

TECHNICAL OVERVIEW OF THE AMERISPEAK® PANEL NORC'S PROBABILITY-BASED HOUSEHOLD PANEL

Updated February 8, 2022

This technical overview provides the basic information about AmeriSpeak®, a large probability-based panel funded and operated by NORC at the University of Chicago. AmeriSpeak is designed to be representative of the U.S. household population, including all 50 states and the District of Columbia. U.S. households are randomly selected with a known, non-zero probability from the NORC National Frame as well as address-based sample (ABS) frames, and then recruited by mail, telephone, and by field interviewers face to face. AmeriSpeak panelists participate in NORC studies or studies conducted by NORC on behalf of governmental agencies, academic institutions, the media, and commercial organizations.

The construction of the AmeriSpeak panel started in 2014 with pilot samples. In 2015, about 7,000 households were recruited from a sample of around 60,000 addresses. In 2016, about 128,000 addresses were sampled to expand the panel to around 20,000 recruited households. About 51,000 addresses were selected for the 2017 recruitment, which led to the expansion of the regular AmeriSpeak panel to 23,000 recruited households. The AmeriSpeak Panel expanded to approximately 30,000 households in 2018 and 35,000 households in 2019 through further recruitment efforts. The current panel size is 54,001 panel members aged 13 and over residing in over 43,000 households.

In addition to the regular panel for general population studies, AmeriSpeak also contains sub-panels to support studies of special populations, including AmeriSpeak Latino, AmeriSpeak Teen, and AmeriSpeak Young Adult 18-34 (which features an oversample of African Americans, Hispanics, and Asians). AmeriSpeak is also the probability sample source for TrueNorth®, the NORC calibration solution for combining probability and non-probability samples for estimation through small area modeling that leverages data from AmeriSpeak, the American Community Survey, Current Population Survey, and other data sources for improved statistical efficiency.¹ AmeriSpeak is also the sample source for the Foresight 50+ panel, which is a partnership between AARP and NORC that provides a high-quality panel for organizations looking for insights from older adults living in the United States.²

Panel Sample Frame

The primary sampling frame for AmeriSpeak is the 2010 NORC National Frame, a multistage probability sample that fully represents the U.S. household population. We provide a brief description of how the National Frame was constructed after the 2010 Census.

The primary sampling units (PSUs) in the first stage sample selection are 1,917 National Frame Areas (NFAs), each of which is an entire metropolitan area (made up of one or more counties), a county, or a group of counties with a minimum population of 10,000. A total of 126 NFAs are selected in the first stage, including 38 certainty NFAs, 60 urban NFAs, and 28 non-urban NFAs. The largest 38 NFAs, those with a

¹ For more information about TrueNorth, see <http://amerispeak.norc.org/our-capabilities/Pages/TrueNorth.aspx>.

² For more information about Foresight 50+, see <https://www.norc.org/Research/Capabilities/Pages/Foresight50.aspx>

population of at least 1,543,728 (0.5 percent of the 2010 Census U.S. population), were selected into the National Frame with certainty.

Within the 126 selected NFAs, the secondary sampling units (SSUs) are segments defined from Census tracts or block groups, where each segment contains at least 300 housing units according to the 2010 Census. Within the certainty NFAs, a sample of 896 segments was selected using systematic PPS sampling, where the size of a segment is the number of housing units. Implicit stratification was achieved by sorting the segments by location (NFA, state, and county), principal city indicator, and by ethnic and income indicators. From each urban and rural NFA, a sample of 8 and 5 segments was selected, respectively, using systematic PPS sampling where the measure of size is the number of housing units per segment. A total of 618 segments are selected from the non-certainty NFAs³. Overall, a stratified probability sample of 1,514 segments was selected into the National Frame in the second stage sampling.

Within the selected segments, all housing units are listed using the U.S. Postal Service Delivery Sequence File (DSF). In the 123 segments where the DSF coverage is deemed inadequate, the DSF address list is enhanced with an in-person field listing to improve coverage. The final National Frame, consisting of all listed households in the sample segments, is estimated to provide over 97 percent coverage of the U.S. household population. It contains almost 3 million households, including over 80,000 rural households that are added through the in-person listing. In addition to NORC's National Frame, the DSF is used as a supplemental sample frame in four states. Although nationally representative, the National Frame does not include households from Alaska, Iowa, North Dakota, and Wyoming. Since 2016, the annual panel recruitment sample has included a small address-based sample from these four states to assure AmeriSpeak presence in all U.S. States and Washington, D.C.

In 2017, an enhanced DSF frame was also used to develop a new Latino Panel with adequate representation of Spanish-language-dominant Hispanics. Census tracts with a high incidence (at least 30%) of Spanish-dominant Hispanics were targeted for this recruitment. Furthermore, within these Census tracts, households that were flagged as Hispanic based on consumer vendor data (that are typically used for direct-mail marketing) were oversampled.

Panel Sample Selection

For panel sample selection between 2014 and 2018 and in 2020, National Frame segments were stratified into six sampling strata based on the race/ethnicity and age composition of each segment, as below:

- Hispanic, high youth segments
- Hispanic, not high youth segments
- Non-Hispanic Black, high youth segments
- Non-Hispanic Black, not high youth segments
- Other, high youth segments
- Other, not high youth segments

Hispanic segments are those where Hispanics make up at least a third of the population and the Hispanic share in the population is greater than that of non-Hispanic Black. Similarly, non-Hispanic Black segments are those where non-Hispanic Black make up at least a third of the population and the non-Hispanic Black share in the population is greater than that of Hispanics. Finally, High Youth refers to segments in which 18-24-year-old adults are at least 12% of the total adult population. The above stratification is used to oversample housing units in areas with a higher concentration of young adults, Hispanics, and non-Hispanic African

³ A sample of 5 segments was selected from each of the 28 non-urban NFAs. However, 2 sample segments were later subsampled out in Montana due to cost.

Americans. The resulting household sample is referred to as the initial AmeriSpeak sample or sample for initial panel recruitment.

To support the second stage of panel recruitment, initially sampled but nonresponding housing units are subsampled for a nonresponse follow-up (NRFU)⁴. At this stage, consumer vendor data are matched to the pending housing units, and housing units that are flagged as having a young adult⁵ (18-34 years of age) or minority (Hispanic⁶, non-Hispanic Black⁷) are oversampled for the NRFU sample. Overall, approximately one in five initially nonresponding housing units are subsampled for NRFU using the same six sampling strata defined above. Due to NRFU, these initially nonresponding housing units have a higher selection probability compared to the housing units that were recruited during the first stage of panel recruitment.

A two-phase state-based ABS sample design was used for the 2019 AmeriSpeak recruitment. NORC's National Frame is designed to represent the U.S. household population nationally. At the state level, however, the panel may have more significant clustering effects from the use of the National Frame, especially for states with a small population. The primary objective of the 2019 design is to improve state-level representation by selecting the recruitment sample mostly from areas that are outside the National Frame. A stratified systematic sample was selected in the first phase, where each state constitutes a sampling stratum, and the sample was allocated to the strata proportional to the square root of the state population. In the second phase, young adults, Hispanic, non-Hispanic Black, and conservatives are oversampled based on commercial data sources to improve their representation in the panel. Because the 2019 design did not use NRFU face-to-face recruitment, the 2019 design did not involve geographic clustering.

In 2020 we returned to the “standard” sampling strategy employed in 2014 through 2018, with intentions to conduct a robust NRFU. However, the COVID-19 pandemic prevented NORC from utilizing field interviewers and the NRFU was limited to its usual first stage, a Federal Express mailing to 20% of the total sample. After an analysis of state-level representativity after 2019 recruitment, it was determined that further statewide representativity was needed in four states: WI, MO, WA, and CO. As such, statewide samples using the USPS DSF file were generated for supplemental recruitment.

In 2021, NORC also recruited into an AmeriSpeak probability sample of persons aged 50 and older using a random national consumer address file (estimated 96% sample coverage of all households in the U.S.). AmeriSpeak re-empaneled approximately 6,000 study participants in this initiative.

It was clear at the start of 2021 that NORC would not immediately be able to conduct in-person interviewing given the ongoing COVID-19 pandemic. However, NORC sought to test new sampling strategies (noted below) early in 2021 in the hopes of documenting their efficacy and continuing and improving on them for the rest of 2021. Additionally, it was hoped that NORC would be able to conduct in-person interviewing in the second half of 2021. As such, the 2021 recruiting sample was split into five replicates, the first of which utilized DSF sample, and was released early in the calendar year, while future replicates were sampled using the NORC National Frame and were held until mid-year for recruiting.

⁴ A small fraction of initially nonresponding housing units is not eligible for NRFU, including “hard refusals” and those with an appointment for a call back from NORC.

⁵ A young adult flagged household refers to a household where MSG or TargetSmart indicated there was an 18-24-year-old adult in the household. In 2016 and 2017, a slightly different definition was used, and a young adult flagged household was defined as having an 18–34-year-old adult in the household by MSG or 18–30-year-old adult by TargetSmart.

⁶ A Hispanic flagged household refers to a household where MSG or TargetSmart indicated the presence of a Hispanic adult in the household.

⁷ A non-Hispanic Black-flagged household refers to a household where MSG or TargetSmart indicated the presence of a non-Hispanic Black adult in the household.

At the end of 2020, a major assessment of panel representativeness was conducted to inform the 2021 sampling strategy. This analysis again explored representativity by state, but as well explored a full range of demographic variables. This analysis was conducted both with the full panelist dataset as well as by assessing “effective panelists,” a measure of the likely demographic distributions that would occur among complete cases in any typical AmeriSpeak survey. This analysis found that AmeriSpeak could benefit from additional panelists in seven groups: households earning over \$200,000, household with children, Hispanics, Hispanics that specifically speak Spanish, African Americans, persons ages 18 to 24, and persons with less than a High School education. As such, the sample was stratified using NORC Big Data Classifiers (Dutwin et al, 2022), a technique utilizing available consumer and other public Big Data to make predictions on a range of household attributes during survey sampling. Households predicted to be one of these seven attributes were oversampled, while households predicted only hold persons aged 50 and older, or otherwise not predicted hold someone with one of the seven attributes, were under sampled. This sampling method was tested in the first sampling replicate, and given very positive results, was continued in all other 2021 replicates.

NORC’s strategy of “waiting it out” was effective, as the sample replicates released mid-year allowed NORC, to wait for an effective “COVID window” to conduct in-person interviewing. In short, in-person interviewing commenced after the peak of the Delta variant in 2021 and concluded with the peak of the Omicron variants. NORC was able to conduct a full NRFU in-person effort during this time.

Panel Recruitment Procedures

AmeriSpeak Panel recruitment is a two-stage process: (i) initial recruitment using USPS mailings, telephone contact, and modest incentives, and (ii) a more elaborate NRFU recruitment using FedEx mailings, enhanced incentives, and in-person visits by NORC field interviewers.

For the initial recruitment, sample households are invited to join AmeriSpeak online by visiting the panel website AmeriSpeak.org or by calling a toll-free telephone line (inbound/outbound supported). Both English and Spanish languages are supported for online and telephone recruitment. The initial recruitment data collection protocol features the following: an over-sized pre-notification postcard, a USPS recruitment package in a 9”x12” envelope (containing a cover letter, a summary of the privacy policy, FAQs, and a study brochure), two follow-up postcards, and contact by NORC’s telephone research center for sample units with a matched telephone number.

For the second stage NRFU recruitment, a stratified random sample is selected from the nonrespondents of the initial recruitment. Units sampled for NRFU are sent a new recruitment package by Federal Express with an enhanced incentive offer. Shortly thereafter, NORC field interviewers make personal, face-to-face visits to the pending cases to encourage participation. Once the households are located, the field interviewers administer the recruitment survey in-person using CAPI or else encourage the respondents to register online or by telephone.

Panel Recruitment Response Rate and Other Panel Statistics

A sample household is considered recruited or responded if at least one adult in the household joins the panel. The weighted household response rate (AAPOR RR3) is about 6% for initial recruitment and 28% for NRFU recruitment. We report two recruitment response rates: one for all the panel recruitment years (2014-2021) and one for the recruitment years with NRFU (2014-2018 and 2021). For all recruitment years, the cumulative weighted household response rate is 21.9%; for recruitment years with NRFU, and the cumulative weighted household response rate is 34.0%.⁸ For client studies requiring a panel recruitment response rate exceeding 30%, the sampling frame may be restricted to the panelists recruited in the NRFU years. The panel

⁸ As the 2021 NRFU is continuing to wind down, response rates noted here are estimated for 2021 sample cases.

recruitment response rate calculation methodology is consistent with AAPOR guidelines and fully documented.⁹ The annual panel retention rate is about 85%.

For individual client surveys based on the AmeriSpeak Panel, the AAPOR RR3 response rate is between 10% to 20% depending on specific study parameters such as target population, survey length, time in the field, salience of subject, and the like. This response rate takes into account panel recruitment rate, panel retention rate, and survey participation rate.¹⁰

Other important panel statistics with respect to the 2014-2019 and 2021 recruited households are as follows: 68% are recruited in the initial stage and 32% are recruited via NRFU; 92% of the active panelists prefer to do web or online surveys, while 8% prefer to participate in telephone surveys; 16% of the recruited households are non-Internet¹¹; 82% are cell phone only or cell phone mostly; 17% are African-American and 18% Hispanic; and 29% have household income below \$30,000 (compared to CPS benchmark of 26%).¹²

Impact of Non-Response Follow-Up

NRFU is instrumental in producing the industry-leading response rate for AmeriSpeak Panel recruitment. Moreover, due to the more intensive effort, NRFU recruitments better represent hard-to-reach groups and are therefore more representative of the target population. For example, initial recruitment tends to under-represent young adults 18-34 years of age. NRFU recruitment corrects for this bias by bringing the age distribution of the panel closer to ACS benchmarks.

Overall, NRFU recruitment significantly improves the representation of the panel with respect to demographic segments that are under-represented among the respondents to the initial recruitment, including young adults (persons 18 to 34 years of age), African Americans, Hispanics, lower-income households, renters, cellphone-only households, and persons with lower educational attainment (e.g., no college degree). To the extent that these demographic characteristics are correlated with substantive survey variables, NRFU helps to reduce potential non-response bias in the sample estimates. NORC's research indicates that NRFU respondents are indeed somewhat different from initial respondents for many common survey variables. For example, compared to the panelists recruited during the initial stage, NRFU panelists tend to be more conservative politically, more likely to attend church, less interested in current events or topics in the news report, less knowledgeable about science, less likely to be in favor of gun control policies, less likely to read a print newspaper (more likely to read the news online and use social media), more likely to eat at fast-food restaurants, and so on¹³. These observations illustrate that NRFU recruitment is critical for achieving a more

⁹ See http://amerispeak.norc.org/research/Pages/WhitePaper_ResponseRateCalculation_AmeriSpeak_2016.pdf

¹⁰ A properly calculated cumulative AAPOR response rate for panel-based research takes into account all sources of non-response at each stage of the panel recruitment, management, and survey administration process (see https://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf, page 48-9). A common misapplication of the term “response rate” in online panel surveys is to represent the survey-specific cooperation rate as the “cumulative survey response rate.” See “Response Rate Calculation Methodology for Recruitment of a Two-Phase Probability-Based Panel: The Case of AmeriSpeak” authored by Robert Montgomery, J. Michael Dennis, N. Ganesh. The paper is available at <https://amerispeak.norc.org/research/>.

¹¹ The non-internet households (HHs) are those that do not select “High-speed, broadband internet at home (such as cable or DSL)” or “Dial-up internet at home” response options when they are asked “What kind of internet access do you have? Please select all that apply” item in the recruitment survey. The non-internet HHs include those that only use internet on a cell connection or mobile phone.

¹² For transparency purposes, unweighted percentages are presented in this section. Hence, these results do not take into account selection probabilities. The base weighted distributions that take into account selection probabilities can be provided upon request.

¹³ See “The Undercounted: Measuring the Impact of ‘Nonresponse Follow-up’ on Research Data and Outcome Measures” authored by Ipek Bilgen, J. Michael Dennis, N. Ganesh. The paper will be soon available at <https://amerispeak.norc.org/research/>.

balanced panel and for making the substantive estimates in AmeriSpeak studies more accurate. Even though NRFU panelists are more reluctant to complete surveys, the addition of NRFU panelists reduced total absolute bias on average 5 to 21 percentage points when compared to the initial stage recruits (among examined surveys).¹⁴

Mixed-Mode Data Collection

The AmeriSpeak Panel supports mixed-mode data collection to improve response rate and the representativeness of the complete surveys. During the recruitment survey, AmeriSpeak panelists are offered an opportunity to choose their preferred mode—web or phone—for future participation in AmeriSpeak surveys. A recruited household can consist of both web- and phone-mode panelists. Panelists predominantly prefer web over phone mode. As of February 2020, 92% of the active panelists prefer to do web or online surveys, while 8% prefer to participate in telephone surveys. The telephone mode encompasses panelists without internet access, panelists whose only internet access is via a smartphone, and panelists with internet access but are unwilling to share an email address.

To the extent that non-internet households or “net averse” persons are different from the rest of the population, mixed-mode surveys have better population coverage and produce more accurate population estimates. NORC’s telephone interviewers administer the telephone surveys using a data collection system supporting both the phone and web modes, providing an integrated sample management and data collection platform. For panelists using smartphones for web-mode surveys, the NORC survey system renders an optimized presentation of the survey questions for these mobile users.

Panel Management and Maintenance

Panel management and maintenance are crucial for panel health and efficiency. NORC maintains strict panel management rules to limit respondent burden, reduce panel attrition, and minimize the risk of panel fatigue. On average, AmeriSpeak panelists are invited to participate in client studies two to three times a month. AmeriSpeak works with NORC clients to create surveys that provide an appropriate user experience for AmeriSpeak panelists. AmeriSpeak will not field surveys that in our professional judgment will result in a poor user experience for our panelists. AmeriSpeak also has a designated website and a telephone number for panelist communications.

Panel maintenance is a dynamic process because the AmeriSpeak Panel is supplemented and refreshed regularly over time to grow the panel, compensate for panel attrition, and improve panel representation for specific subpopulations. For example, the Latino Panel and Teen Panel are created to support studies of Hispanics and teenagers, respectively; the 2019 recruitment is primarily designed to improve sample representation at the state level. As panelists are added or/and removed from the panel, the panel refreshment process takes place to ensure that the refreshed panel fully represents the corresponding target population.

ABOUT NORC AT THE UNIVERSITY OF CHICAGO

As one of the world’s foremost independent research institutions, NORC at the University of Chicago delivers objective data and meaningful analysis to help decision-makers and leading organizations make informed choices and identify new opportunities. Since 1941, NORC has applied sophisticated methods and tools, innovative and cost-effective solutions, and the highest standards of scientific integrity and quality to conduct and advance research on critical issues. Today, NORC expands on this tradition by partnering with

¹⁴ See “Nonresponse Follow-up Impact on AmeriSpeak Panel Sample Composition and Representativeness” authored by Ipek Bilgen, J. Michael Dennis, N. Ganesh. The paper is available at <https://amerispeak.norc.org/research/>.

government, business, and nonprofit clients to create deep insight across a broad range of topics and to disseminate useful knowledge throughout society.

Headquartered in downtown Chicago, NORC works in over 40 countries around the world, with additional offices on the University of Chicago campus, the DC metro area, Atlanta, Boston, and San Francisco.

ADDITIONAL RESOURCES

To learn more about AmeriSpeak or to share an RFP, please contact AmeriSpeak at AmeriSpeak-BD@norc.org. Information about AmeriSpeak capabilities and research papers are available online at AmeriSpeak.NORC.org.