

Representation in Research: Diversity and Online Samples

**Frances M. Barlas, Ph.D.,
Randall K. Thomas, and Beatrice Abiero, Ph.D.**

Ipsos Public Affairs

Study Background



In 2020, we have seen a broader awakening to the continued systemic racism throughout all aspects of our society and renewed calls for racial justice. For the survey and market research industries, this has raised questions about how well our industry does in ensuring that our public opinion research captures the full set of diverse voices that make up the United States. These questions were reinforced in the wake of the 2020 election with the scrutiny faced by the polling industry and the role that voters of color played in the election. We conducted a study to assess how well online samples represent communities of color and their diversity. While past studies have found lower bias in probability-based samples with online panels compared to opt-in samples (MacInnis et al., 2018; Yeager et al., 2011) there has been little investigation into representativeness among subgroups of interest. In Sept. 2020, we fielded parallel studies on Ipsos' probability-based KnowledgePanel which is designed to be representative of the US and an opt-in nonprobability online sample with approximately 3,000 completes from each sample source. The questionnaire included a number of measures that could be benchmarked against gold standard surveys such as the Current Population Survey, the American Community Survey, and the National Health Interview Survey. We found that across all race/ethnicity groups KnowledgePanel had lower bias than opt-in sample. In both sample sources bias was lowest among white respondents and higher among Black and Hispanic respondents. A calibration adjustment using variables related to the self-selection bias of opt-in samples helped reduce the bias in the opt-in sample across all race/ethnicity groups. We provide recommendations to improve representativeness with online samples.

Presenter



Frances Barlas, Ph.D. (she/her)



Dr. Frances Barlas is the lead methodologist for KnowledgePanel at Ipsos. She has extensive experience conducting survey research across all modes of administrations. In her current role, she is charged with overseeing and advancing the statistical integrity and operational efficiency of the KnowledgePanel.® She leads research on research that advances industry knowledge on collecting the most valid and reliable survey data. She holds a Ph.D. in Sociology from Temple University.

Study Background



- **In 2020, we saw a broader awakening to the continued systemic racism throughout all aspects of our society and heard renewed calls for racial justice.**
- **In the wake of the murders of George Floyd and Breonna Taylor, we saw widespread protests 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

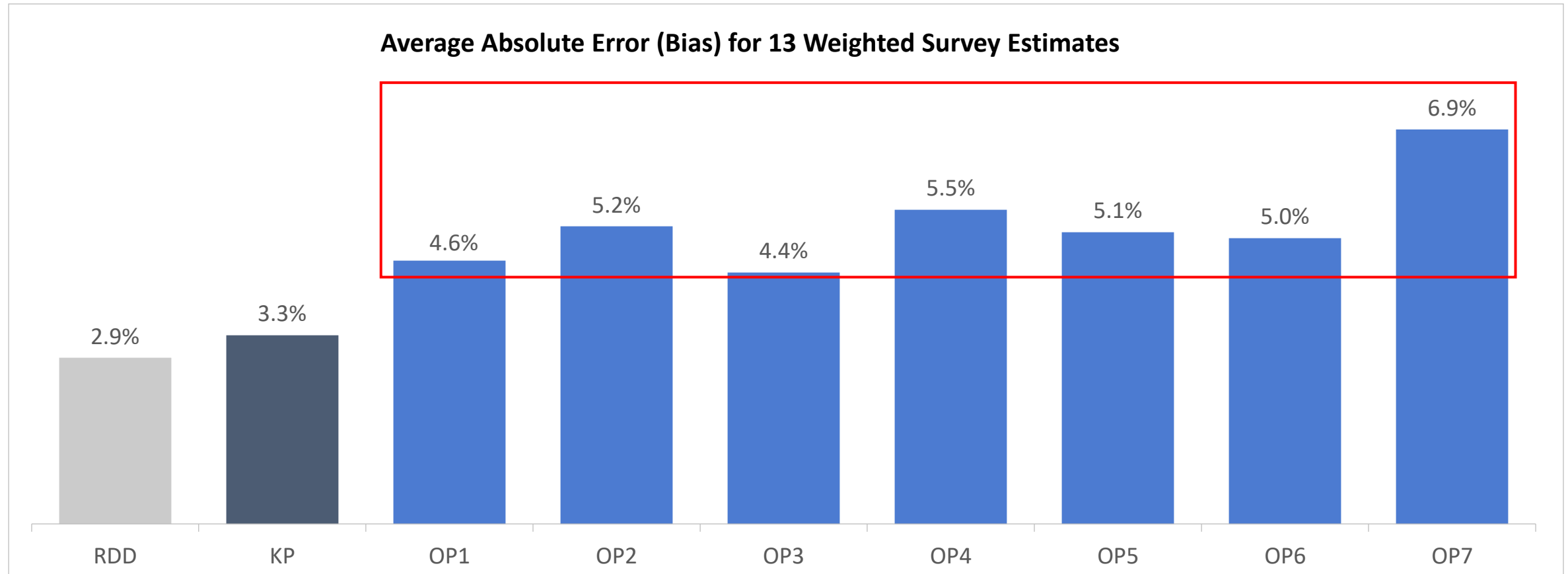


- **Survey research has played a key role in documenting the impacts of the COVID-19 pandemic on society. We know that it has had a differential impact by race and ethnicity, with Black, Hispanic, and Native Americans facing disproportionately high rates of infection, hospitalization, and death.**
- **We've 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.**
- **These questions were reinforced in the wake of the 2020 election with the scrutiny faced by the polling industry and the role that voters of color played in the election.**

Comparing Probability-based to Opt-in Data



A 2011 study comparing 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)

Study Background



- **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 level.**
- **There has been less investigation into representativeness among subgroups of interest.**
- **We wanted to assess how well online samples represent communities of color and the diversity within each community.**

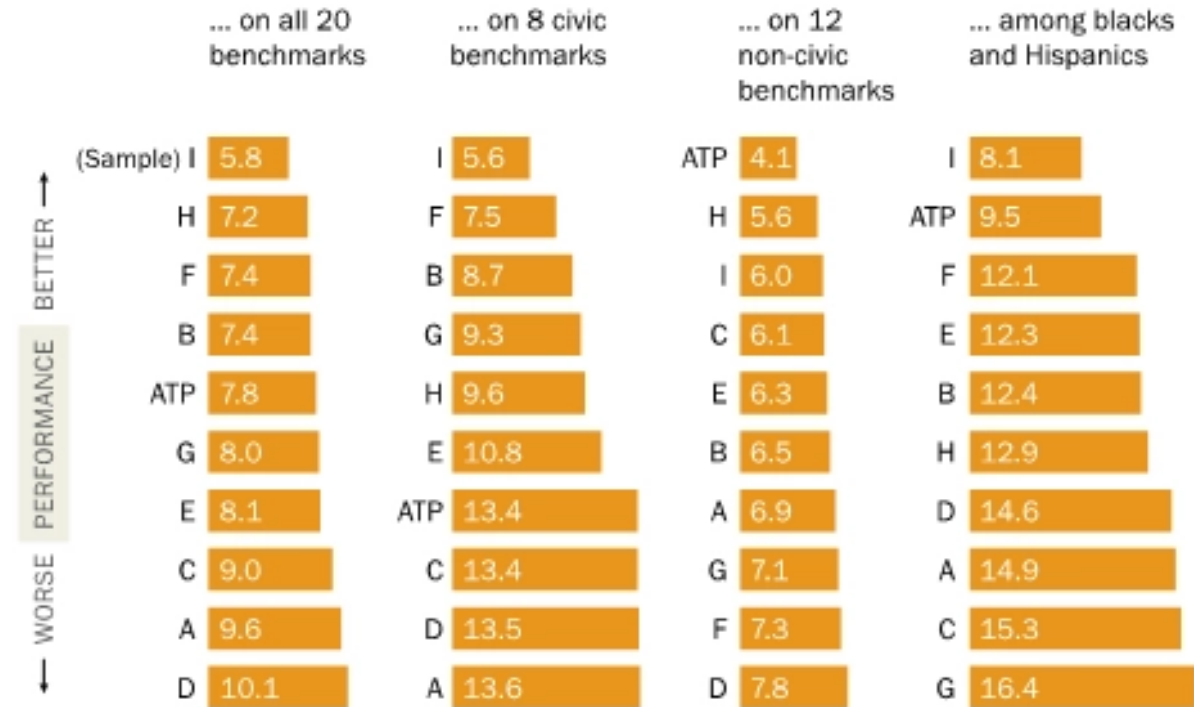
Study Purpose

- 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 at the overall level.

Notable differences in data quality across online samples

Average estimated bias in benchmarking analysis ...

Values for each sample represent the average of the absolute differences between the population benchmarks and weighted sample estimates



Source: Pew Research Center analysis of nine online nonprobability samples and the Center's American Trends Panel data. See Appendix A for details. "Evaluating Online Nonprobability Surveys."

KnowledgePanel Recruitment

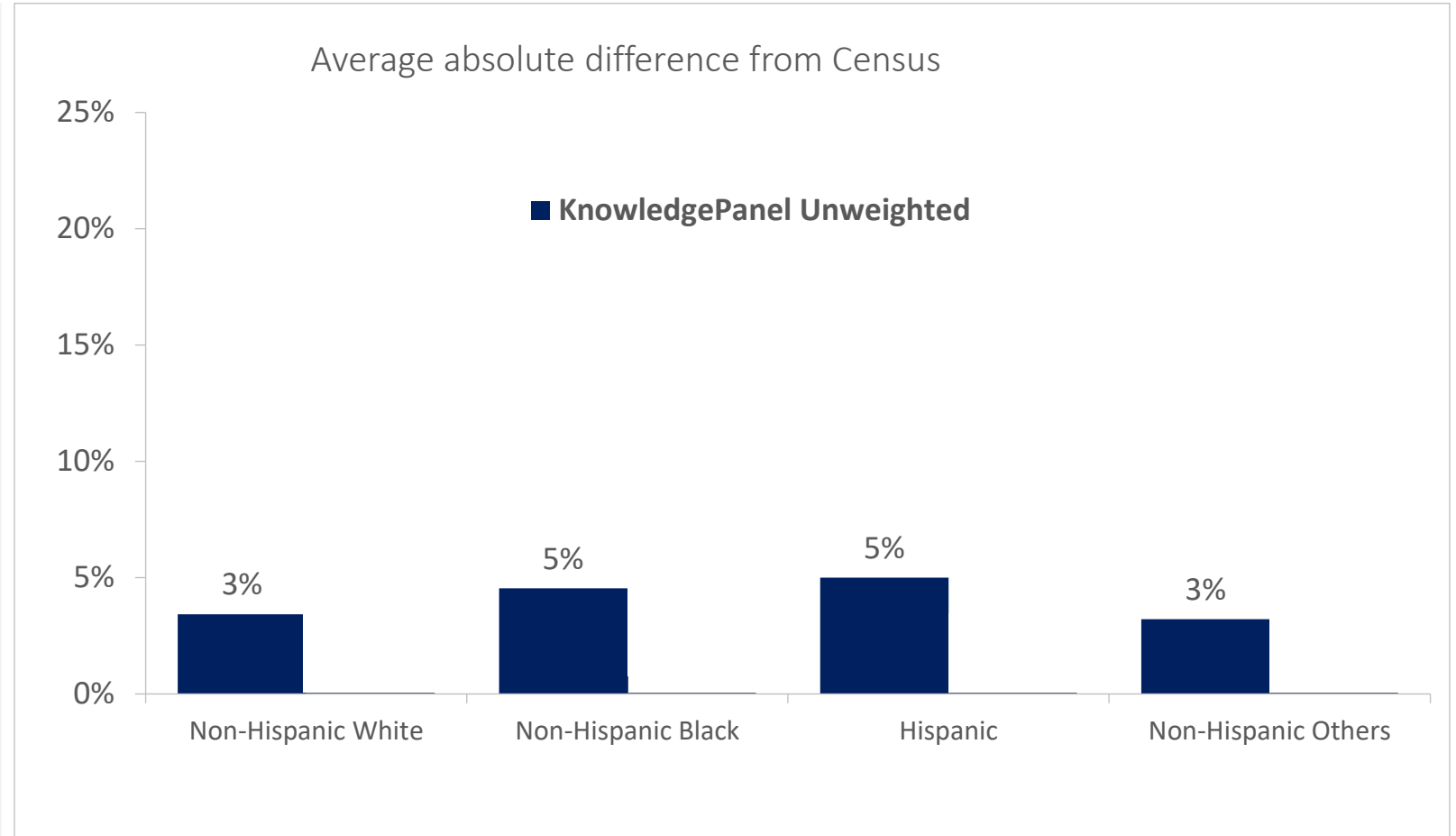


In comparing KnowledgePanel to Census benchmark data, the unweighted panel is fairly close to benchmarks on average.



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 Recruitment

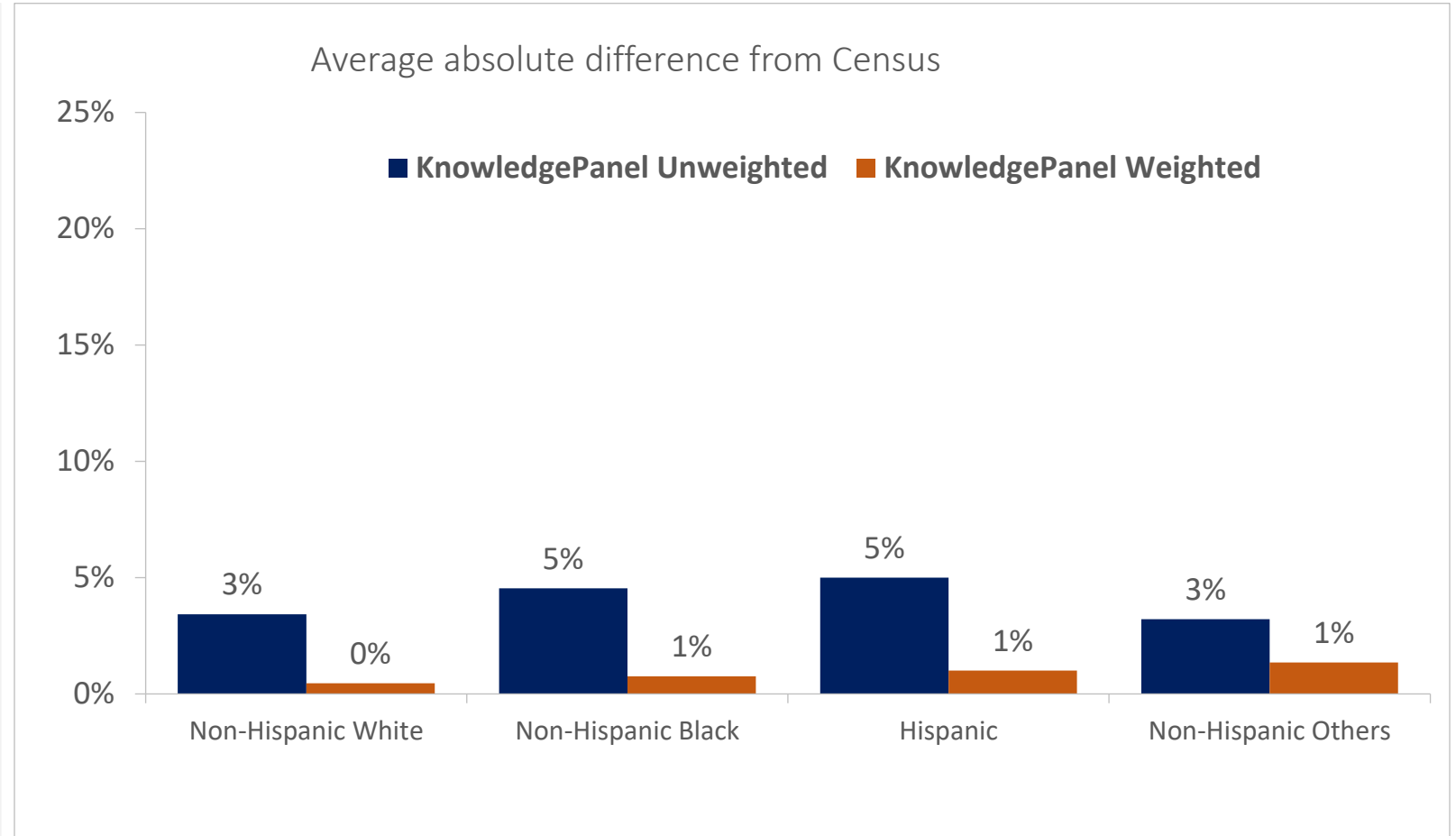


In comparing KnowledgePanel to Census benchmark data, the unweighted panel is fairly close to benchmarks on average. The panel weighting that is used in study-level sample selection brings the subgroups even more in line with benchmarks.



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



How about Survey Results and Adequacy of Weighting?



- **We wanted to see how well our general population samples reflect basic demographics of racial and ethnic subgroups after weighting.**
- **Our typical general population weighting adjusts by race/ethnicity overall, but we do not do any nested adjustments within racial/ethnic subgroups unless there are sizable oversamples of those groups.**
- **So if we look at weighted results within race/ethnicity – how well do the basic demographics align with population distributions for those groups?**

How about Survey Results and Adequacy of Weighting?

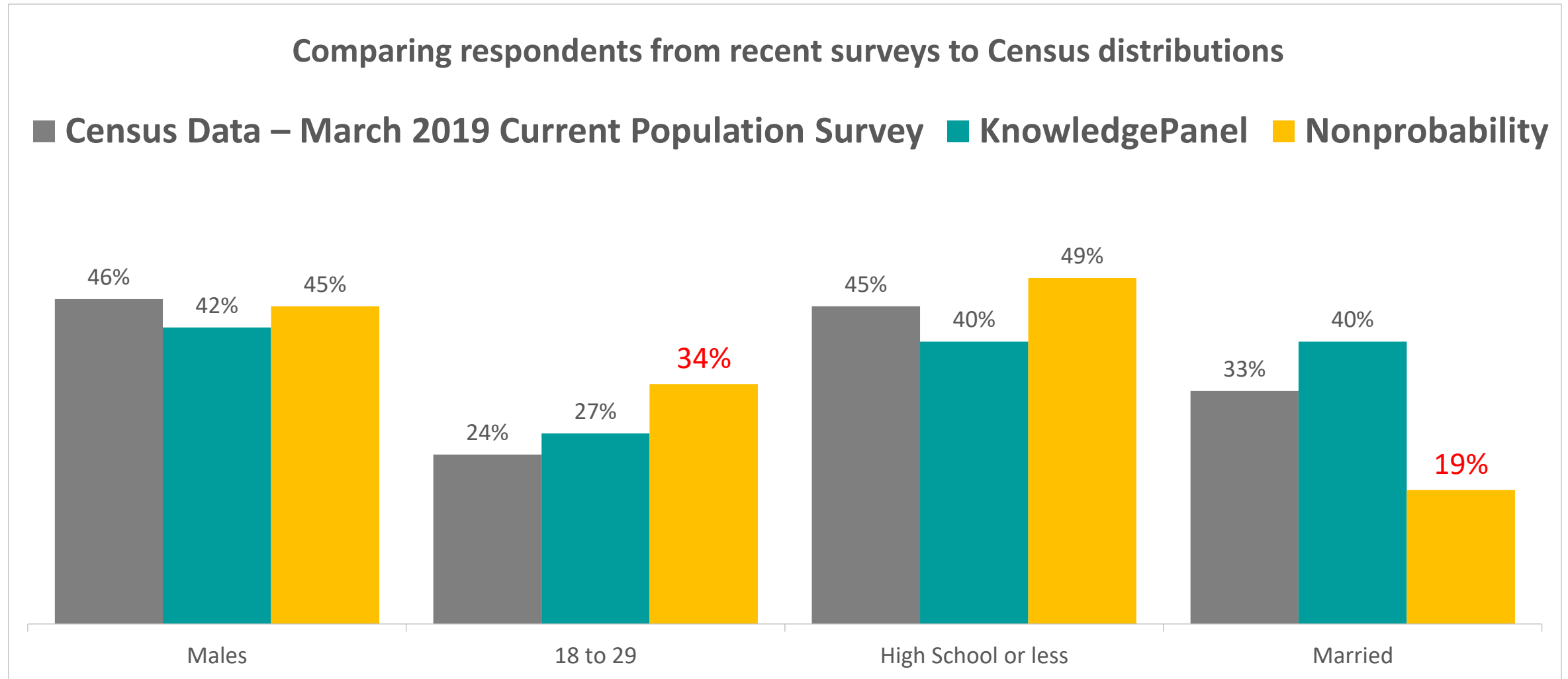


- In our first study, we compared the subsets of Black and Hispanic respondents from two general population samples –

Sample Type	Overall	Black	Hispanic
KP	1,548	145	188
Opt-in	1,215	177	186

Weighted Demographics Among Black/African American

KP respondents who were Black/African American were more representative than opt-in sample.



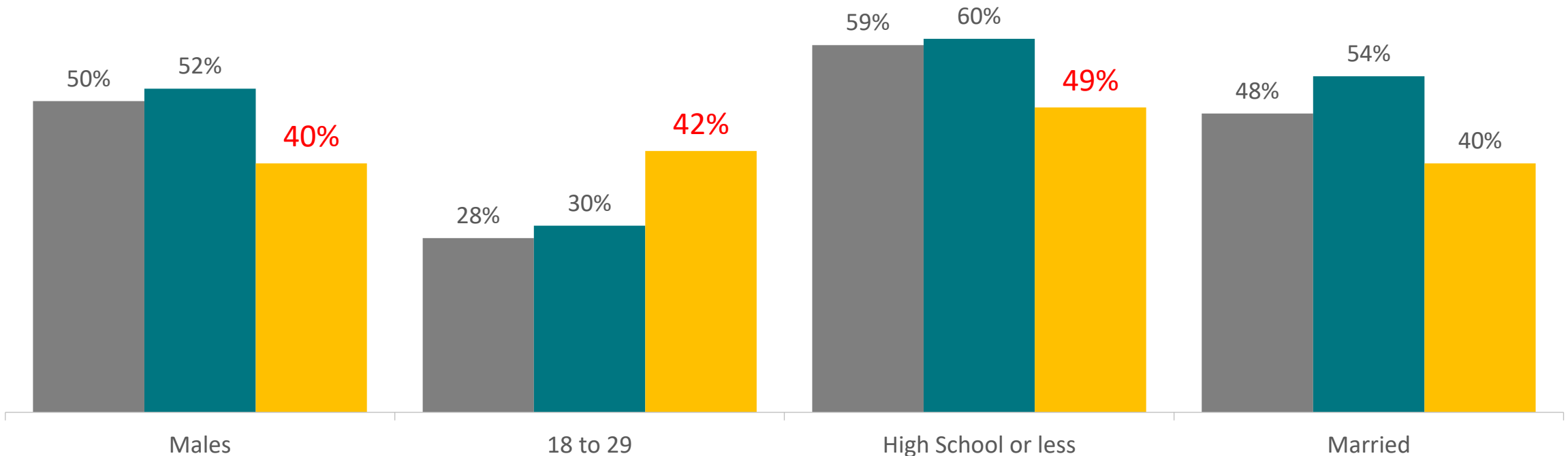
Weighted Demographics Among Hispanic/Latinx



KP respondents who were Hispanic/Latinx showed greater representativeness than those from opt-in sample.

Comparing respondents from recent surveys to Census distributions

■ Census Data – March 2019 Current Population Survey ■ KnowledgePanel ■ Nonprobability



How about Survey Results Beyond Weighting Variables?



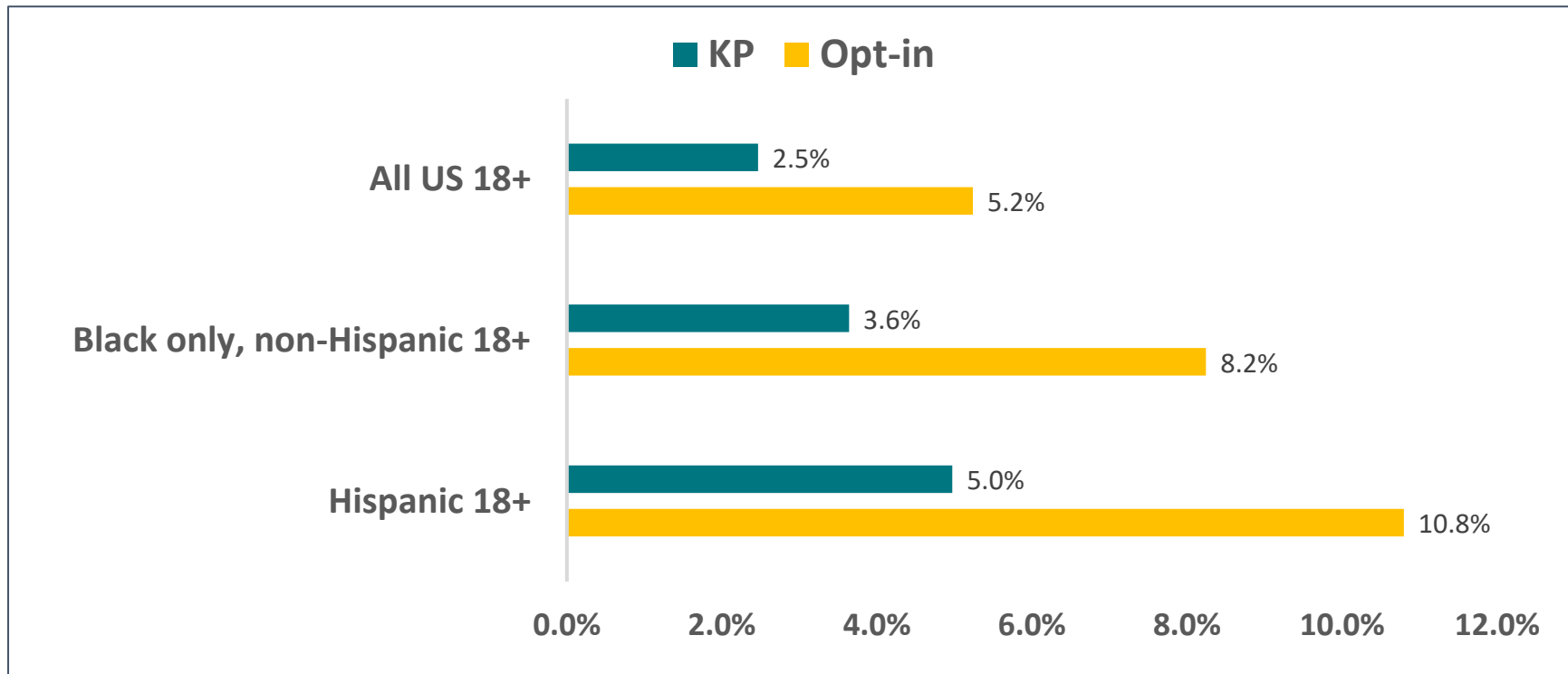
- We then designed a custom study in Sept. 2020, in which we fielded parallel surveys on both KP and opt-in
- Approximately 3,000 completes per sample type
- Length of interview (LOI) was approximately 12 minutes
- Questionnaire included 10 benchmarkable items that could be compared to Census data, including:
 - **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)**

How Well Do Survey Estimates Align with Benchmarks

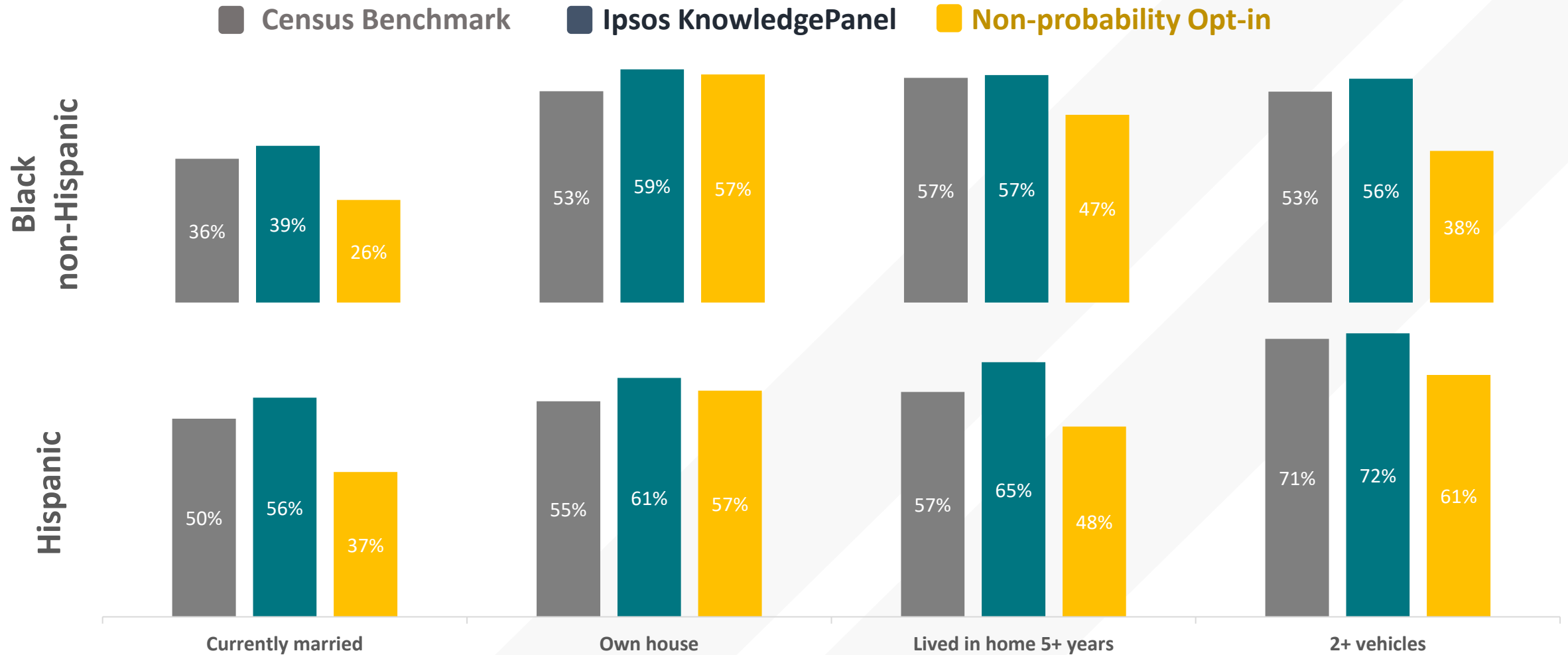


Comparing findings against 10 government demographic benchmarks – on average – KnowledgePanel has less bias than opt-in sample – it is closer to benchmarks by about 5 percentage points for Black and Hispanic respondents. In both samples, bias is higher among Black and Hispanic subgroups versus general population.

Average Deviation from 10 Census Benchmarks for Weighted Survey Results



Deeper Dive into a few of the Benchmarks



Based on a weighted comparison of sample types vs 10 government benchmark measures not used in sample weighting

Conclusions and Discussion



- **Overall, KnowledgePanel aligns fairly closely with Census data on key demographics and misalignments virtually disappear when the panel is weighted using our typical weighting used for sample selection.**
- **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 Hispanic subgroups.**
- **Both samples showed highest bias among Hispanic respondents – KP was 5% pts off and opt-in near 11% pts off on average.**

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**
 - **Panel engagement and satisfaction survey – will analyze results by race/ethnicity for any differential experiences to potentially inform recruitment and engagement**

Thank you!

Frances M. Barlas, Ph.D.
frances.barlas@ipsos.com