Using Probability Panels to Enhance AANHPI Representation in Cross-Sectional Surveys

Prepared by Brian M. Wells, Lisa Lee, Ipek Bilgen

Introduction

Survey researchers face persistent challenges in collecting and reporting detailed data on Asian American, Native Hawaiian, and Pacific Islander (AANHPI) communities (Kauh, 2023; Shimkhada, Scheitler, & Ponce, 2021). AANHPI comprise about seven percent of the population in the United States and are a rapidly growing, diverse group, representing multiple countries of origin (Wu & Moy, 2025; Shimkhada et al., 2021).

However, given the small size of the AANHPI population in the United States, it is difficult to obtain large enough samples in general population surveys to analyze and report on AANHPI broadly or as separate subgroups. Approaches to better survey AANHPI in general population or targeted studies present challenges in terms of cost and recruitment efforts. As a result, AANHPI attitudes, behaviors, and experiences are often missing from survey results.

One approach to increasing their representation in surveys is to oversample these groups (Shimkhada et al., 2021). Oversampling involves selecting additional AANHPI respondents so they comprise a larger proportion of the survey sample than they do in the population, while applying appropriate weighting adjustments to the full sample to produce results that remain generalizable to the general population.

This approach is often used to study harder-to-reach and harder-to-survey communities. Oversampling can be accomplished, for example, by increasing the probability of selection for respondents in certain geographies known to have a higher incidence of the population of interest. A drawback of this sample design is increased design effects and reduced effective sample size. And even with the use of geographic clustering or other methods of targeting certain populations, study costs often increase substantially.

In this brief, we present a case study on how the number of AANHPI respondents in a nationally representative, cross-sectional study can be expanded by leveraging the strengths of a multi-purpose probability-based survey panel. This type of panel uses a randomly selected and representative sample of a population to conduct a series of surveys covering a variety of research topics.

By using a multi-purpose probability-based survey panel, researchers can preselect AANHPI respondents from the panel and include them in the study sample. This targeted approach can yield efficiencies in identifying and collecting data from AANHPI persons. This case study from the General Social Survey (GSS) team at NORC at the University of Chicago, summarizes the 2022 GSS, which used NORC's AmeriSpeak® Panel to oversample AANHPI respondents.

GSS-AmeriSpeak Case Study

The GSS is a nationally representative, cross-sectional survey conducted since 1972 that

monitors public opinion trends on a repeated core of behavioral, attitudinal, and topical questions. For most of its history, GSS has relied on in-person data collection, occasionally supplementing with phone interviews.

Following a temporary shift to web data collection during the COVID-19 pandemic, the 2022 GSS was fielded as a multimode survey. In-person interviews and web questionnaires were the primary modes of data collection, supplemented by a small number of phone interviews (Davern et al., 2024). The expansion to self-administration opened the GSS to alternative frames and data collection approaches to achieve its purposes of representation.

As part of the 2022 design, the GSS included an expanded sample of select subgroups that reflect the national population obtained from NORC's AmeriSpeak® Panel (Davern et al., 2024; Wells & Sparkman, 2024). AmeriSpeak is a probability-based multi-purpose survey panel recruited from a national address-based frame via multiple modes of sample contact, including mail, text, telephone, and in-person (NORC, 2024). One goal of this expanded sampling effort was to obtain an additional 200 completed interviews of the AANHPI population.

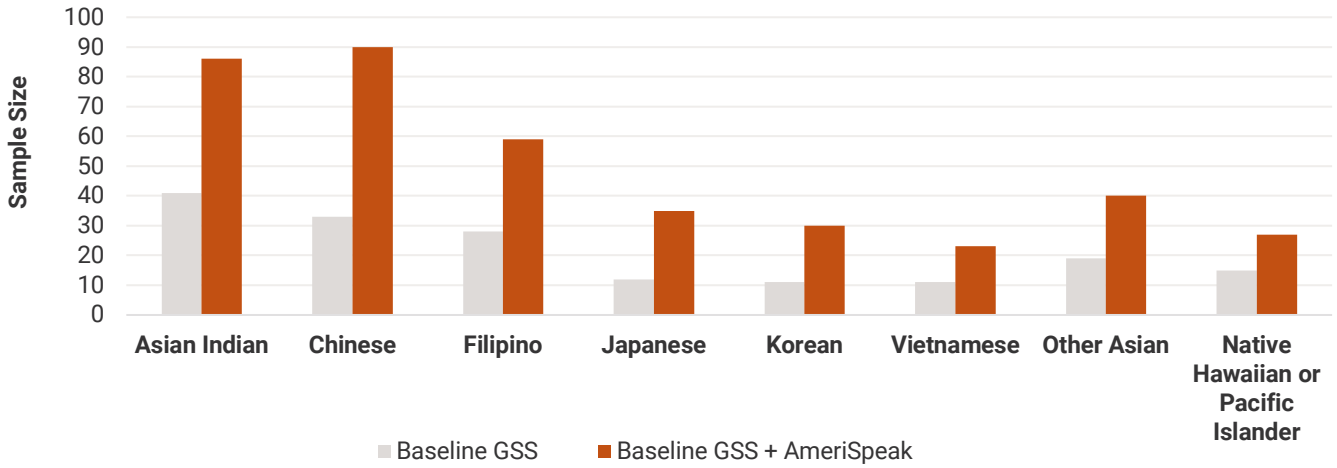
While the AmeriSpeak oversample aimed to increase representation of the AANHPI population overall, the GSS also examined how the oversample increased representation across the six major Asian origin groups traditionally identified by the U.S. Census Bureau—Asian Indian, Chinese, Filipino, Japanese, Korean, and Vietnamese—plus the combination of all Native Hawaiian and Pacific Islander (NHPI) groups (e.g., Native Hawaiian, Samoan, and Chamorro). An "Other Asian" category captured all remaining Asian origin groups not listed above. The GSS race question follows the same structure as the Census race question, allowing for multiple race selections.

Figure 1 shows how the addition of the AmeriSpeak Panel sample to the baseline GSS cross-sectional sample consistently increased the proportion for every AANHPI group. Comparing the baseline sample to the combined sample, the AmeriSpeak addition more than doubled the unweighted sample sizes for all Asian subgroups (see Table 1), ranging from a 192 percent increase for Japanese persons (n = 12 to n = 35) to a 109 percent increase for Vietnamese persons (n = 11 to n = 23). The AmeriSpeak sample also increased the sample size for NHPI persons by 80 percent (n = 15 to n = 27).

Given concerns over larger design effects when expanding subgroup samples, we also estimated effective sample sizes for each of the AANHPI groups. Larger AANHPI groups, such as Asian Indian, Chinese, and Filipino, saw larger design effects when considering the baseline sample. This resulted in a greater loss of information, as indicated by the smaller ratios of effective and unweighted sample sizes (see Table 1).

While there are some fluctuations in these ratios when including the AmeriSpeak sample, we did not observe a systematic change across AANHPI groups. This suggests the addition of the AmeriSpeak design does not negatively impact the design effect for AANHPI persons relative to the standard GSS design.

Figure 1: Unweighted Sample Sizes by AANHPI Subgroup in the 2022 GSS



Source: General Social Survey, 2022

Table 1: Unweighted and Effective Sample Sizes by AANHPI Subgroup in the 2022 GSS

Subgroup	Baseline sample			Baseline + AmeriSpeak sample			Percent increase over baseline sample alone (unweighted)	Percent increase over baseline sample alone (effective)
	Unweighted sample size	Median effective sample size*	Ratio of effective sample size and unweighted sample size	Unweighted sample size (increase from baseline alone)	Median effective sample size* (increase from baseline alone)	Ratio of effective sample size and unweighted sample size		
Asian	148	78	0.53	344 (+196)	198 (+120)	0.58	132%	154%
Asian Indian	41	27	0.66	86 (+45)	50 (+23)	0.58	110%	85%
Chinese	33	24	0.73	90 (+57)	78 (+54)	0.87	173%	225%
Filipino	28	22	0.79	59 (+31)	46 (+24)	0.78	111%	109%
Japanese	12	12	1.00	35 (+23)	31 (+19)	0.89	192%	158%
Korean	11	11	1.00	30 (+19)	30 (+19)	1.00	173%	173%
Vietnamese	11	11	1.00	23 (+12)	23 (+12)	1.00	109%	109%
Other Asian	19	18	0.95	40 (+21)	40 (+22)	1.00	111%	122%
Native Hawaiian or Pacific Islander	15	15	1.00	27 (+12)	23 (+8)	0.85	80%	53%

Source: General Social Survey, 2022
* Median effective sample sizes are rounded to the nearest whole number and are not allowed to be larger than the unweighted sample size. To calculate a median effective sample size, we followed the procedures described by Wells and Sparkman (2024) in evaluating the 2022 AmeriSpeak oversample.

We similarly observed a consistent pattern of sample size increases ranging from 53 percent for NHPI (down from 80 percent unweighted) to 225 percent for Chinese (up from 173 percent unweighted).

In summary, the 2022 design demonstrated the value of using the AmeriSpeak Panel to increase the number of AANHPI respondents in the GSS. The inclusion of additional AANHPI AmeriSpeak respondents improved representation for this group, increased effective sample sizes, and better enabled the analysis and reporting of findings for these subgroups. The success of this approach encouraged its continued use in the 2024 GSS.[1]

The Center's Perspective

The need to represent growing and diverse population subgroups, such as AANHPI, is an important consideration for survey research. This case study substantiates the value of using probability-based multi-purpose panels to enhance the representation of smaller subgroups in cross-sectional national studies.

To support meaningful subgroup analysis, researchers should determine early in the study design process whether their sampling plans will yield sufficient sample sizes for analysis of smaller population segments. When needed, supplementing with additional cases from multi-purpose probability-based survey panels can be a timely and cost-effective strategy to strengthen representation and produce reliable estimates.

In addition to national probability-based survey panels, we also view probability-based survey panels created for and composed of communities that are often underrepresented in survey research as a promising resource for increasing the representation of these groups and expanding the reach of survey studies. Accordingly, we see a continued need for the development and growth of representative subgroup panels to meet this increasing demand.

[1] Following the fielding of the 2022 GSS case study, AmeriSpeak launched Amplify AAPI, a dedicated probability-based panel of AANHPI households in the United States. (NORC, 2023). Amplify AAPI enables the collection of nationally representative data for AANHPI. Respondents can be interviewed in English, Chinese, Vietnamese, and Korean languages.

References

Davern, M., Bautista, R., Freese, J., Herd, P., & Morgan, S. L. (2024). *General Social Survey, 1972-2022* (Release 4) [Data set and codebook]. NORC. https://gss.norc.org/Documents/codebook/GSS%202022%20Codebook.pdf

Kauh, T. (2023, December 7). *Every community is worth collecting data on.* Robert Wood Johnson Foundation. https://www.rwjf.org/en/insights/blog/2023/12/every-community-is-worth-collecting-data-on.html

NORC. (2023, September 11). *Technical Overview of the Amplify AAPI Panel: NORC's Probability-Based Household Panel of AANHPI.* https://amplifyaapi.norc.org/content/dam/amplifyaapi/pdfs/Amplify-AAPI-Technical-Overview.pdf

NORC. (2024, August 13). *Technical Overview of the AmeriSpeak® Panel: NORC's Probability-Based Household Panel*. https://amerispeak.norc.org/content/dam/amerispeak/about-amerispeak/pdf/amerispeak-technical-overview.pdf

Shimkhada, R., Scheitler, A. J., & Ponce, N. A. (2021). Capturing racial/ethnic diversity in population-based surveys: Data disaggregation of health data for Asian American, Native Hawaiian, and Pacific Islanders (AANHPIs). *Population Research and Policy Review, 40*, 81-102. https://doi.org/10.1007/s11113-020-09634-3

Wells, B. M. & Sparkman, R. (2024). GSS 2022 AmeriSpeak Oversample. (General Social Survey Methodological Report No. 138). NORC. https://gss.norc.org/content/dam/gss/get-documentation/pdf/reports/methodological-reports/GSS%20MR138%20AmeriSpeak%20Oversample.pdf.

Wu, L. W. & Moy, R. H. (2025). Strengthening Asian/Asian American, Native Hawaiian, and Pacific Islander leadership in cancer research. *Cancer Discovery, 15*(2), 267-270. https://doi.org/10.1158/2159-8290.CD-24-1618

ACKNOWLEDGEMENTS

Thanks to René Bautista, David Dutwin, Jeannie Novak, Mark Sheehy, Micah Sjoblom, and Ting Yan for their valuable feedback and insights on this brief.

ABOUT NORC

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.