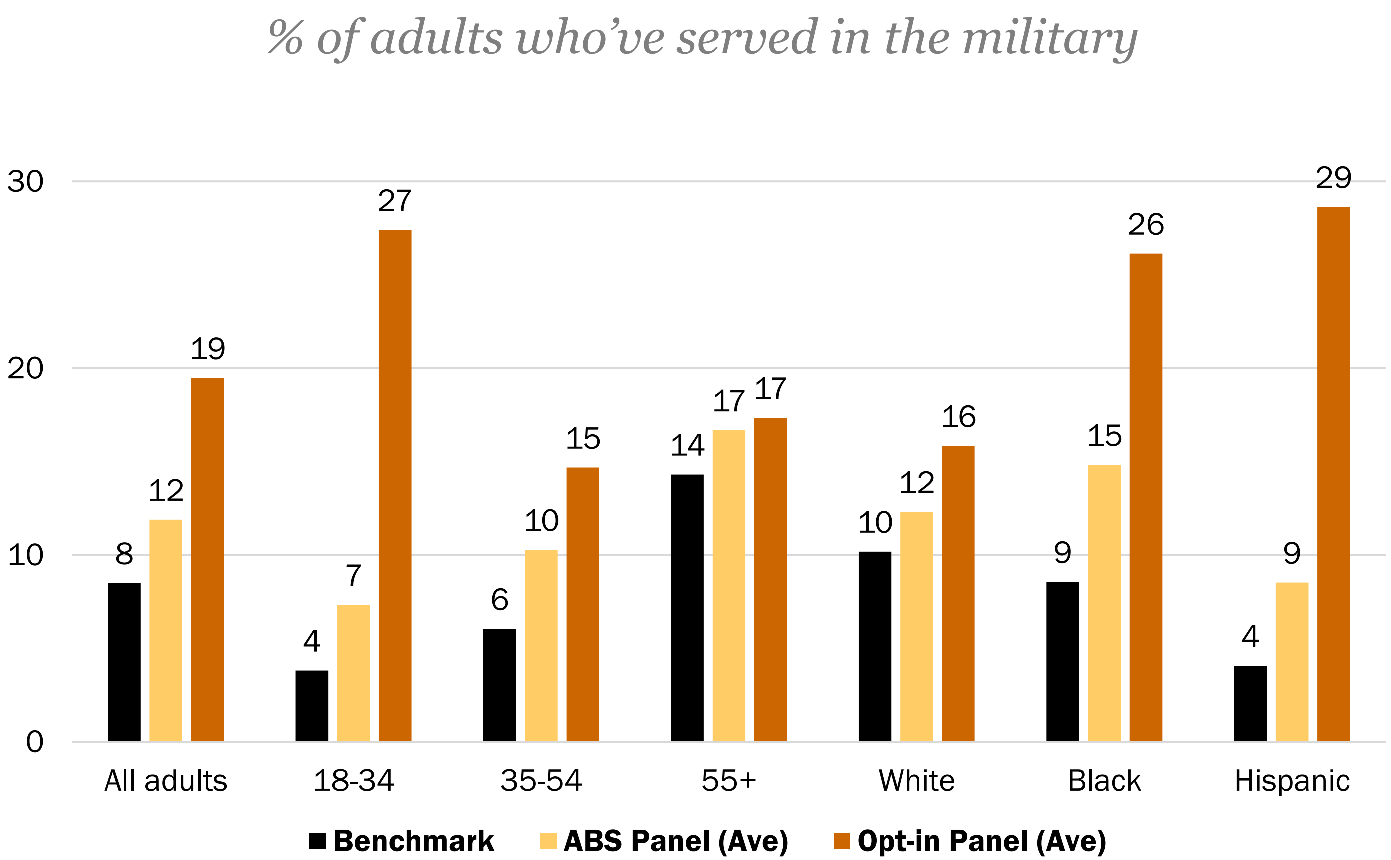
# Do opt-in errors stem from respondents being adorably unique or bogus?



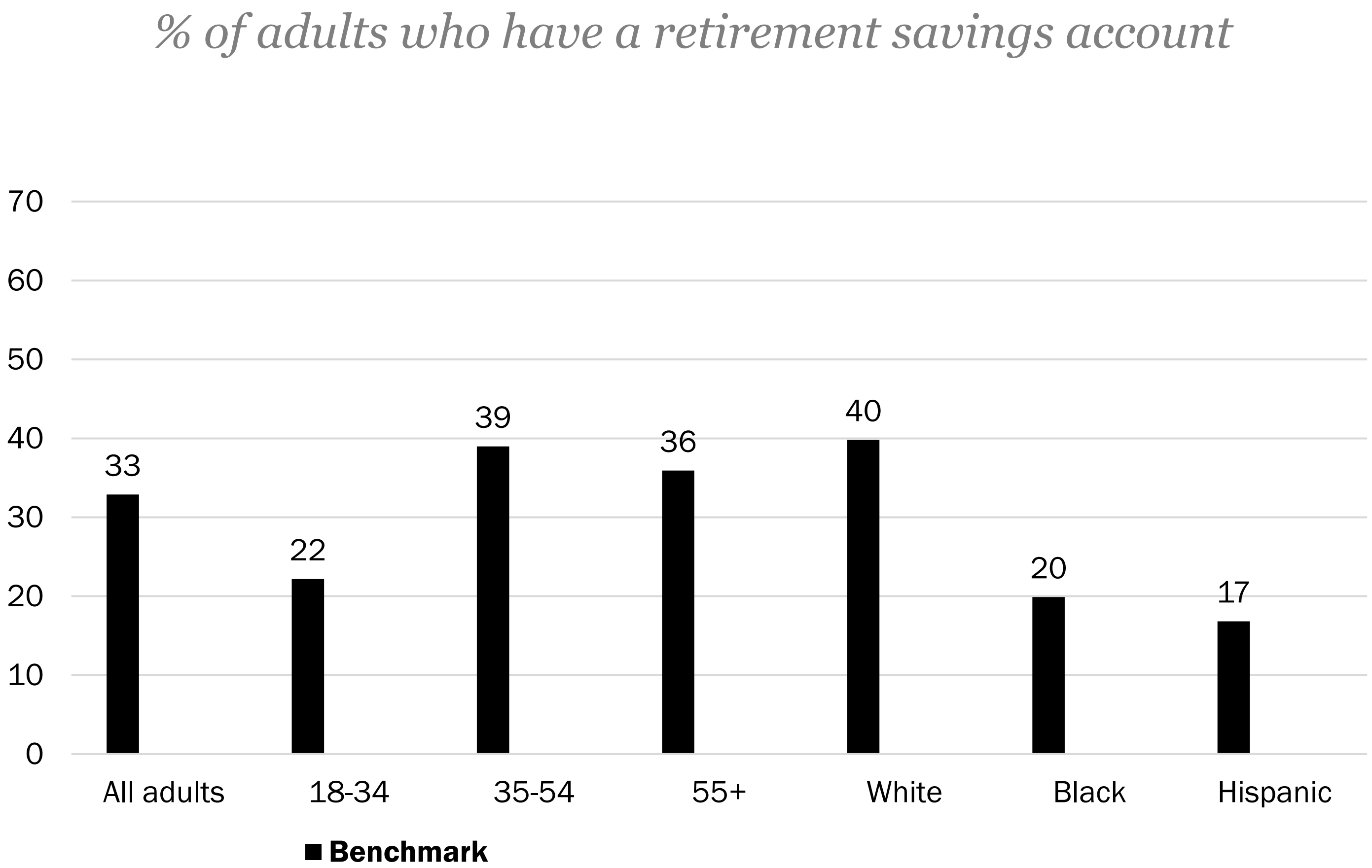
# Do opt-in errors stem from respondents being adorably unique or bogus?



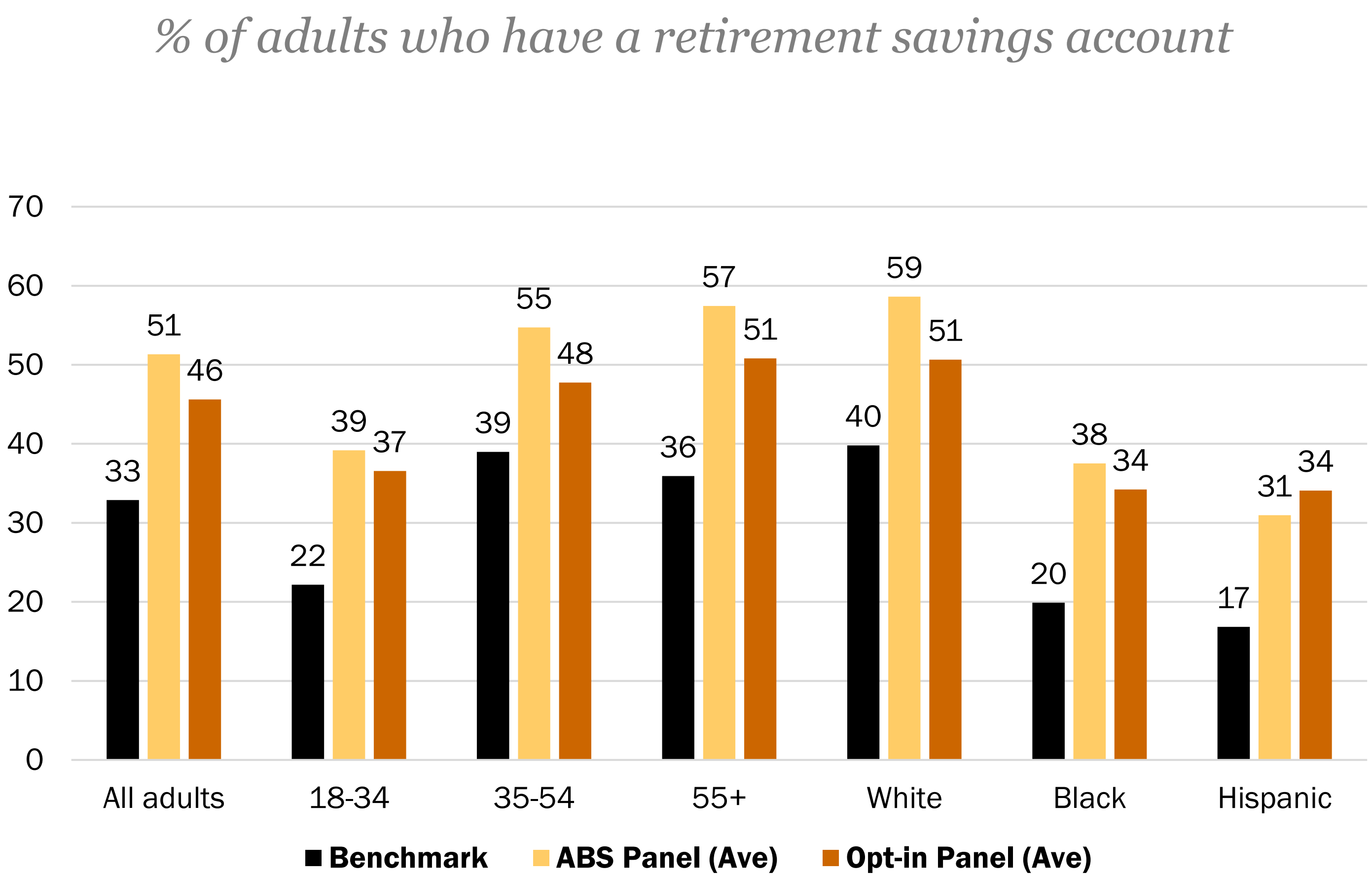
# Do opt-in errors stem from respondents being adorably unique or bogus?



# A counter example



# A counter example






Pew Research Center

# ARE WE SURE THESE HISPANIC OPT-IN CASES ARE ACTUALLY HISPANIC?

# Why might a substantial share of Hispanic opt-in cases not actually be Hispanic?

1. Lying to qualify for more surveys (e.g., Downes-Le Guin et al. 2006)

Are you of Hispanic, Latino, or Spanish origin, such as Mexican, Puerto Rican or Cuban?

○ Yes

○ No

## Why might a substantial share of Hispanic opt-in cases not actually be Hispanic?

1. Lying to qualify for more surveys (e.g., Downes-Le Guin et al. 2006)
2. Positivity bias (e.g., Kennedy et al. 2021)

Are you of Hispanic, Latino, or Spanish origin, such as Mexican, Puerto Rican or Cuban?

- ○ Yes
- ○ No

# Why might a substantial share of Hispanic opt-in cases not actually be Hispanic?

1. Lying to qualify for more surveys (e.g., Downes-Le Guin et al. 2006)
2. Positivity bias (e.g., Kennedy et al. 2021)
3. Haphazard responding

Are you of Hispanic, Latino, or Spanish origin, such as Mexican, Puerto Rican or Cuban?

- ○ Yes
- ○ No

# Testing for bogus cases more directly

- We fielded an opt-in survey using a different vendor
- n=569 completed interviews with U.S. adults
- Included questions to detect clearly bogus data

# Testing for bogus cases more directly

Are you licensed to operate a class SSGN submarine?

- ○ Yes
- ○ No

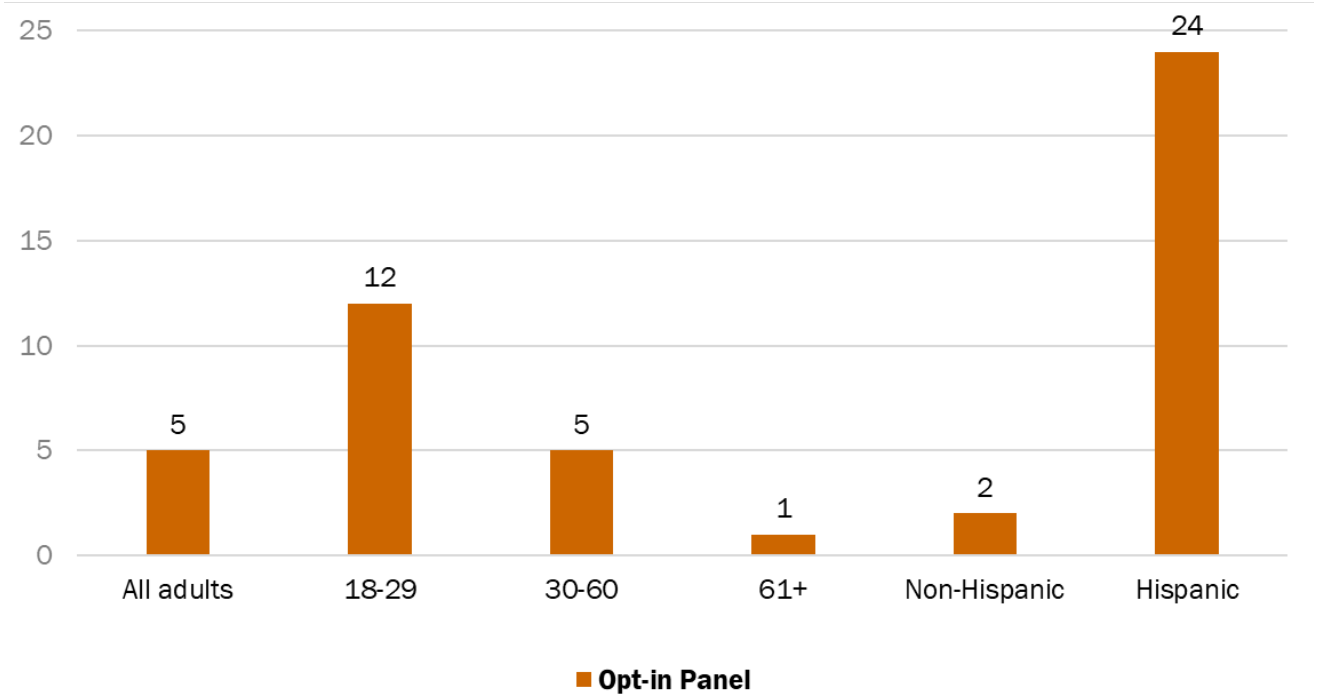
# Testing for bogus cases more directly

Are you licensed to operate a class SSGN submarine?

- ○ Yes
- ○ No

# Testing for bogus cases more directly

*% of respondents saying they are licensed to operate a class SSGN sub*



# Testing for bogus cases more directly

**Which of the following did you do in the past week?**
***Check all that apply.***

- ☐ Purchased a private jet
- ☐ Climbed a peak in the Karakoram Mountains
- ☐ Watched TV
- ☐ Learned to cook halusky
- ☐ Played jai alai
- ☐ Read a book
- ☐ None of the above

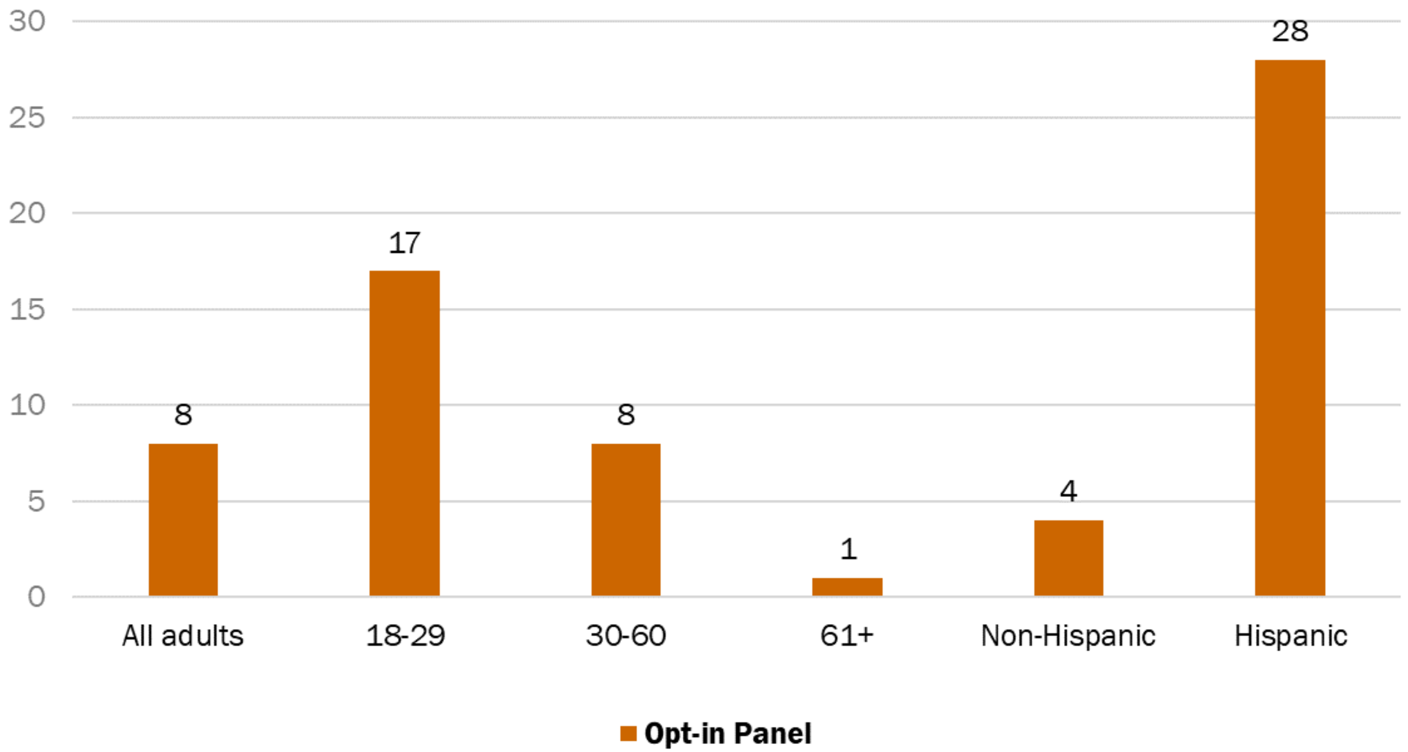
# Testing for bogus cases more directly

**Which of the following did you do in the past week?**
***Check all that apply.***

- ☐ Purchased a private jet
- ☐ Climbed a peak in the Karakoram Mountains
- ☐ Watched TV
- ☐ Learned to cook halusky
- ☐ Played jai alai
- ☐ Read a book
- ☐ None of the above

# Testing for bogus cases more directly

*% of respondents saying they purchased a jet, climbed a Karakoram mountain, cook halusky, or played jai alai in the past week*



# Conclusions

- Statistical models for nonprobability data tend to assume that respondents are who they say they are.
- But we find evidence invalidating that assumption for a sizable portion of young and/or Hispanic nonprobability cases.
- This leads to erroneous patters concerning effects from age and ethnicity.
- We do not think that ethnicity or race is a causal mechanism.
- Nonprobability estimates for older adults are much more accurate. We see few bogus respondents in that age group.
- Whether supplementing a probability sample with nonprobability cases increases or decreases MSE may vary by subgroup. More research is needed to test that.

PewResearchCenter

**Courtney Kennedy**
*Director of Survey Research*
**ckennedy@pewresearch.org**

# Opt-in estimate errors are particularly high for low-incidence outcomes