INSTITUTE FOR SOCIAL RESEARCH
PROGRAM IN SURVEY AND DATA SCIENCE
UNIVERSITY OF MICHIGAN

Ipsos

Representative Research: Assessing Diversity in Online Samples?

Frances M. Barlas, Ph.D.

Ipsos Public Affairs

Presented at University of Michigan Webinar
November 10, 2021

Presenter

Frances M. Barlas, Ph.D.
(she/her)
SVP and Head of KnowledgePanel
Ipsos Public Affairs

How did I get here?

- **Education:**
  - BS in Psychology – University of Pittsburgh
  - MA in Sociology – Temple University
  - PhD in Sociology – Temple University
- **Career:**
  - NERA Economic Consulting – worked for testifying expert in surveys and sampling
  - ICF International – federal contracting
  - GfK – KnowledgePanel
  - Ipsos – KnowledgePanel

Study Background

- In 2020, we saw a broader awakening to the continued systemic racism throughout all aspects of society and heard renewed calls for racial justice.
- In the wake of the murders of George Floyd and Breonna Taylor, widespread protests erupted in the US and in countries around the world.
- Many began to question the role we play and the work we should be doing to dismantle white supremacy.
- For the survey and market research industries, this has raised many questions including how well our industry does to ensure that our public opinion research captures the full set of diverse voices that make up the United States.

Study Background

- **Survey research has played a key role in documenting the impacts of the COVID-19 pandemic on society.**
- **We know that it has had a differential impact by race and ethnicity, with Black, Latinx, and Native Americans facing disproportionately high rates of infection, hospitalization, and death.**
- **We have also done a lot of work to help unpack COVID vaccine hesitancy.**
- **As such, the stakes of our industry getting this right could not be higher.**

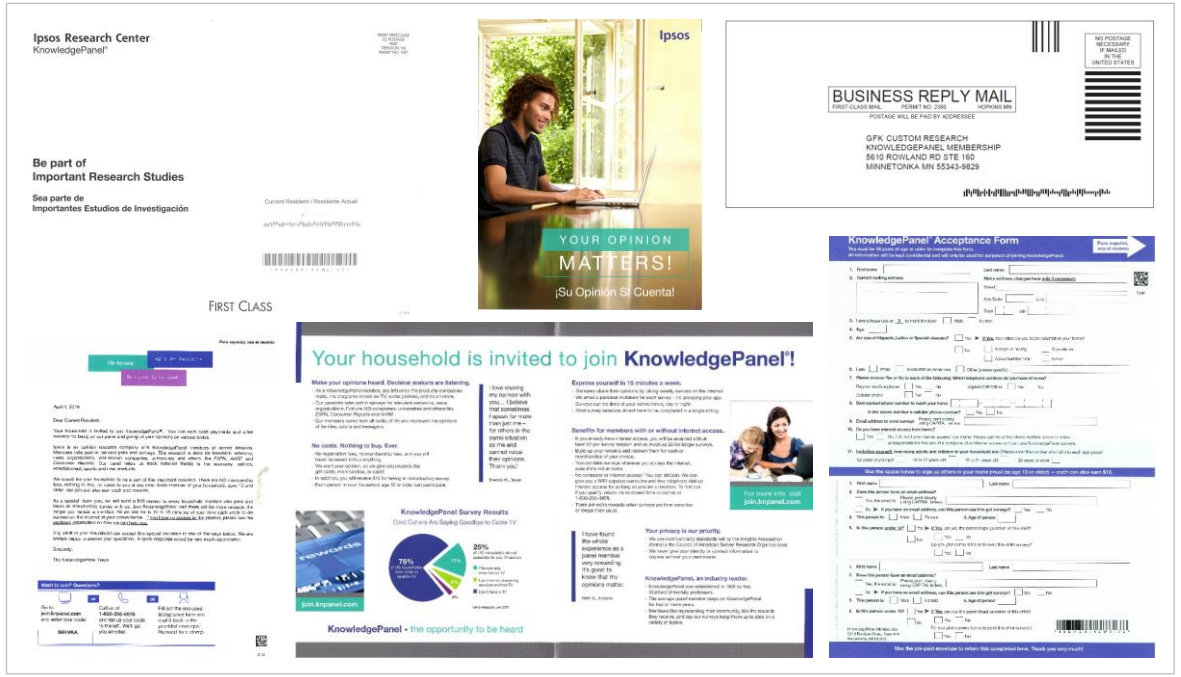
Overview of the Research

- Focus on online samples
- Approach this question from several angles:
  - Recruitment for probability-based panels
  - Online samples:
    - Different online sample sources
    - Demographic variables and other benchmarks
  - The impact of weighting

Types of Internet Samples

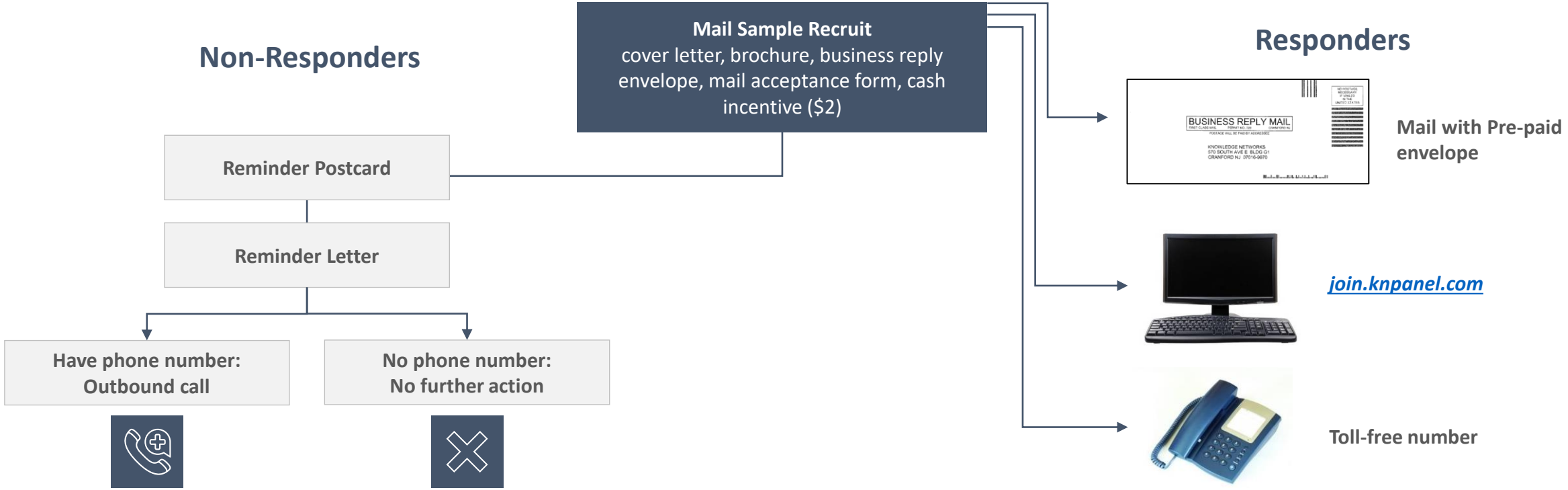
As RDD phone surveys declined, use of online modes of interviewing increased, leading to the rise of both non-probability opt-in samples and probability-based samples

Non-probability based or “opt-in”	Probability-based or “Invitation only”



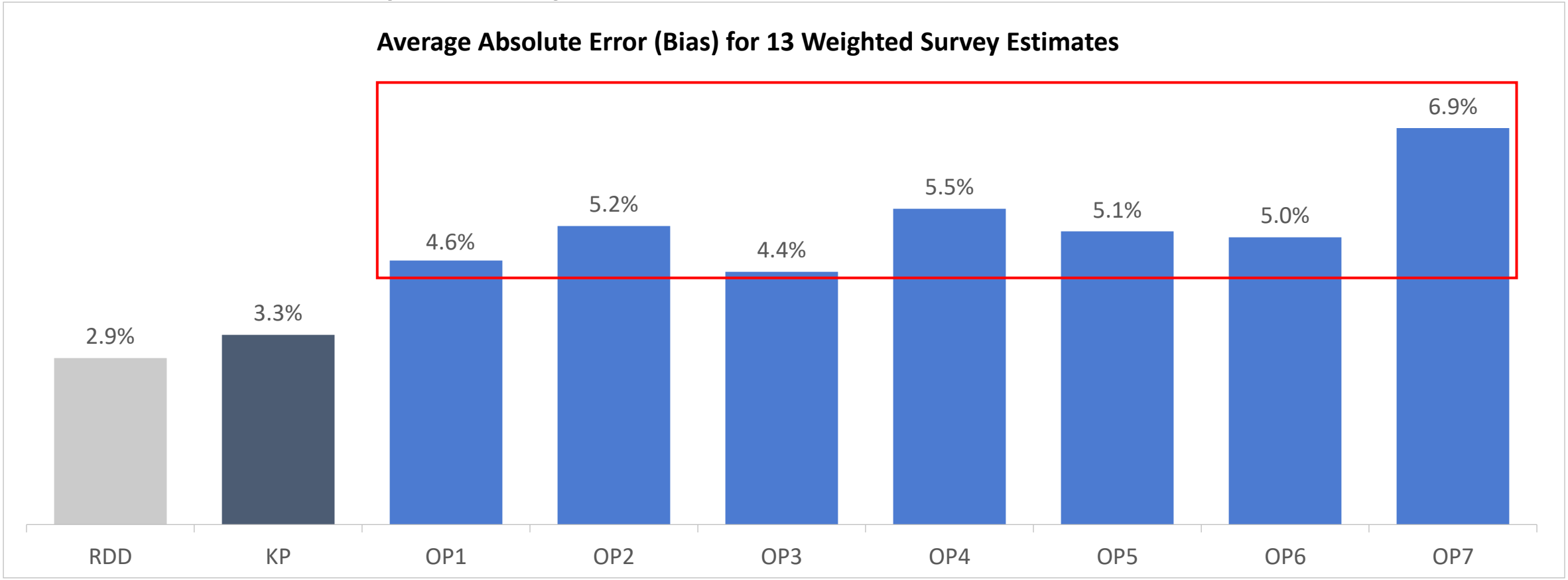
KnowledgePanel

- **KnowledgePanel is an online probability-based panel with about 60,000 members from about 52,000 households**
- **Recruitment is primarily through Address-Based Sampling (ABS)**



Comparing Probability-based to Opt-in Data

A 2011 study comparing general population online samples – both probability-based and opt-in samples – to telephone found that probability-based was closer to RDD and had lower bias than opt-in samples.



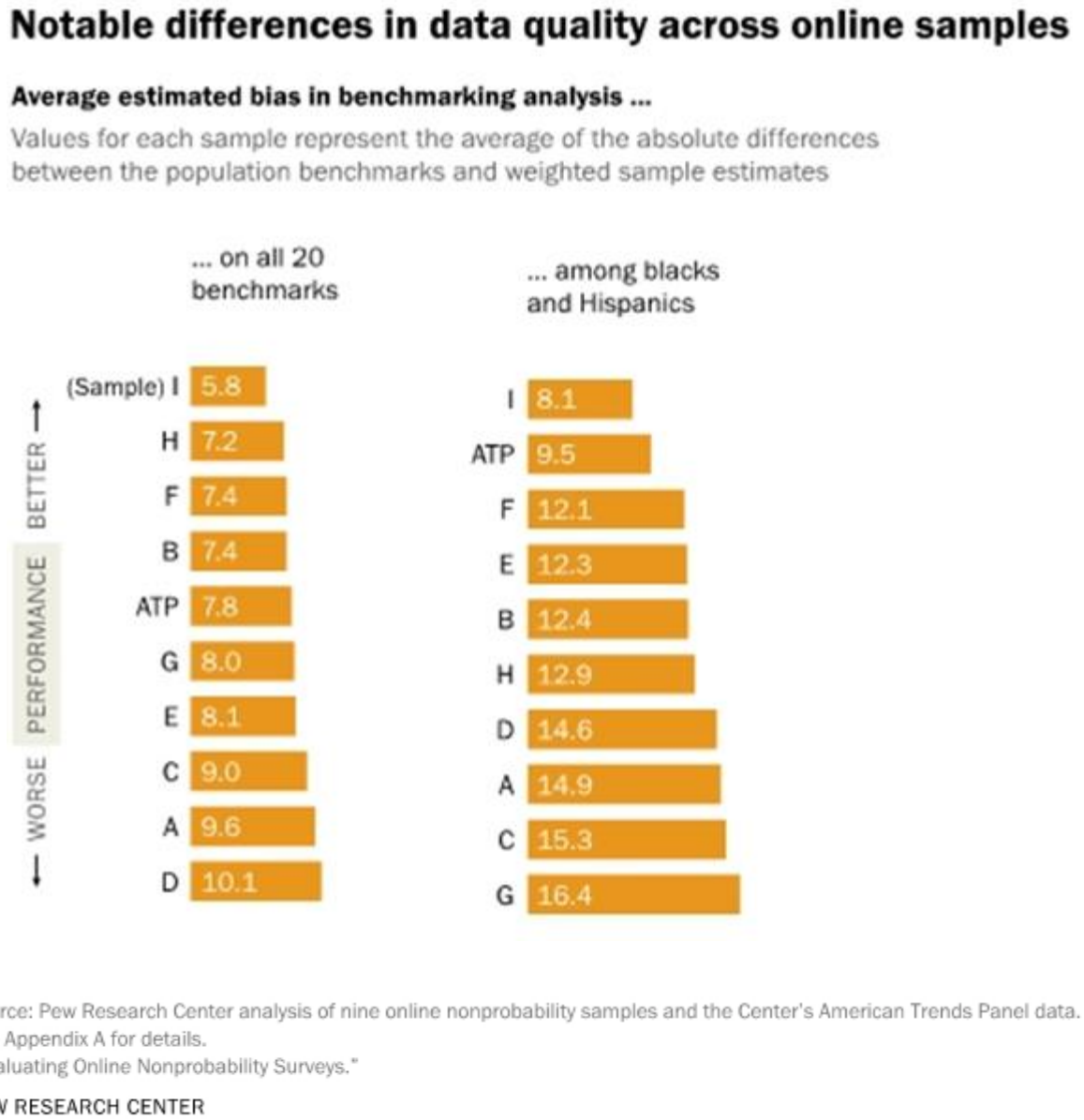
Source: Yeager & Krosnick, et al. "Comparing the Accuracy of Probability & Nonprobability Samples" 2011. Public Opinion Quarterly, 75(4)

Investigations into Bias Focus on Total

- **While past studies have found lower bias in probability-based online panel samples compared to opt-in samples (MacInnis et al., 2018; Yeager et al., 2011), these studies have focused on the overall general population level, without regard to subgroup representation.**
- **There has been less investigation and documentation in the extent of representativeness among subgroups of interest.**

One Study Went Deeper

- A Pew Study by Kennedy et al (2016) found very large divergence (about 10% pts or more) from benchmarks for survey estimates among Black and Hispanic subgroups across 10 online samples.
- Divergence from benchmarks was greater among these subgroups than for the general population level.



KnowledgePanel As A Whole

- Probability-based panels are designed to be miniature representations of the US.
- However, due to differential recruitment and retention rates the panel does have some demographic misalignments compared to Census benchmarks.
- We first wanted to see how does the panel look in terms of representing people of color.

Within each Race/Ethnicity group, we compared the demographics to census Data:

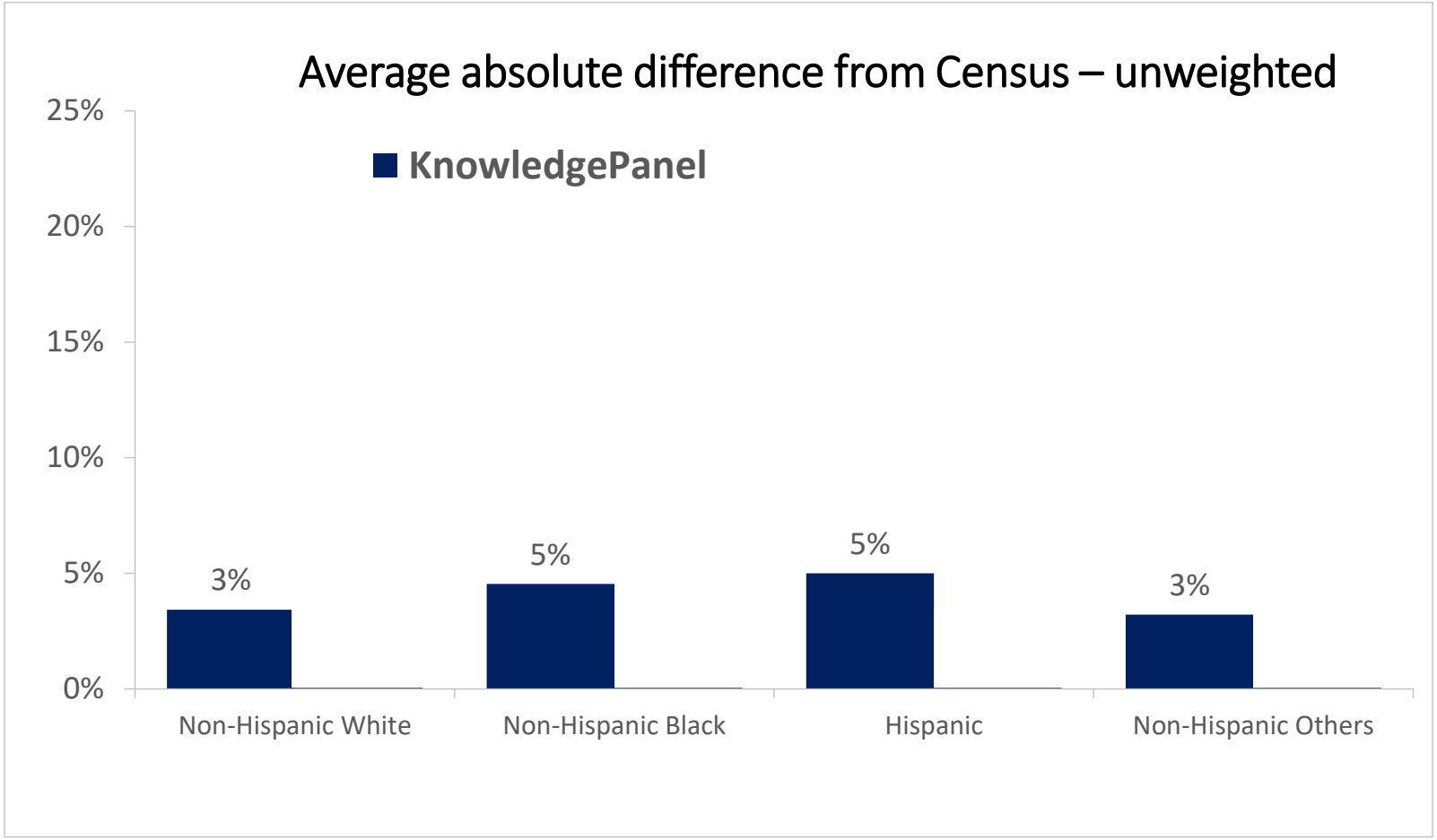
- Age
- Gender
- Education
- Income
- Employment status
- Region of US
- Metropolitan status
- Marital status

KnowledgePanel As A Whole

In comparing KnowledgePanel results for subgroups to Census benchmark data, the unweighted panel subgroups are fairly close to benchmarks on average.

Within each Race/Ethnicity group, we compared the demographics to census Data:

- Age
- Gender
- Education
- Income
- Employment status
- Region of US
- Metropolitan status
- Marital status

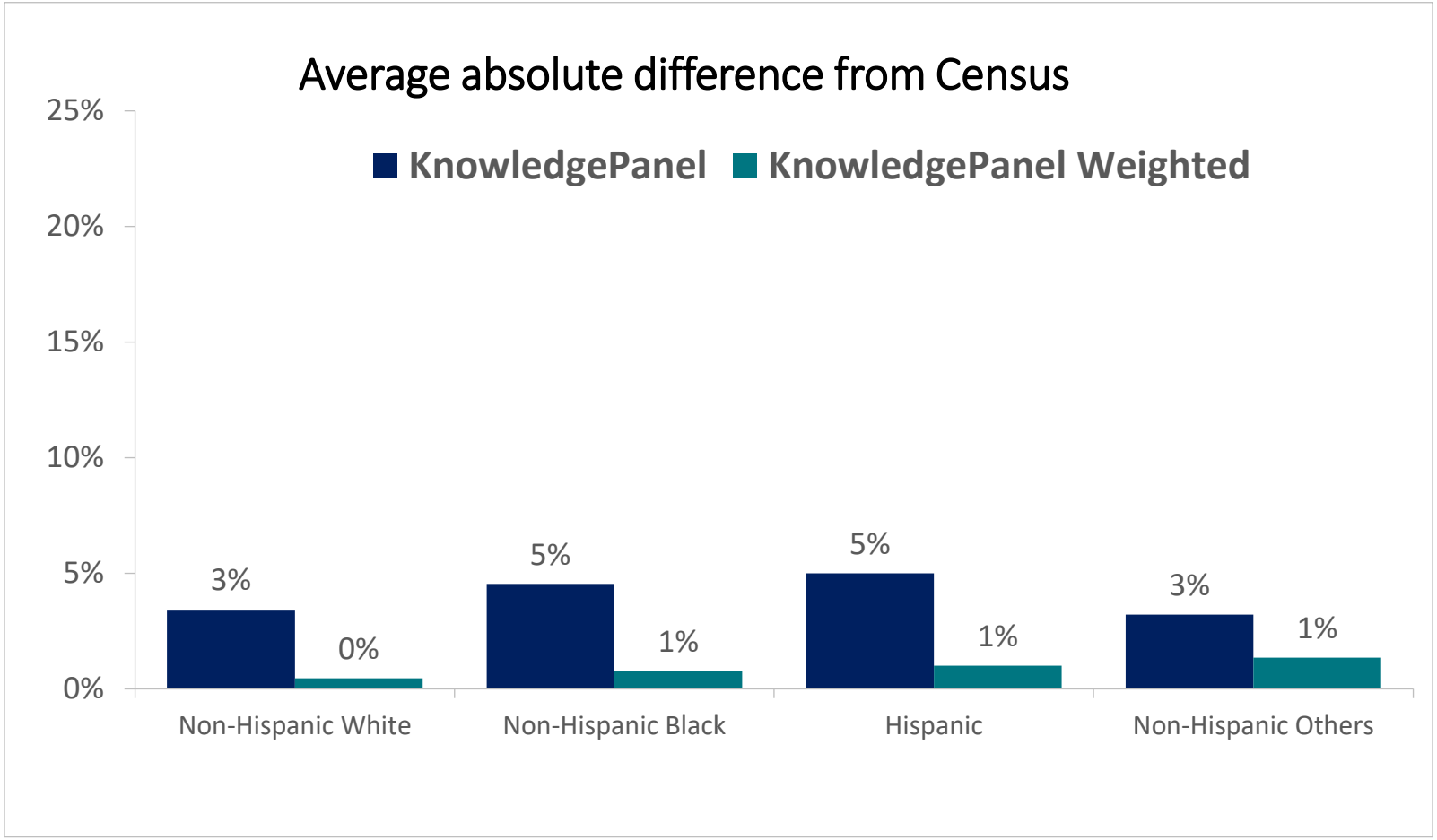


KnowledgePanel As A Whole

When applying weighting typically used for general population study-level samples, the subgroups are even more in line with benchmarks.

Within each Race/Ethnicity group, we compared the demographics to census Data:

- Age
- Gender
- Education
- Income
- Employment status
- Region of US
- Metropolitan status
- Marital status



Survey Results – A Sample Comparison Study

- **We then designed a custom study in Sept. 2020, in which we fielded parallel surveys on both KP and opt-in**
- **Three sample types:**

Sample Type	Sample Details	Number of completes
KnowledgePanel	Standard weighted (PPS) sample	3,344
Opt-in no quota	No quotas	3,293
Opt-in with quota	Quota by age, gender, race/ethnicity, education	2,677

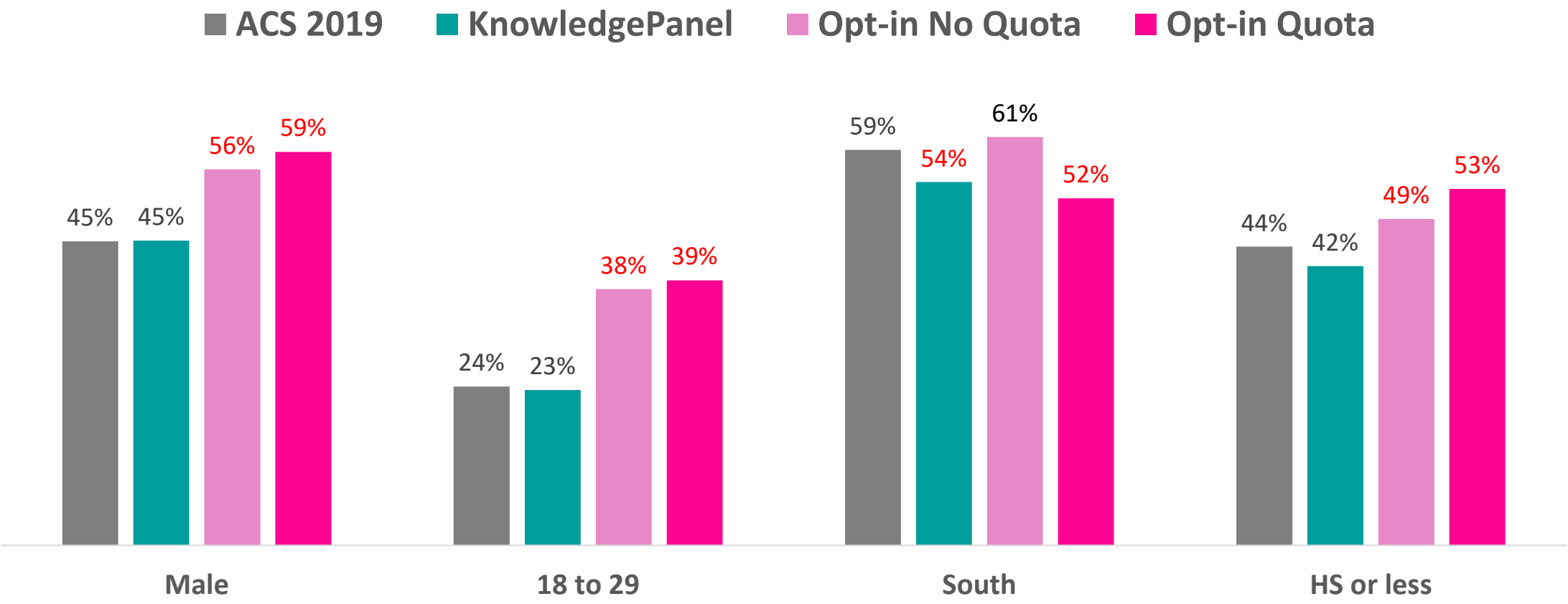
- **Data were weighted to benchmarks from the 2019 American Community Survey for age by gender, race/ethnicity, education, income, and Census region.**

Is Weighting by Race/Ethnicity Overall Adequate?

- **We wanted to see how well our general population samples reflect basic demographics of racial and ethnic subgroups after general, overall weighting.**
- **Standard weighting ensures that the proportion of white, Black, and Latinx matches benchmarks overall, but does not touch demographic distributions within white, Black, and Latinx respondents.**
- **The question is - if we look at weighted results within race/ethnicity, how well do the basic demographics align with population distributions for those groups?**

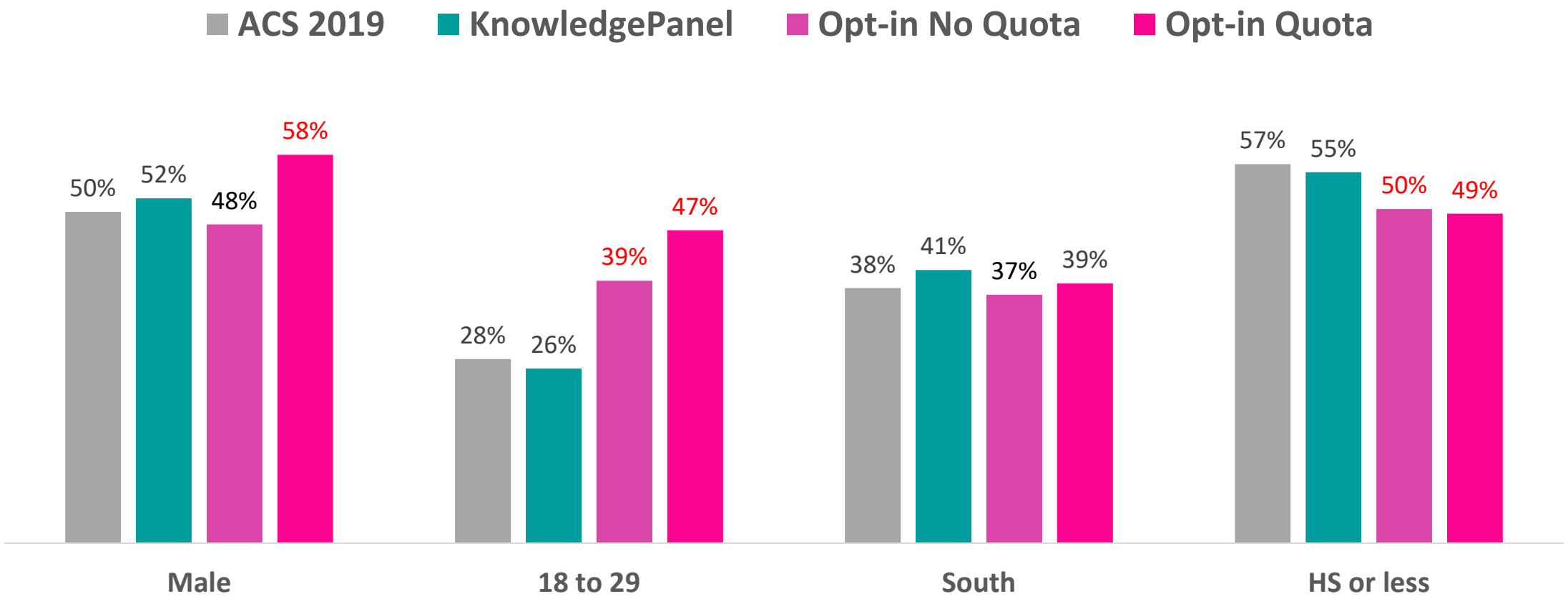
Weighted Demographics Among Black/African American

KP respondents who were Black/African American were more representative than opt-in sample. Opt-in sample with Quotas were furthest from the benchmark.



Weighted Demographics Among Latinx

KP respondents who were Latinx showed greater representativeness than those from opt-in sample. Once again, the Opt-in with quota was furthest from benchmarks.

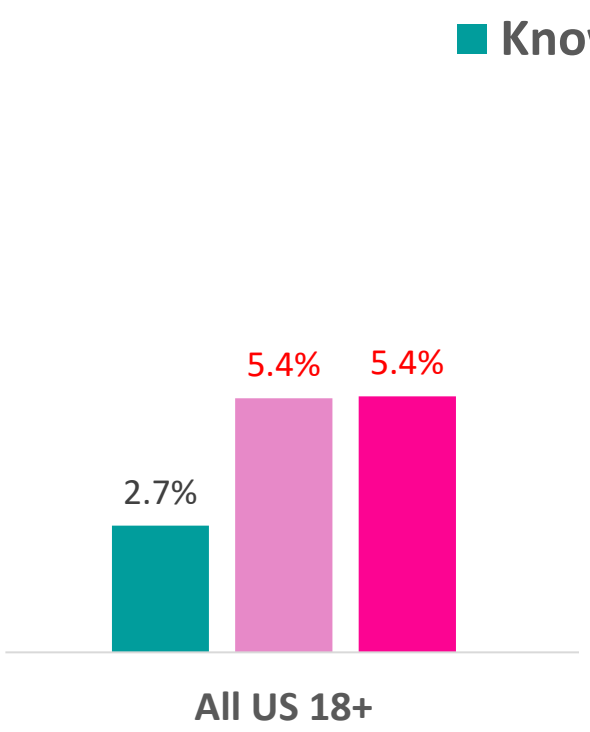


Sample Comparison Study

- **Included 10 benchmarkable items largely from Census data (ACS or CPS)**
- **These items were secondary demographics – not used in weighting**
  - Currently married
  - Citizenship
  - 2 or more in HH
  - At least 1 child under 18 in HH
  - Own house
  - 3 bedrooms or more in HH
  - Moved in current home more than 5 years ago
  - 2 or more vehicles one ton or less
  - Speaks a language other than English at home
  - Has landline phone (NHIS)
- **Calculated the average absolute deviation from benchmarks across 10 measures within race/ethnicity**

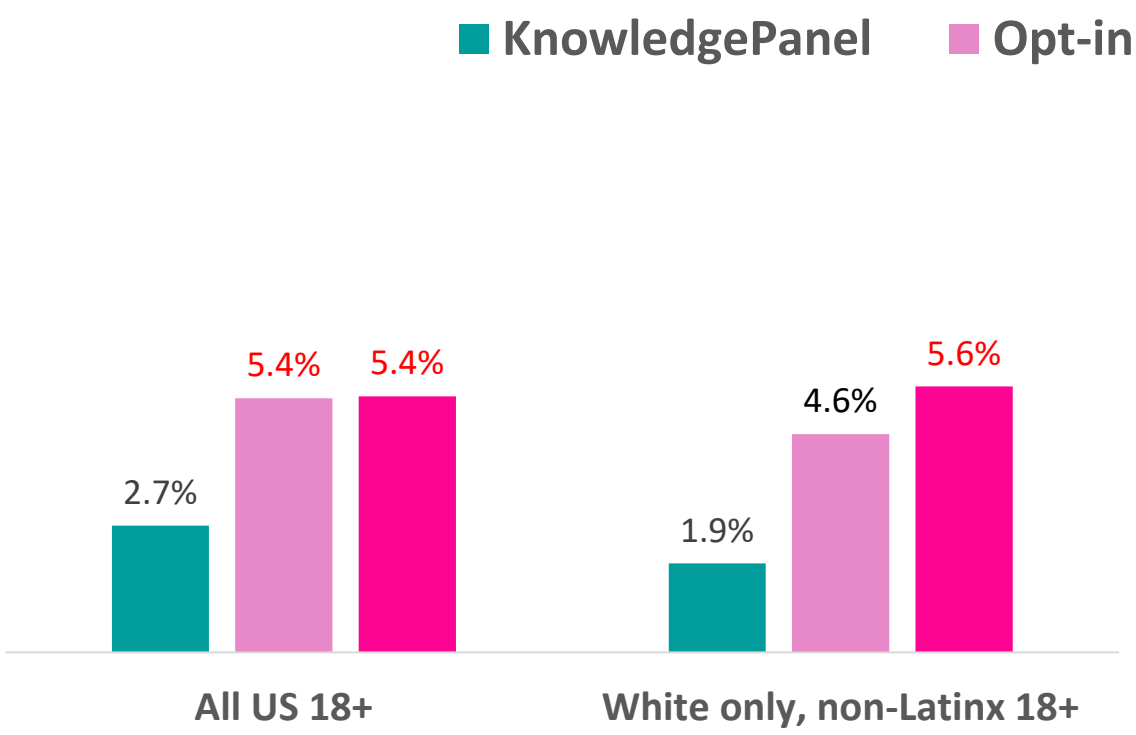
Results – Average Deviation from Benchmarks

We found that bias was lowest for the overall sample and for White participants but was higher for both Black and Latinx participants for both types of sample (KP and Opt-in).



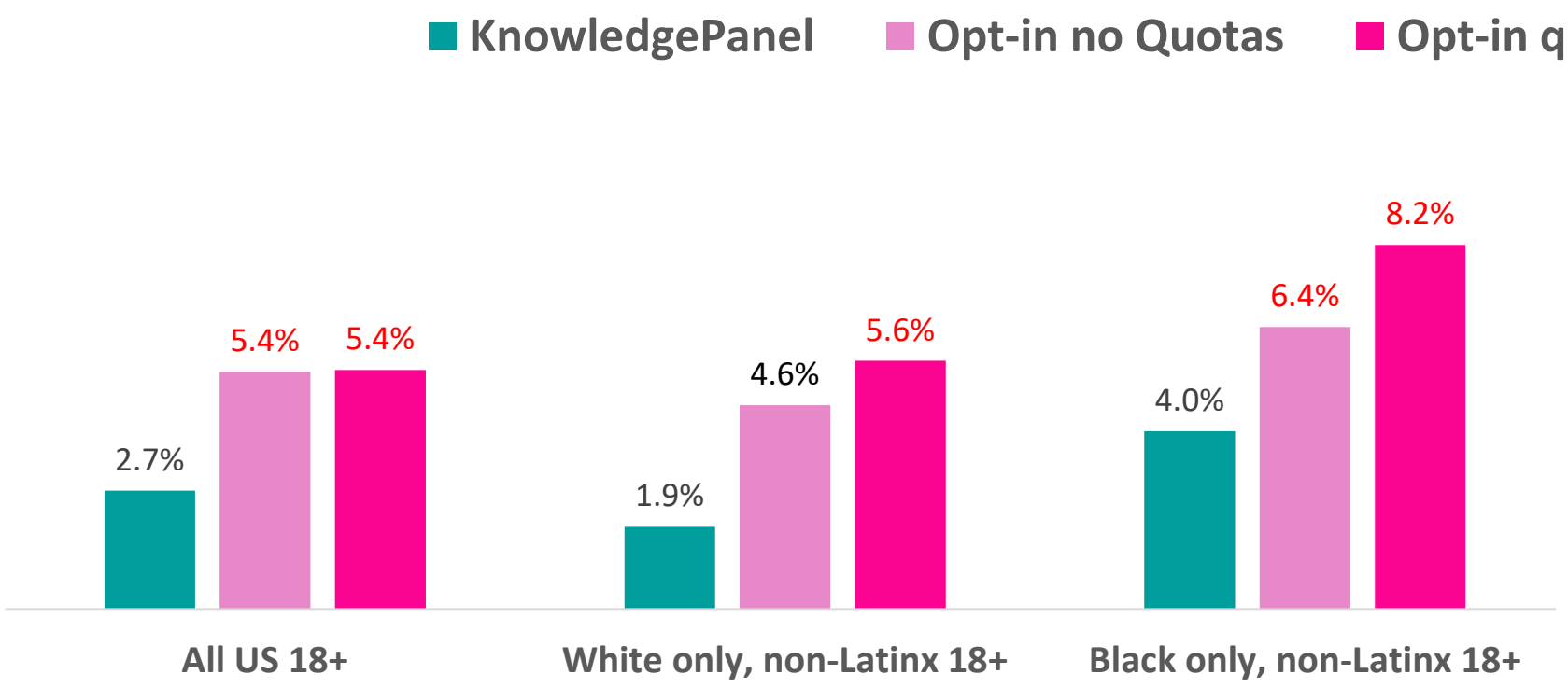
Results – Average Deviation from Benchmarks

We found that bias was lowest for the overall sample and for White participants but was higher for both Black and Latinx participants for both types of sample (KP and Opt-in).



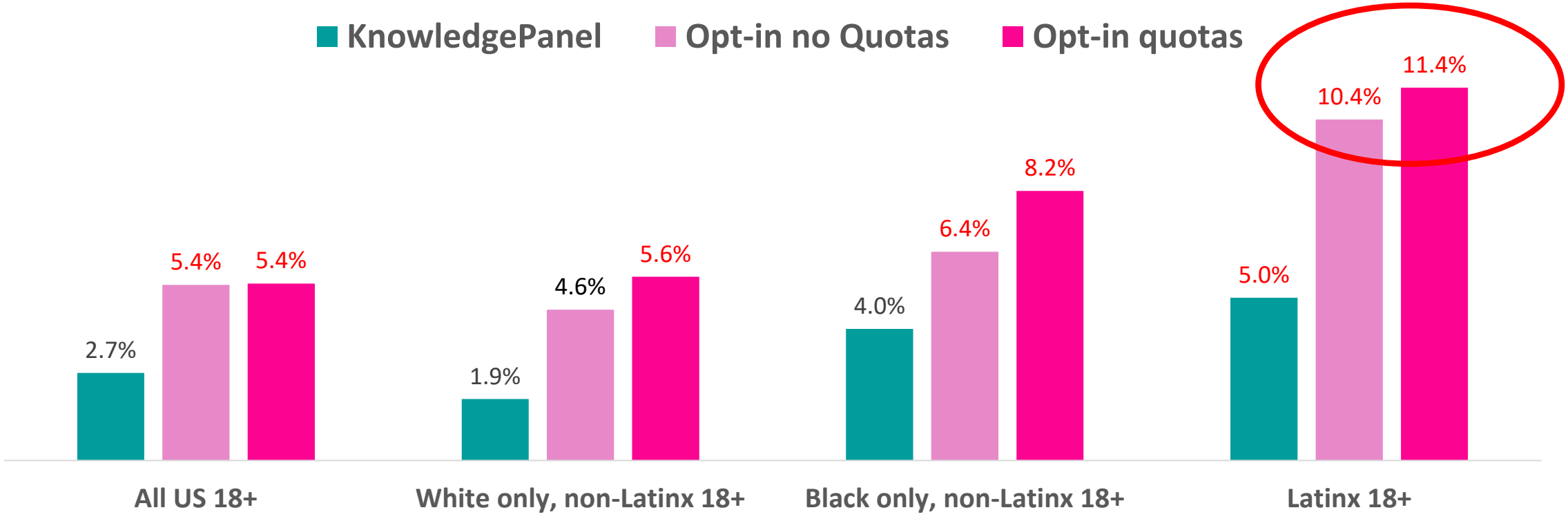
Results – Average Deviation from Benchmarks

We found that bias was lowest for the overall sample and for White participants but was higher for both Black and Latinx participants for both types of sample (KP and Opt-in).



Results – Average Deviation from Benchmarks

We found that bias was lowest for the overall sample and for White participants but was higher for both Black and Latinx participants for both types of sample (KP and Opt-in).



Would More Extensive Weighting Help?

- We've seen that we do find higher bias among Black and Latinx subgroups – average divergence from benchmarks is larger.
- Typical geodemographic weighting at the overall level does not sufficiently align distributions with benchmarks even among weighting variables when looking within Black and Hispanic subgroups.
- While KnowledgePanel exhibited lower bias than opt-in sample, we found that both KP and opt-in showed higher divergence from benchmarks among Black and Latinx subgroups.
- Both samples showed highest bias among Latinx respondents – KP was 5% pts off and opt-in near 11% pts off on average.
- What happens when we extend weighting to adjust within race and ethnicity?

Does Additional Nested Weighting Help?

Example of Normal Weighting, no nesting of groups for race-ethnicity

Age 18-29 Male
Age 18-29 Female
Age 30-44 Male
Age 30-44 Female
Age 45-59 Male
Age 45-59 Female
Age 60+ Male
Age 60+ Female

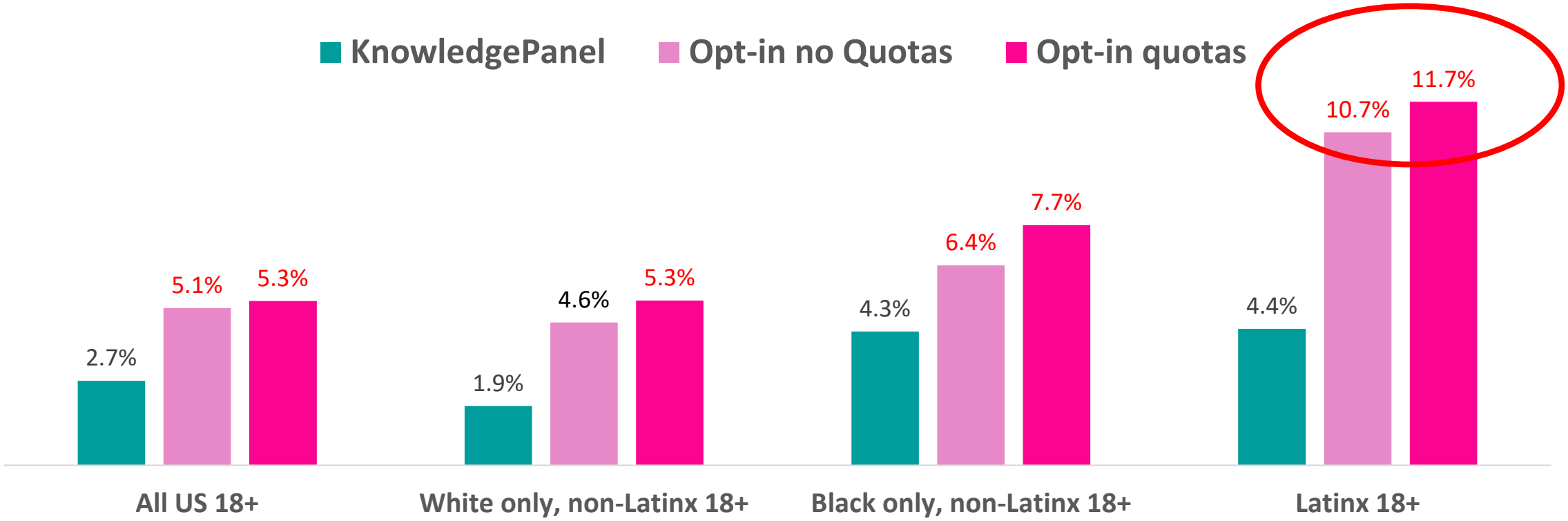
White, Non-Hispanic
Black, Non-Hispanic
Hispanic
Other, 2+ Races, Non-Hispanic

Example of Nested Weighting, nesting of categories within race-ethnicity

Black Male 18 to 44
Black Female 18 to 44
Black Male 45 to 59
Black Female 45 to 59
Black Male 60+
Black Female 60+
Hispanic Male 18 to 44
Hispanic Female 18 to 44
Hispanic Male 45 to 59
Hispanic Female 45 to 59
Hispanic Male 60+
Hispanic Female 60+
Other Male 18 to 44
Other Female 18 to 44
Other Male 45 to 59
Other Female 45 to 59
Other Male 60+
Other Female 60+

Nested Weighting – Average Deviation from Benchmarks

We found very little change in bias when expanding weighting to include adjustment nested within race/ethnicity. Bias was still higher for both Black and Hispanic participants for both types of sample (KP and Opt-in).



Conclusions and Discussion

- **It was reassuring to see how closely the panel as a whole aligned, but the results show that even among KP we can make some improvements to better represent people of color in our samples.**
- **Some things we are exploring include:**
    - **An investigation into recruitment methods to ensure most representative sample is coming in the door**
    - **Differential incentives and additional reminder protocols for some groups with lower study-level completion rates which could also help prevent attrition**
    - **Panel engagement and satisfaction survey – analyze results by race/ethnicity for any differential experiences to potentially inform recruitment and engagement**

Conclusions and Discussion

- For opt-in sample, the considerations are a bit different.
- Fit for purpose – already accepting higher bias – but sheds new light on studies that might not be appropriate with opt-in sample.
- The opt-in no quota design having a lower bias than the sample with quotas was a surprise. This needs to be replicated before drawing any conclusions.
- Also note that no quotas comes at the expense of a lower weighting efficiency.
- Opt-in samples are affected by the other samples in field at the time of the study – unlike probability-based panels which tend to follow a one sample, one survey approach, opt-in samples match the survey to the panelist at the time someone is in the survey environment.
- It's possible that differential incentives might help, but need to ensure higher incentive does not induce people to lie about demographics.

Conclusions and Discussion

- **While the additional nested weighting that we applied did not seem to reduce bias, it is possible that we need to extend weighting even further.**
- **Perhaps some of the secondary demographics might be better suited for adjustment within race/ethnicity – e.g., homeownership or marital status.**
- **Regardless of sample source, if we want to better represent the full and rich voices of people of color, these findings suggest that we need to increase our attention to issues in how respondents are recruited, who is being recruited, and how they are being retained in panels has taken on a more important priority.**

Thank you

Randall K. Thomas
Elisa Chan
Leticia Maciel
Xiaolu Yan
Ge Tang
Maya Grosul
John Lien
Fei He
Beatrice Abiero
Linda McPetrie
Liju James
Kristine Janus
Jim Bernier
Kip Davis
Raleigh Cobb
Tanya George
Robert Petrin
Annie Weber
April Jeffries
Caroline Hewett
Chris Deeney
Kathleen Santos

Thank you!

Frances M. Barlas, Ph.D.

frances.barlas@ipsos.com