

TECHNICAL OVERVIEW OF THE AMPLIFY AAPI PANEL NORC'S PROBABILITY-BASED HOUSEHOLD PANEL OF AANHPI

Updated September 11, 2023

This technical overview provides the basic information about Amplify AAPI, a large probability-based panel operated by NORC at the University of Chicago. Amplify AAPI is designed to be representative of the U.S. Asian American, Native Hawaiian and Pacific Islander (AANHPI) household population, including all 50 states and the District of Columbia. U.S. households are randomly selected with a known probability from a national frame of addresses and then recruited by mail and telephone, and in limited cases, face-to-face interviews. Amplify AAPI 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 Amplify AAPI started in the spring of 2022 with a pilot study to recruit an initial 150 households from a sample of 30,000 and to test feasibility of a sample design and assess participation rates. In 2023, a number of recruiting efforts were executed to accommodate separate funding commitments. Specifically, in the spring, NORC conducted a recruit of 18,000 sampled addresses with a goal to recruit approximately 400 panelists in California. In the summer, a second funding source allowed for a national effort that included both fresh recruiting from a national sample as well as the empanelment of households that recently participated in a large scale probability-based Asian American Survey and agreed to be recontacted by NORC for future surveys. As well, a third funding source was used to generate another recruit of approximately 750 Californians, with a minimum of 75 recruits from six groups: Chinese, Filipino, Asian Indian, Vietnamese, Korean, and NHPI. Future additional recruits are anticipated in 2024 and 2025. As well, there are approximately 850 active AANHPI panelists in NORC's flagship probability panel, AmeriSpeak, which serve as dual members with the Amplify AAPI panel as well.

All told, AAPI Amplify is able to support surveys of 1,500 presently.

Panel Sample Frame

Given multiple sources and funders, each with different geographic and AANHPI subgroup quota expectations, Amplify AAPI is a blend of multiple sample designs. Each is designed to provide maximum feasible coverage of the AANHPI population, and all are combined through weighting adjustments to represent the overall cross-section of AANHPI in the U.S.

For the active AANHPI panelists in AmeriSpeak, the primary sampling frame is the 2010 NORC National Frame, a multistage probability sample that fully represents the U.S. household population, with supplements from the USPS computerized delivery sequency file (CDS). AmeriSpeak uses a very high-quality recruiting protocol that includes three recruitment mailings and a nonresponse follow up protocol of Federal Express mailings and door-to-door recruiting. As a nationally representative general population panel, AmeriSpeak recruits both adults and teens, age 13 and older, and presently has over 70,000 panelists. A detailed description of AmeriSpeak sampling and other technical information is provided at <https://amerispeak.norc.org/content/dam/amerispeak/research/pdf/AmeriSpeak%20Technical%20Overview%202019%2002%2018.pdf>. As a general random sample, AANHPI panelists in AmeriSpeak are similarly

attained through random selection with over 96% coverage of all households in the U.S. AmeriSpeak however does not cover Asian-language “linguistically isolated” households, that is, households in which no adult can speak English or Spanish at least “well.” As such recruiting linguistically isolated AANHPI households is a particular goal of most other recruits, past and future.

The initial Pilot study was designed to test a principal sampling approach to be used as a general framework for all national samples. This comprised of using big data-modelled predictive variables that indicate the potential presence of an AANHPI household, utilizing available public and consumer data (for more on big data modelling see Dutwin at al., 2023 in the Journal of Survey Statistics and Methodology). The address frame is based on a vendor file that matches the number of households in the U.S. and is considered a frame with near-complete coverage of U.S. households. Coupled with this frame is the use of the CDS to sample addresses not predicted to be AANHPI through modelling. Analysis using AmeriSpeak and other sources (see Table 1 below) finds that the predictive modeled data do quite well at covering the AANHPI population and that selective sampling of non-predicted households is effective and attaining large coverage of all AANHPI in the U.S.

The CDS file is organized into three strata based on census block group and specifically the percent of AANHPI in each block group. The “high” incidence strata include all block groups for which at least 30% of households are AANHPI, with the “medium” inclusive of households 10-30%. AANHPI households that are not identified with the modelling and are not in a block group of at least 10% incidence are not covered due to the exorbitant cost of doing so.

Table 1: AAPI Amplify Pilot Sample Design

Sample Strata	% of AANHPI Population	Prevalence of AANHPI Households
Asian Language Model	19.6%	58.9%
Asian English Model	43.9%	46.3%
High Census Blocks AAPI	6.2%	11.8%
Mid Census Block AAPI	11.7%	3.8%
NHPI Model	0.6%	37.5%
Other (not sampled)	18.0%	1.3%

Overall, as shown in Table 1, the Pilot design included five sampling strata: three strata using big data predictions (Asian language, Asian English, and NHPI) as well as two strata to cover residual census block group households in block groups with AANHPI prevalence above 10 percent. The design covers about 82% of the U.S. AAPI population. Sample was allocated proportionally with the exception of the Mid Census Block Group AAPI strata receiving a probability of selection half its population prevalence, and the High CBG AAPI strata holding a probability of selection twice its population prevalence, again due to the very high cost of screening.

The California recruit focused specifically on predicted big data sample only, and divided sample between predicted Chinese households and other households. This division is guided by findings from our Pilot and the recent large scale survey of Asian Americans that Chinese households exhibit much higher survey participation rates compared to the other groups. As such, the sample fraction of Chinese predicted households was half of all other Asian Americans. As well, the funding for this recruit was specifically for English-only households and the sample was limited as such.

The second California sample required a more balanced recruit by subgroup and thus utilized big data predictions across a range of subgroups. As well, another outcome of the Pilot was a strong tendency of AANHPI with higher educational attainment to participate versus those of lower educational attainment. As such, again modelled predictions were utilized on education and those predicted to have high education were undersampled by a factor of two. As well, the design created in-language strata to ensure a high representation of non-English recruits.

Table 2: AAPI Amplify CA Sample Design

Strata	% of Sample Universe	% of Sample	Sample Fraction	Expected % of Recruits	Subgroup
1 Chinese Low Educ	7.5%	3.4%	0.45	5.1%	18.2%
2 Chinese High Educ	15.3%	3.5%	0.23	5.3%	
3 Chinese Language	11.5%	5.2%	0.45	7.9%	
4 Asian Indian Low Educ	12.0%	14.5%	1.21	10.5%	14.4%
5 Asian Indian High Educ	9.0%	5.5%	0.61	3.9%	
6 Filipino Low Educ	7.9%	17.5%	2.22	11.8%	13.4%
7 Filipino High Educ	2.2%	2.5%	1.11	1.7%	
7 Vietnamese Low Educ	2.2%	3.4%	1.55	2.7%	13.7%
8 Vietnamese High Educ	4.7%	3.7%	0.77	3.0%	
9 Vietnamese Language	6.5%	10.0%	1.55	8.1%	
10 Korean Low Educ	1.5%	3.4%	2.30	3.5%	14.3%
11 Korean High Educ	2.5%	2.9%	1.15	2.9%	
12 Korean Language	3.3%	7.7%	2.30	7.9%	
13 NHPI Low Educ	1.0%	8.8%	8.47	13.4%	15.2%
15 NHPI High Educ	0.3%	1.2%	4.24	1.8%	
16 Residual	12.6%	7.0%	0.56	10.6%	10.6%

The first national recruit also leveraged insights from the Pilot to create a more representative sample. This includes (a) undersampling high education households by half; (b) undersampling Chinese-predicted households ; and (c) modest undersampling of Japanese households and oversampling of Filipino households. Again, these sample fractions are principally designed to account for expected differential nonresponse so as to attain a cross-section of Asian sub-groups that is reflective of true population distributions.

Table 3: AAPI Amplify National Sample Design

Strata Name	% of Sample
Asian Language	8.8%
Asian Language High Education	2.6%
Asian Language Chinese	8.9%
Asian Language High Educ Chinese	3.8%
Asian (English)	22.6%
Asian High Educ Chinese	8.5%
Asian Chinese	7.9%

Asian High Educ Chinese	2.6%
Asian Filipino	4.9%
Asian High Educ Filipino	0.7%
Asian Japanese	3.4%
High CBG	14.6%
Medium CBG	7.0%
NHPI	3.7%

Future recruitment sample designs will likely combine the best features of all past designs so that they will ensure over-representation of smaller AANHPI groups. This will allow the panel to be able to both represent the overall AANHPI population as well as to “drill down” to the largest 5 AAPI groups as well as NHPI specifically.

Panel Recruitment Procedures

Amplify AAPI recruitment is a two-stage process: (i) initial screening using USPS mailings, telephone contact, and modest incentives, and (ii) recruitment.

For the initial screening, sample households are invited to an online or phone (respondent’s choice) survey by visiting a panel website or by calling a toll-free telephone line (inbound/outbound supported). English, Chinese dialects of Mandarin and Cantonese, Vietnamese, and Korean were offered in both printed materials and by telephone recruiters in samples that offered non-English languages. The initial recruitment data collection protocol features an over-sized pre-notification mailing card (9 x 12). This card urged respondents to go online or call in and provided instructions in English and the four languages above.

The recruitment survey is specifically designed to identify whether a household is AANHPI, commencing with a number of “warm-up” questions and then asking about race/ethnicity. The survey identifies whether the respondent is AANHPI as well as whether there are other members of the household who may be AANHPI. If such a member is present, the survey then moves to recruitment, explaining the importance of the panel to represent the AANHPI community, and asking the respondent to join the panel. By joining the panel, the respondent is informed they will receive a \$25 incentive and that each survey they take will further provide them \$3-5 (depending on survey length) for their participation.

Panel Recruitment Response Rate and Other Panel Statistics

NORC is currently developing the infrastructure to report response metrics and such data will be widely published when available.

Multi-Modality

Amplify AAPI supports mixed-mode data collection to improve response rate and the representativeness of the complete surveys. During the recruitment survey, Amplify AAPI panelists are offered an opportunity to choose their preferred mode—web or phone—for future participation in Amplify AAPI surveys. A recruited household can consist of both web- and phone-mode panelists. Panelists predominantly prefer web over phone mode. 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 are expected to 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, Amplify AAPI panelists are invited to participate in client studies once a month. NORC researchers work with NORC clients to create surveys that provide an appropriate user experience for AAPI panelists. NORC will not field surveys that in our professional judgment will result in a poor user experience for our panelists.

Weighting

NORC is a leader in advanced weighting procedures to minimize survey bias, particularly in probability-based panels. Below denotes a general approach to weighting surveys from Amplify AAPI.

Base weights

For each of the Amplify AAPI samples other than the AmeriSpeak Panel, the base weights are computed as the inverse of the sample selection probabilities under the respective sample design per stratum. For a sample household i in stratum h , the base weight is computed as $1/p_{hi}$, where p_{hi} is the ratio of the sample count to the frame count per stratum, n_h/N , in stratum h .

We denote the base weights as W_{1i} .

Normalized W_{1i} sum to the recruitment sample size per sample source, we denote these normalized base weights as W_{2i} .

Person-level weights

Starting from basewt W_{2i} , normalize basewt W_{2i} to the respondent sample size per sample source, we denote the base weights as W_{3i} . We denote the person-level weights as W_{4i} .

Raking adjustments

The total respondents will be raked to the AAPI population benchmarks from the American Community Survey, 5 years datafile.

The following variables are used in raking, where the total count along each dimension is equal to the total AAPI population.

- a. Age x Gender : age (18-34, 35-49, 50-64,65+) by gender
- b. Education: High School diploma or less, some college, 4-year degree, graduate degree
- c. U.S. born (yes/no)
- d. Region: Northeast, Midwest, South, West
- e. Asian Origin: Chinese, Asian Indian, Filipino, Vietnamese, Korean, Japanese, NHPI, Other

We denote the raked and normalized weights as W_{5i} .

ABOUT NORC AT THE UNIVERSITY OF CHICAGO

NORC at the University of Chicago conducts research and analysis that decision-makers trust. As a nonpartisan research organization and a pioneer in measuring and understanding the world, we have studied almost every aspect of the human experience and every major news event for more than eight decades. Today, we partner with government, corporate, and nonprofit clients around the world to provide the objectivity and expertise necessary to inform the critical decisions facing society. www.norc.org