Novel sampling methodology combining probability sampling with online recruitment - A case study of an Internal Revenue Service (IRS) national taxpayer survey

Tim Castle
Chief, Servicewide Support Section Statistics of Income | Research, Applied Analytics & Statistics

Peter Enns
Cornell University and Verasight

Yvonne Nomizu
Pacific Consulting Group

Jake Rothschild
Verasight

IRS VERASIGHT PCG fast forward to customer success

Traditional probability-based survey methods continue to rise in cost and sometimes struggle to represent the population of interest

The Pew Research Center reports that probability-based samples significantly over-estimate voter turnout. Opt-in samples, by contrast, are extremely accurate on this dimension.

Voter turnout	Benchmark (%)	Prob. mean	Prob. 1	Prob. 2	Prob. 3	Opt-in mean	Opt-in 1	Opt-in 2	Opt-in 3
Voted (among eligible voters)	66.2		+8.9	+7.6	+8,2		-0.6	-0.4	+3.0
Average absolute error		*8.2*	*8.9*	*7.6*	*8.2*	*1.3*	*0.6*	*0.4*	*3.0*

Note: The benchmark for turnout in the 2020 presidential election comes from the U.S. Elections Project as of May 16, 2021. The benchmarks for presidential vote choice are from the Federal Election Commission.

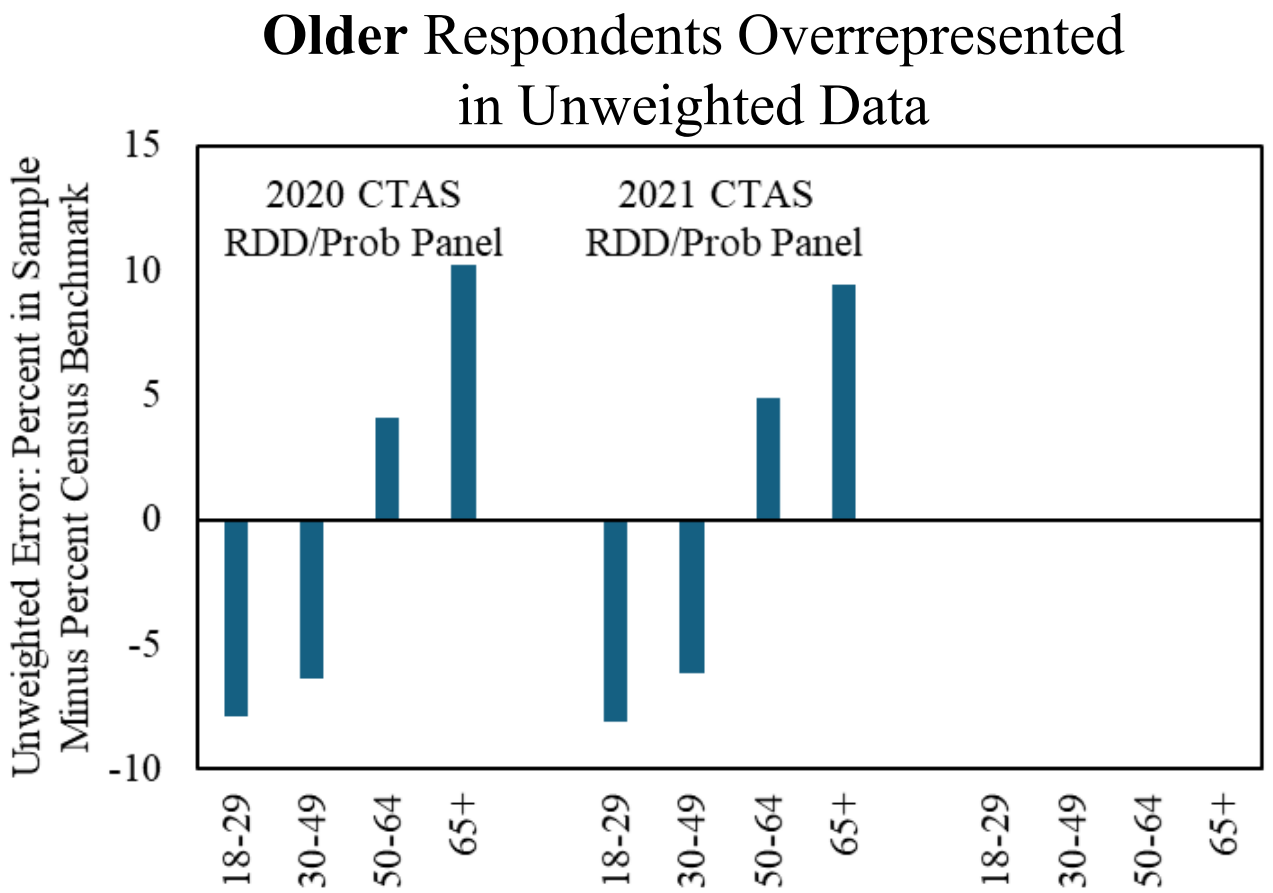
Pew Research Center: Andrew Mercer and Arnold Lau (2023)

Traditional probability-based survey methods continue to rise in cost and sometimes struggle to represent the population of interest

Traditional probability-based survey methods continue to rise in cost and sometimes struggle to represent the population of interest

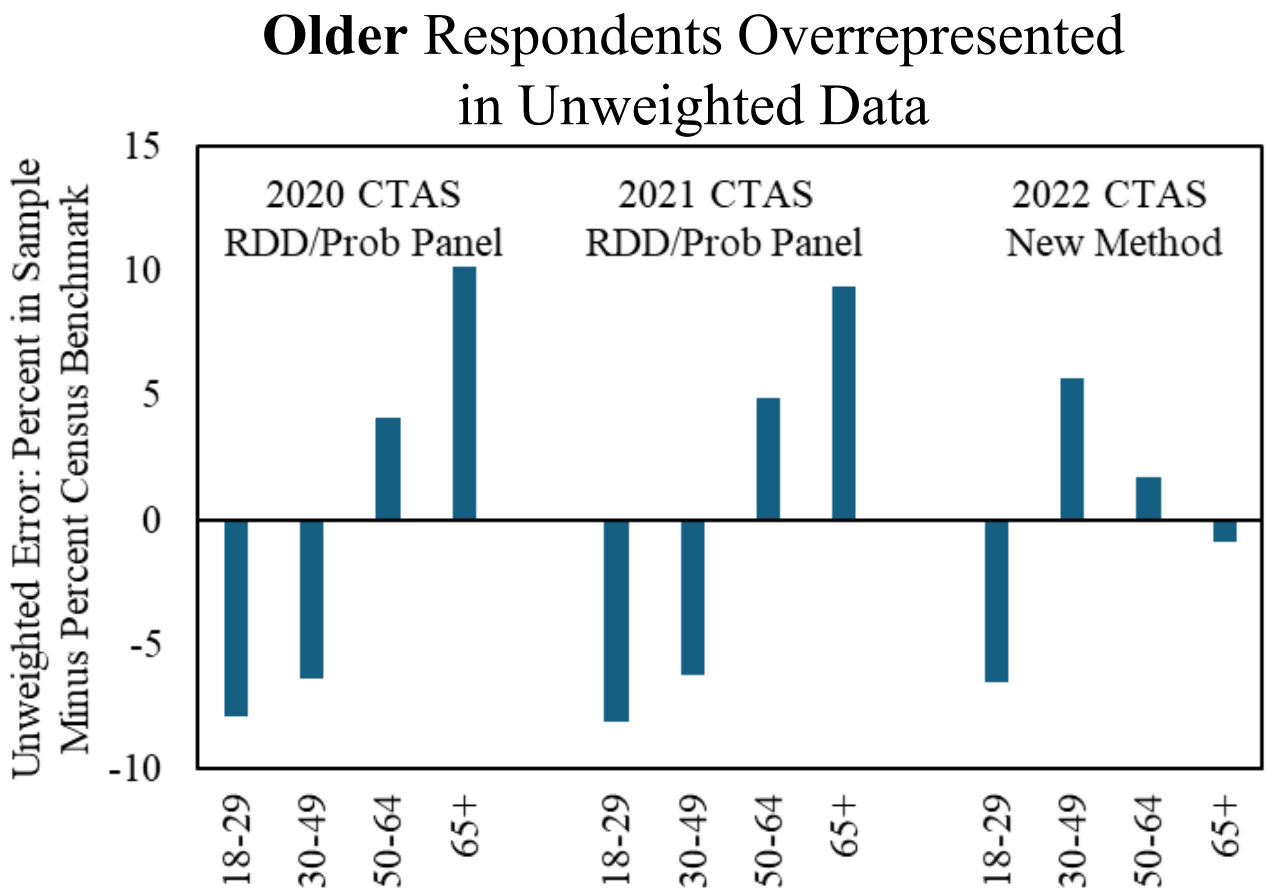
- Comprehensive Taxpayer Attitude Study (CTAS) is a long running survey to understand the taxpayer experience
- While post-stratification weights ameliorate any concerns, these patterns suggest traditional approaches will increasingly struggle to reach certain groups

*RDD: Random digit dialing



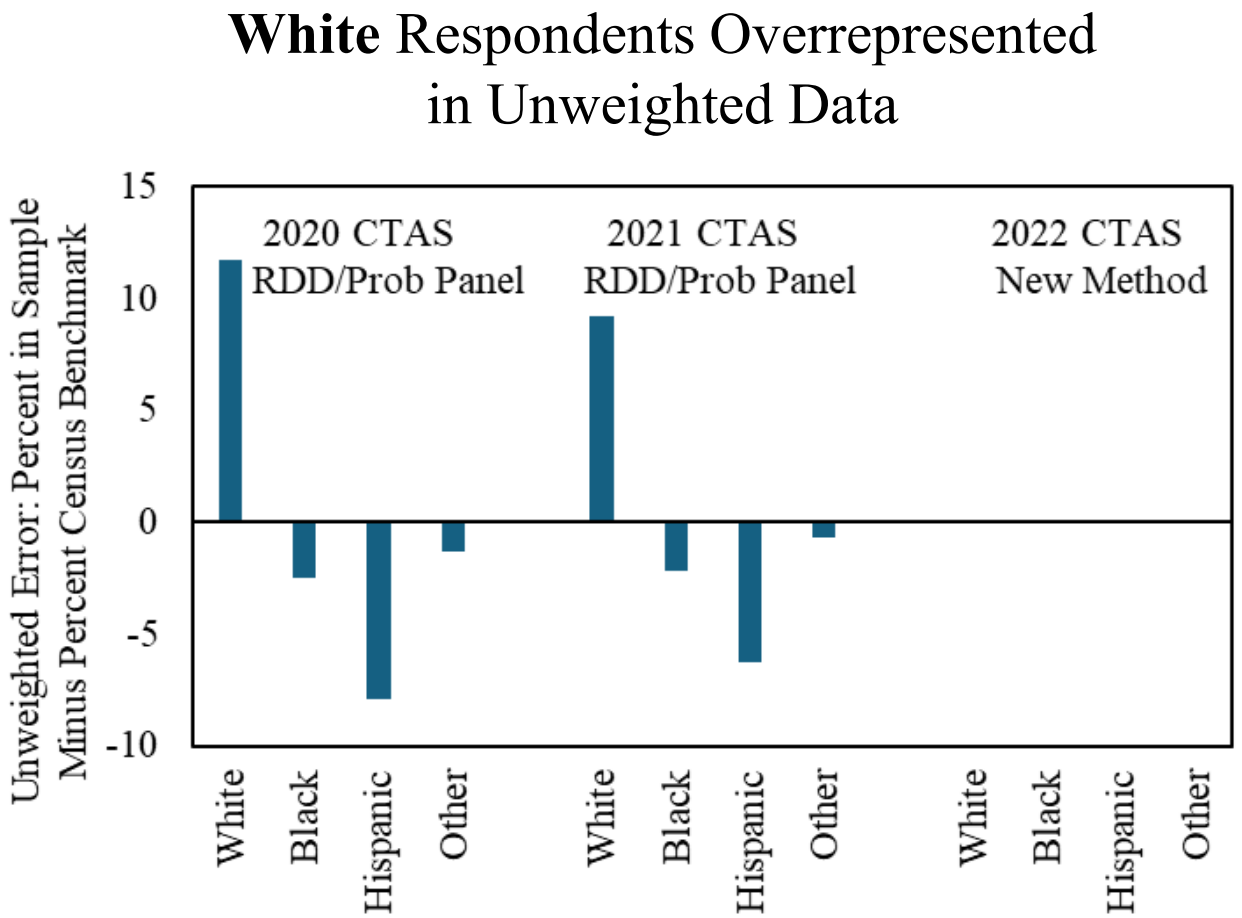
Traditional probability-based survey methods continue to rise in cost and sometimes struggle to represent the population of interest

- Comprehensive Taxpayer Attitude Study (CTAS) is a long running survey to understand the taxpayer experience
- While post-stratification weights ameliorate any concerns, these patterns suggest traditional approaches will increasingly struggle to reach certain groups



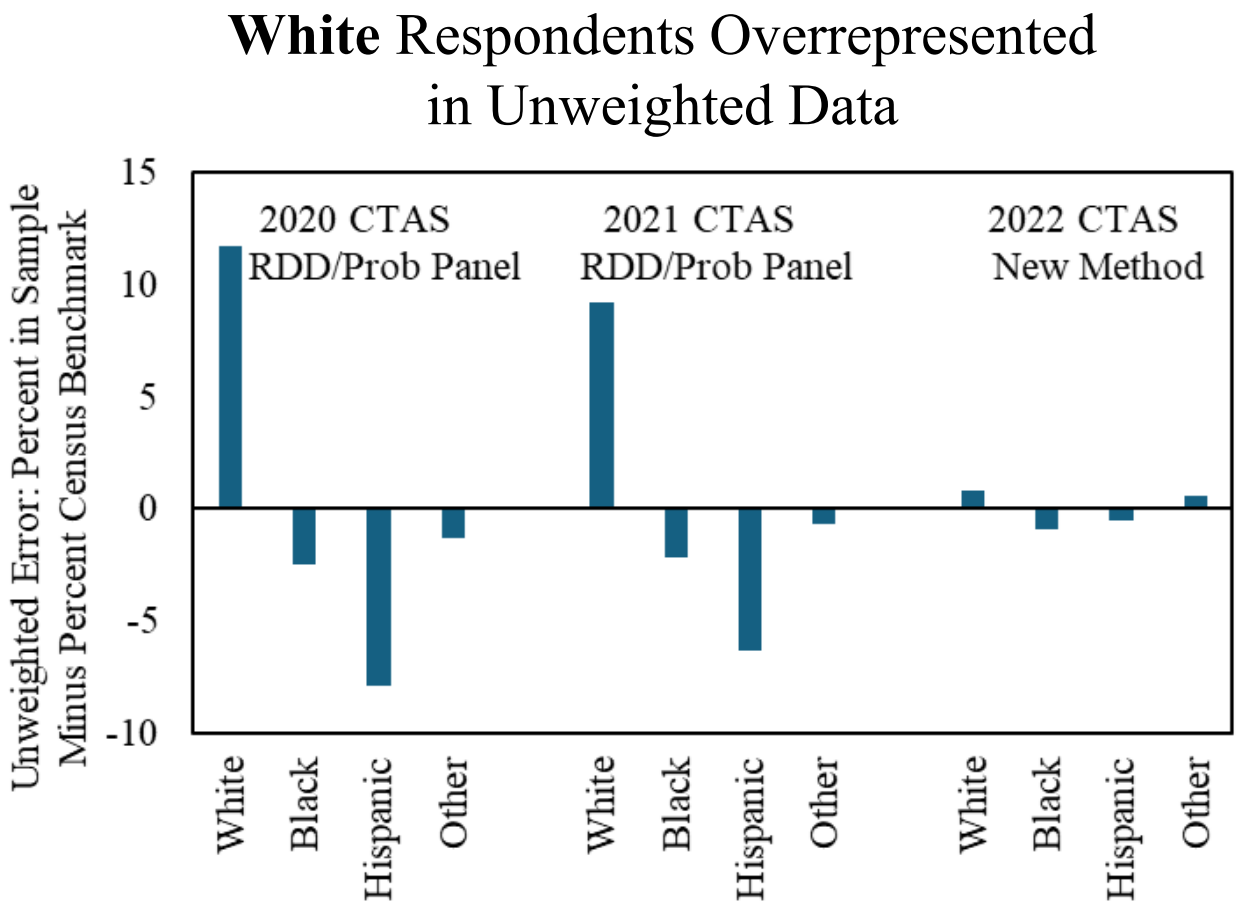
Traditional probability-based survey methods continue to rise in cost and sometimes struggle to represent the population of interest

- Comprehensive Taxpayer Attitude Study (CTAS) is a long running survey to understand the taxpayer experience
- While post-stratification weights ameliorate any concerns, these patterns suggest traditional approaches will increasingly struggle to reach certain groups



Traditional probability-based survey methods continue to rise in cost and sometimes struggle to represent the population of interest

- Comprehensive Taxpayer Attitude Study (CTAS) is a long running survey to understand the taxpayer experience
- While post-stratification weights ameliorate any concerns, these patterns suggest traditional approaches will increasingly struggle to reach certain groups



Our solution: The IRS, PCG, and Verasight partnered to implement a sampling strategy to obtain a representative sample and reach underserved populations in a cost-effective manner

- Continue the multi-sample strategy
- Expand the sampling strategy to include probability (random ABS) and nonprobability samples
  - ABS - allows 1-800 calls in English and Spanish (Inbound IVR).
  - Ensure the nonprob samples are representative and high quality in two ways:
    - Utilize a panel that:
      - Recruits respondents from diverse sources
      - Every respondent is verified in multiple ways
      - Limits how many surveys respondents can take
      - Never routes respondents from one survey to the next
      - Never outsources data collection
      - Incentivizes respondents for every survey they start, regardless of whether or not they qualify
    - Utilize Dynamic Online Targeting

ABS: Address-based sampling
IVR: Interactive Voice Response

Dynamic Online Targeting

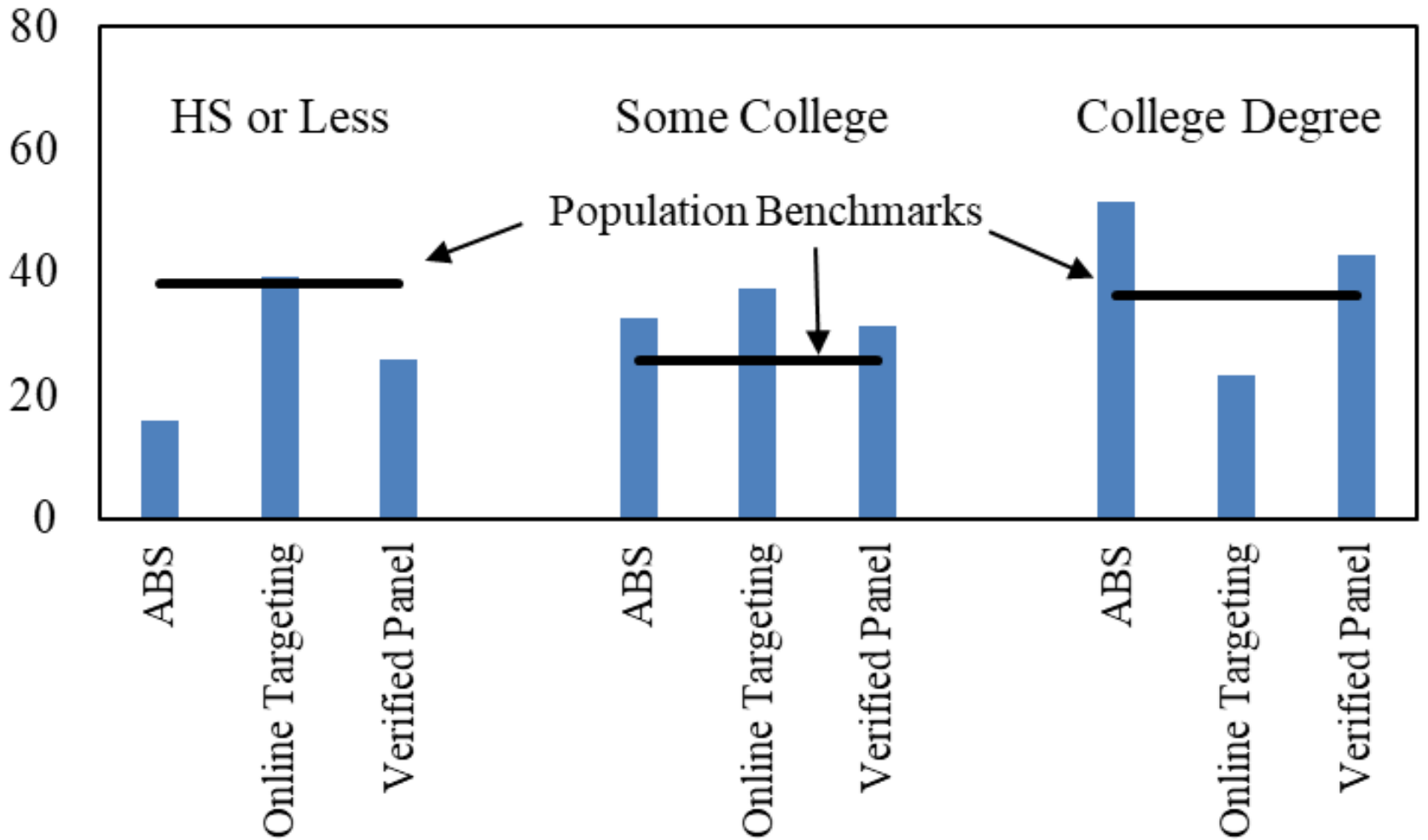
Sampling Frame

- U.S. adults who are active on social media websites (e.g., Facebook, Instagram, TikTok) or search websites (e.g., Google, Bing) during the fielding period.
- >95% of Americans have access to the internet and 97% between the ages of 18 and 49 have a smartphone (Pew Research Center 2024)

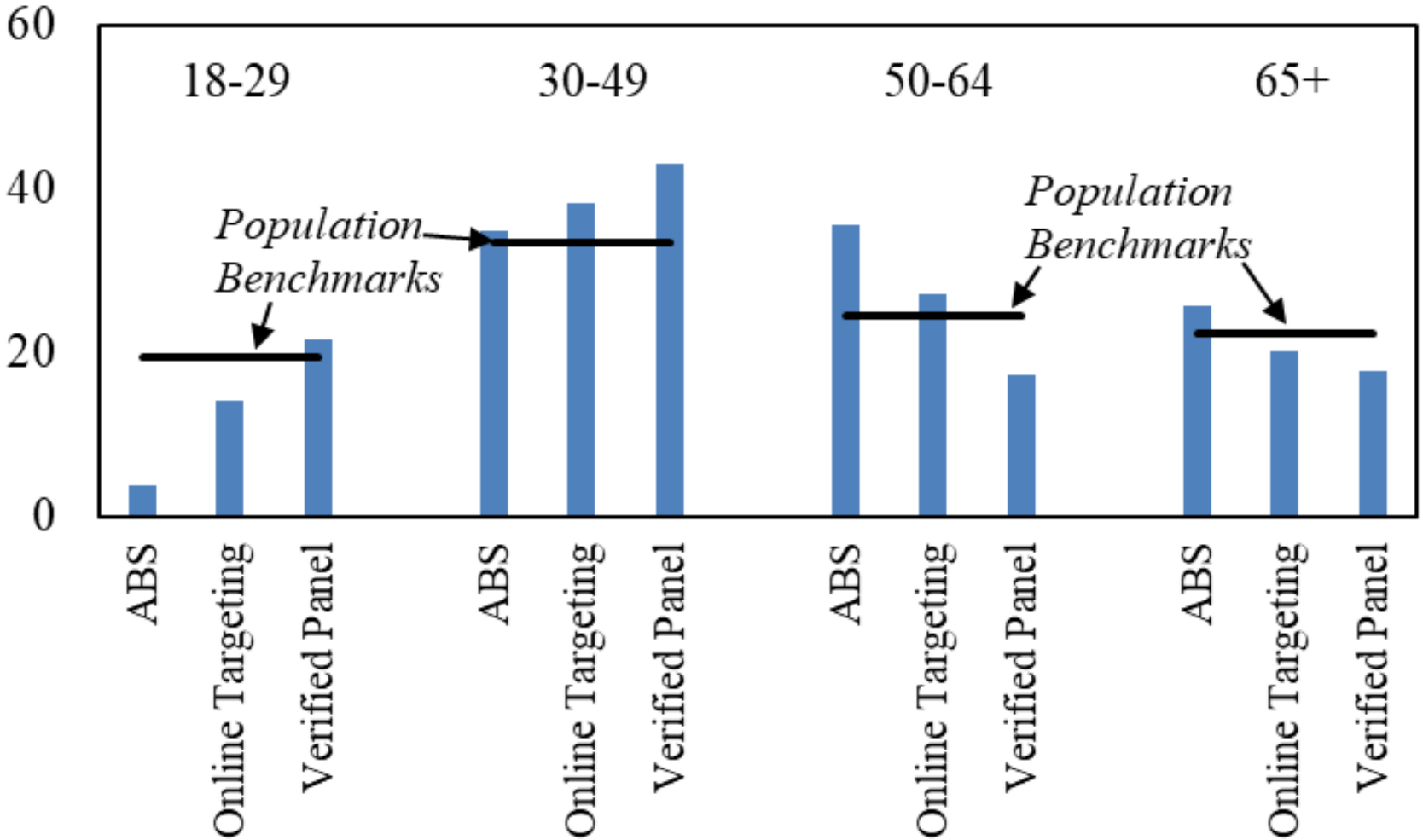
Benefits/Innovations

- Survey invitations targeted by online location data (e.g., zip code), social media registration data (e.g., age, gender, education), and predicted behavioral data (e.g., U.S. adults with browsers set to Spanish are likely to be Hispanic)
- Dynamically adjust survey invitations to reach oversamples and to recruit a representative sample.

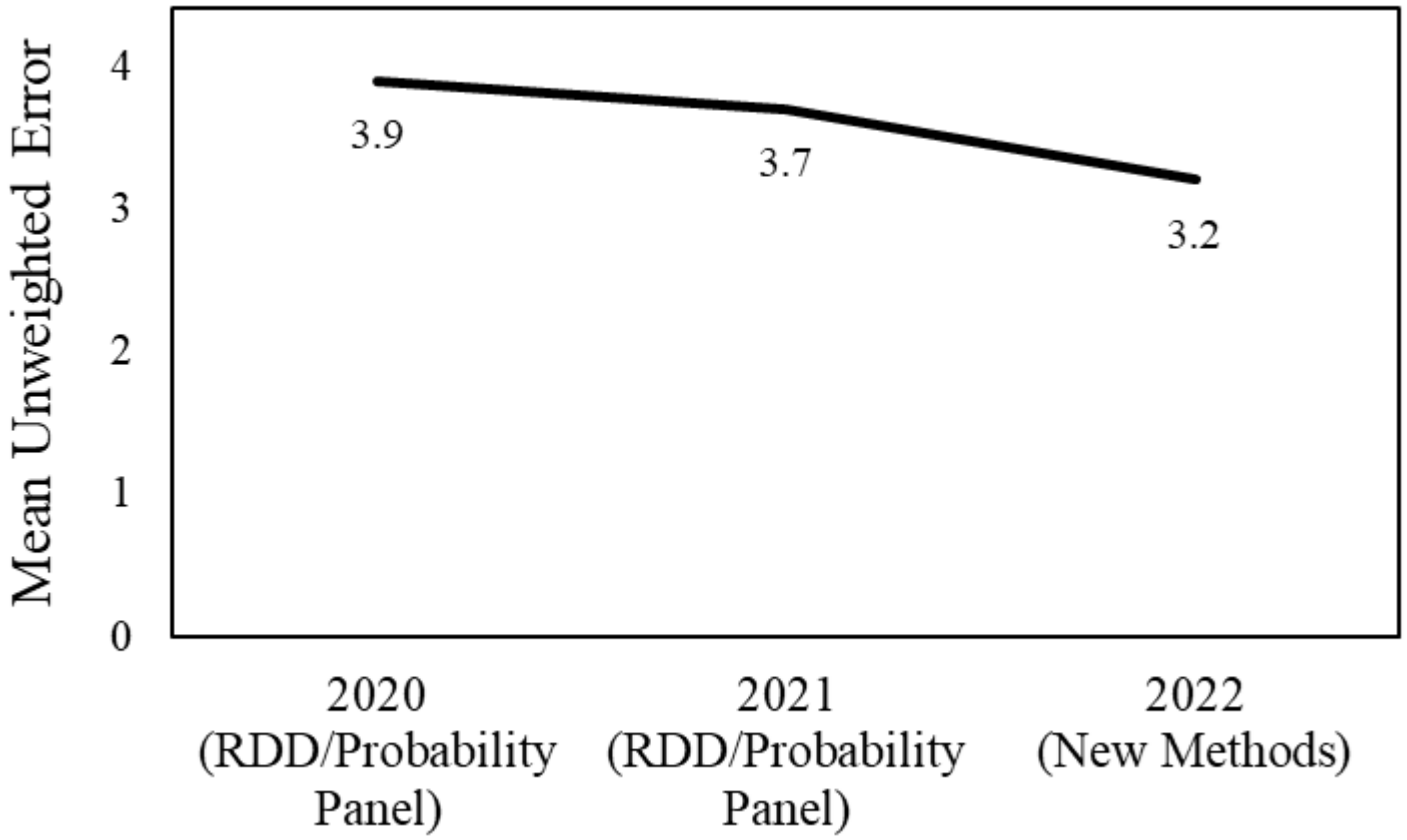
Combining Samples Improves Population Coverage



Combining Samples Improves Population Coverage



The New Approach Improved Population Coverage

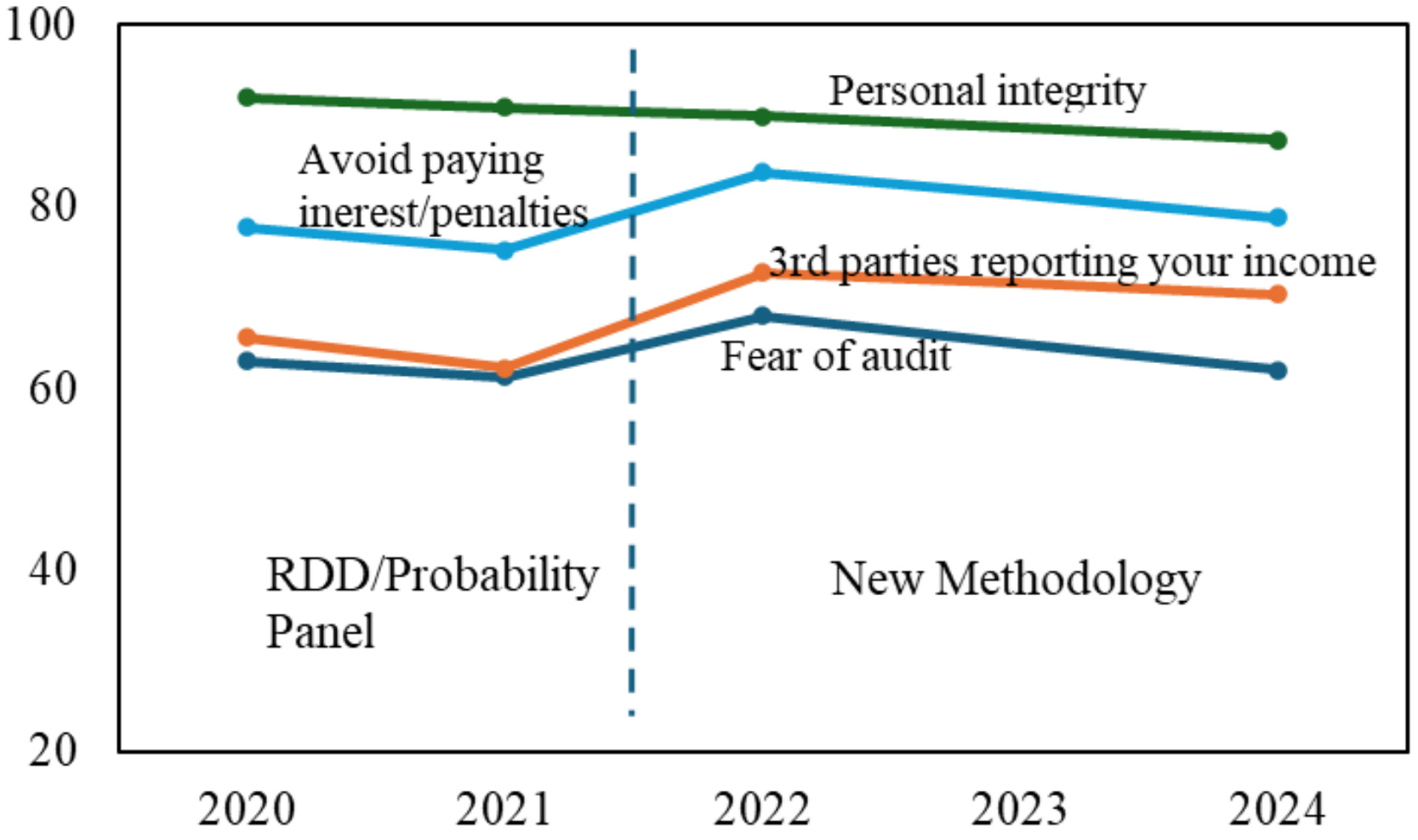


Indications that the New Methods Continue to Accurately Measure Tax-Related Attitudes

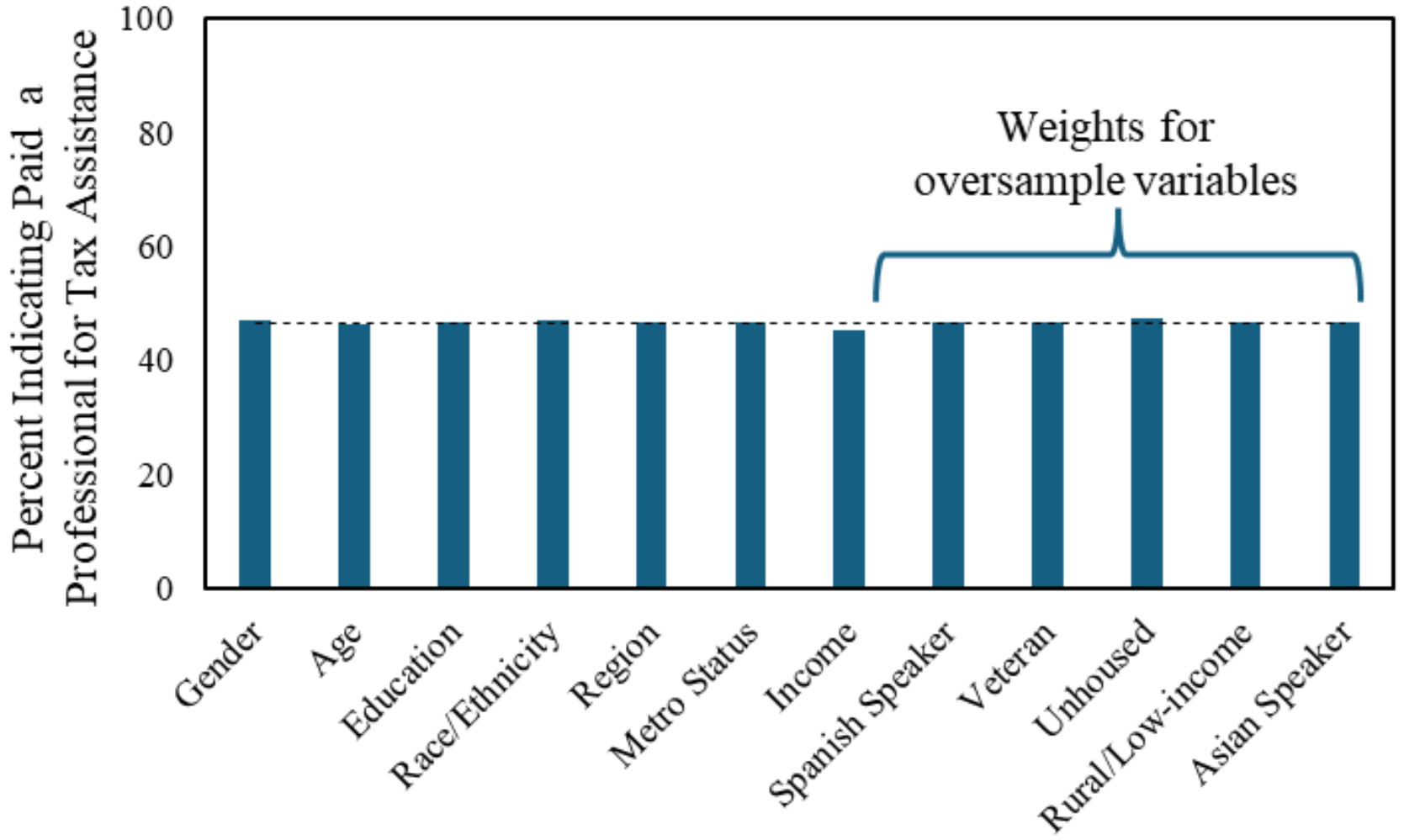
- Attitudes can *change*. This is a primary reason for an ongoing study like CTAS.
- However, we would expect longstanding patterns to persist across methods if both methods reliably measure the same construct.
- A longstanding finding in CTAS is that *Personal Integrity* is the most influential factor in why respondents say they report and pay their taxes honestly. *Fear of an audit* is among the least important factors.

Indications that the New Methods Continue to Accurately Measure Tax-Related Attitudes

How much influence does each of the following factors have on whether you report and pay your taxes honestly?



Results are *Not* Sensitive to Weighting Strategy Used



Weights recalculated dropping each variable

Expand IRS's ability to hear from diverse and underrepresented groups

2022 oversamples focused on Americans with limited English proficiency and low incomes

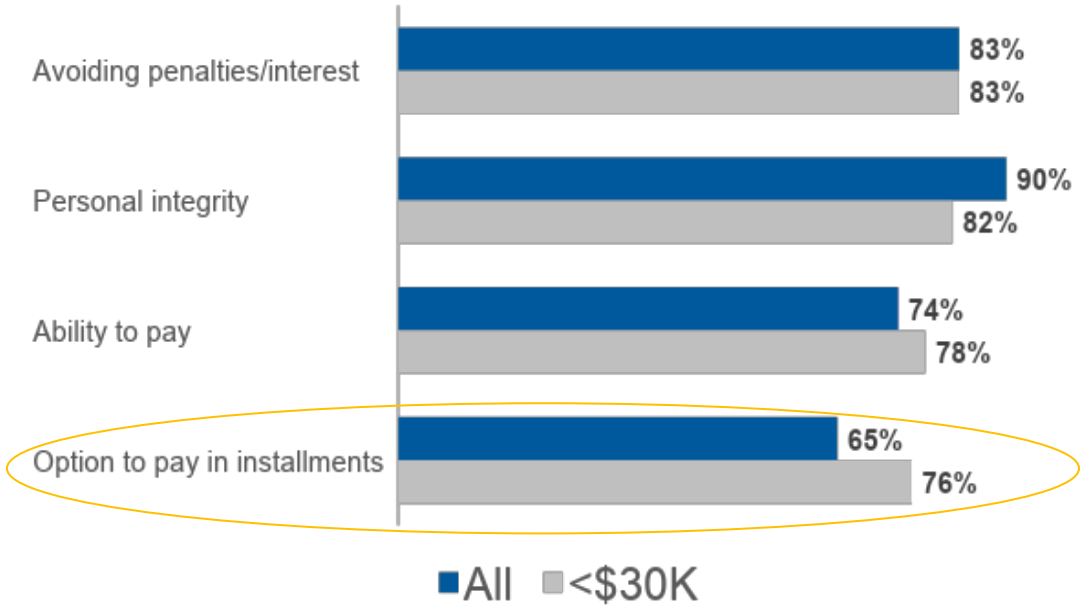
Demographic	Total in Sample
Spanish Limited English Proficiency	101
Asian Language Speakers	70
Unhoused	37
Rural, Low-Income	97
Stateside Military	19
Veterans	120

These groups were targeted to increase responses and thus statistical power to identify potential differences across groups

Expand IRS's ability to hear from diverse and underrepresented groups

Oversample of Americans with annual income <$30,000 2022 CTAS

Top Factors Influencing Whether Taxpayers Report and Pay Taxes Honestly (% "A Great Deal" or "Somewhat" Influence) (Q4)



Conclusions

- Despite the rising costs of probability-based sampling methods, in some contexts these methods struggle to represent the population of interest
- Less expensive methods, also present challenges
- CTAS has benefited from combining samples in the past and is pushing farther in that direction
- Numerous indicators of the efficacy of this approach
  - Improved population coverage (reduced unweighted error)
  - Long term patterns in the data persist, showing stability (where expected) across methods
  - Conclusions insensitive to weighting strategy
  - Enhanced ability to oversample
  - Cost effective

Conclusions and Applications

- Multiple samples can reduce costs and improve population representation over traditional probability-based samples
  - If the budget is sufficient and current probability-based samples align with population benchmarks, continuing with current approaches might be optimal
  - When budgets are limited, diversifying sample type offers potential benefits at a *reduced* cost
  - Further, if certain groups, such as the less educated or younger respondents are increasingly harder to reach, diversifying sample type can provide a solution
- Diversifying sampling strategies can also expand the ability to reach and oversample certain groups in a cost-effective way.

Appendix

The IRS partnered with PCG and Verasight to offer a multi-sample solution

MULTI-STRATEGY SAMPLING APPROACH

	Address-Based Sampling	Dynamic Online Targeting	Verasight Panel
Sampling Frame	Non-institutionalized U.S. adults with a mailing address via the United States Postal Service database	U.S. adults who are active on social media websites (e.g., Facebook, Instagram, TikTok) or search engines (e.g., Google, Bing) during the fielding period	U.S. adults who were previously recruited to Verasight surveys via address-based sampling and dynamic online targeting
Benefits	Multiple Completion Options including QR code, website link, inbound IVR. Inbound IVR enables respondents to complete the survey 24/7 at their chosen time vs. outbound telephone or IVR Ability to oversample rural or low-income areas by incorporating zip-code level census data with address file	Survey invitations targeted by online location data (e.g., zip code), social media registration data (e.g., age, gender, education), and predicted behavioral data (e.g., U.S. adults with browsers set to Spanish are likely to be Hispanic) Dynamically adjust survey invitations to reach oversamples and to recruit a representative sample	Active verification of all panelists via U.S. mobile-phone number and passive verification via response consistency Capped survey opportunities to minimize respondent fatigue panel learning effects Target panelists based on demographic data to oversample hard-to-reach groups at a lower cost