The General Social Survey

GSS

Post-stratification Weights for GSS 1972-2022

GSS Methodological Report #137
April 2024

Authors:
Brian M. Wells
Zachary H. Seeskin
Amy Ihde

NORC at the University of Chicago

SUMMARY

This report describes a new set of post-stratification weights available for users of the 1972-2018 General Social Survey (GSS) cross-sectional surveys to help improve nonresponse bias adjustment. The weight derivation follows the approach applied to 2021 and 2022 GSS Cross-sections. Use of these weights results in weighted totals that, for each GSS cross-sectional sample, equal marginal control totals from the U.S. Census Bureau estimates for education, sex, marital status, age, region of the country, race, U.S. born status, and Hispanic origin when available. NORC recommends that GSS data users use this new weight for all analyses in the future.

These weights also:

(a) correct for the form assignment errors reported in GSS Methodological Report 36 for 1978, 1980, 1982, 1983, 1984, and 1985;

(b) correct for the ballot-and-form assignment errors reported in GSS Methodological Report 134 for 2002, 2010, 2012, 2016, and 2018; and

(c) support person-level analyses of the combined main and Black oversamples for 1982 and 1987.

Given the global trend of declining response rates over the past several years, the use of auxiliary data, such as U.S. Census totals for nonresponse adjustment, is important for improving representativeness of estimates with respect to key demographic characteristics. In addition, this report examines the impact of using the post-stratification weights across all GSS cross-sections. The majority of estimate differences observed include post-stratification variables and their close correlates.

INTRODUCTION

Post-stratification and related survey weighting procedures use high-quality information regarding the population and conduct ratio adjustments to survey weights to align survey estimates with the population information (Lohr, 2021). For studies of the U.S. adult population, like the General Social Survey (GSS), population estimates are available from the U.S. Census, the American Community Survey, and the Current Population Survey. Post-stratification is helpful for nonresponse adjustment and to improve the representativeness of estimates with respect to the selected demographic characteristics. The use of post-stratification is particularly important in recent years due to trends in declining response rates (National Research Council, 2013) and to address disruptions in survey data collections due to the COVID-19 pandemic (Rothbaum et al., 2021).

In this report, we use the term “post-stratification” to refer to aligning survey-weighted distributions with high quality information regarding a set of population characteristics, or population control totals. The GSS post-stratification procedures may also be referred to as “raking.” Raking refers to a procedure to iteratively adjust survey weights to population totals one variable at a time. After iterating over each variable multiple times, the raking algorithm aligns the survey-weighted distributions for all variables used in the procedure with population totals.

NEW POST-STRATIFICATION WEIGHT VARIABLES

The new GSS weights, WTSSPS and WTSSNRPS, are the new post-stratified weights. WTSSPS is available for years 1972-2022,[1] and WTSSNRPS is available for 2004-2022.[2] WTSSPS and WTSSNRPS also account for weighting adjustments to address form assignment errors (GSS Methodological Report 36), ballot-and-form assignment errors (GSS Methodological Report 134), and/or Black oversamples (1972-2018 GSS Codebook Appendix A) for specific years.

WTSSPS and WTSSNRPS were derived by raking input weights to control totals developed from the U.S. Census, American Community Survey, or Current Population Survey data. For the 1972-1998 GSS, the input weight was either WTSSALL or a version adjusted for form assignment errors and/or Black oversamples in specific years (for more details, see Appendix A). For the 2000-2018 GSS, the input weights were either WTSSALL/WTSSNR or versions adjusted form-and-ballot assignment errors in specific years (for more details, see Appendix B). The final WTSSPS and WTSSNRPS were derived by raking the input weights to population control totals from the U.S. Census, American Community Survey, or Current Population Survey for the following characteristics:

- Census division (9-level)
  - New England
  - Mid-Atlantic
  - East North Central
  - West North Central
  - South Atlantic
  - East South Central
  - West South Central
  - Mountain
  - Pacific
- Marital status (2-level)
  - Currently married
  - Not currently married
- Hispanic origin (2-level)[3]
  - Of Hispanic origin
  - Not of Hispanic origin
- Education status (3-level)
  - Less than high school education
  - High school diploma, but no four-year college degree
  - Bachelor's degree or more
- U.S. born status (2-level)[4]
  - Born in the United States
  - Not born in the United States
- Sex (2-level)
  - Male
  - Female

[1] GSS Methodological Report 135 originally introduced the post-stratification weights for 2000-2018.
[2] WTSSNR was introduced in 2004 and, along with WTSSNRPS, is not available for prior years.
[3] Hispanic origin was used in raking for 2000-2018, but not for 1972-1998 GSS as the variable was not yet available.
[4] U.S. born status was used in raking for 1977-2018, but not for 1972-1976 GSS as the variable was not yet available.

- Race (3-level)[5]
  - White
  - Black
  - Other
- Age group (5-level)
  - 18-29 years old
  - 30-39 years old
  - 40-49 years old
  - 50-64 years old
  - 65-89 years old[6]

For a detailed breakdown by year and source, please see Appendix C.

Cases with missing values for these variables were imputed with hot deck imputation prior to raking. For the years 1975, 1976, 1983, and 1993, the GSS sample was drawn from multiple sampling frames and post-stratification was conducted within each sampling frame for these years.

METHODS

In order to examine the impacts of these post-stratification weights for historical estimates from 1972 through 2018, we analyzed a subset of 179 historical variables across these years, including all post-stratification raking dimensions, primarily focused on categorical variables with ten or fewer response categories. Depending on the year, the number of variables ranges from 43 variables with 239 response categories (1972) to 178 variables with 962 response categories (2018), with categories including select reserve codes (i.e., "Don't know", "No answer", "Can't choose"). We performed this analysis using Release 2 of the 1972-2022 GSS cumulative file. All analyses adjusted for strata and clustering.[7]

We estimated using both the historical weight, WTSSALL, and the new post-stratification weight, WTSSPS. For simplicity, we used WTSSALL unadjusted. For the various ballot and form issues (comprising 1978, 1980, 1982, 1983, 1984, 1985, 2002, 2010, 2012, 2016, and 2018), we did not include the corresponding adjustments (e.g., BALLOTFORMWT). For the 1982 and 1987 Black oversamples, we followed the historical recommendation to exclude cases in the Black oversamples comprising codes 4, 5, and 7 of the variable SAMPLE (Davern et al., 2024). Naïve researchers may not exclude these cases and not adjust WTSSALL with the variable OVERSAMP which will result in big shifts in estimates for those years regarding race and related racial issues (e.g., national spending to improve the condition of Blacks, variable NATRACE). No weight adjustments or sample exclusions are needed when estimating with WTSSPS.

In order to compare the two weights, we examined the weighted estimates for all individual responses using WTSSPS relative to WTSSALL. We compared the ratio of absolute difference between the two estimates to the standard error of the estimate using WTSSALL to determine changes in estimates that were large relative to standard errors:

$$\frac{\left|\hat{\theta}_{WTSSPS} - \hat{\theta}_{WTSSALL}\right|}{SE\left(\hat{\theta}_{WTSSALL}\right)}$$

[5] Race included six categories for 2021 and 2022 including white, Black/African American, American Indian/Alaska Native, Asian, Native Hawaiian or other Pacific Islander, and other race.
[6] The GSS does not cap respondents at age 89, but the GSS does not release respondent ages greater than 89.
[7] Strata and clustering are not used for survey estimation in 1972, 1973, and 1974 given quota sampling was used.

The comparison is non-standard for statistical testing, as we compare two estimates based on the same response values but different survey weights. Nonetheless, the ratio is informative regarding estimates with large differences relative to their standard errors, and we note ratios of absolute differences to standard errors greater than 1.96, with the choice of 1.96 motivated by a significance level of 5% in a statistical testing framework. We refer to the above quantity as the "absolute difference-to-standard error ratio" or "ratio" in brief. We refer to differences with such ratios greater than 1.96 as "large."

As post-stratification can lead to increased estimate variance depending on the effectiveness of the adjustment (Dever & Valliant, 2010), we also examine the relative change in standard error of the new WTSSPS weight from the historical WTSSALL weight:

$$\frac{SE(\hat{\theta}_{WTSSPS}) - SE(\hat{\theta}_{WTSSALL})}{SE(\hat{\theta}_{WTSSALL})}$$

RESULTS

Of the 179 variables considered, only 57 variables, or 32 percent of the variables examined, had at least one year where there was an absolute difference-to-standard error ratio greater than 1.96. Of these 57 variables, less than half had more than one large difference (corresponding to ratio greater than 1.96) across the 32 cross-sections of the GSS.

Regarding analysis of response categories, on average 1.8 percent total per year saw absolute difference-to-standard error ratios greater than 1.96, ranging from 0.5 percent of response categories (2008) to eight percent (1972) (see Exhibit 1). The biggest shift was in 1972, which was the inaugural year of the GSS and a year that the GSS used a non-probability quota sample. Excluding 1972, the biggest difference for a given GSS year is less than four percent.

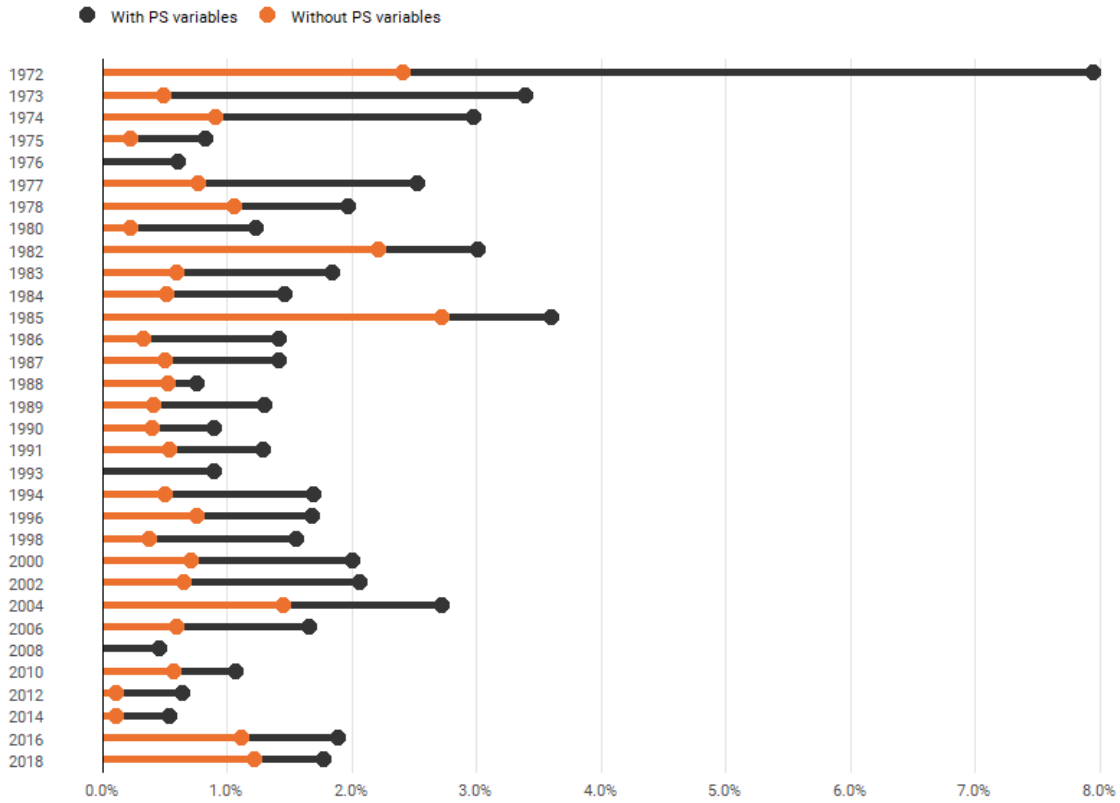
It was important to consider for this analysis whether variables were used as post-stratification dimensions or not. Averaging across the years, an average of 70 percent of the differences with large absolute difference-to-standard error ratios were post-stratification responses (denoted by the length of the black bar in Exhibit 1). Some years were not strongly dominated by post-stratification variables (e.g., 29 percent in 1985), but for three years the large differences (with ratios greater than 1.96) were exclusively among post-stratification variables (1976, 1993, 2008). After removing the post-stratification variables, the average percentage of large differences is less than one percent of response categories, ranging from zero (1976, 1993, 2008) to 2.7 percent (1985).

Post-stratification variables

Given the majority of the variables that saw changes in their estimates were post-stratification dimensions themselves, we examine these variables first. The most common estimate changes we observed were U.S. born status, age, marital status, sex, race, and highest degree obtained.

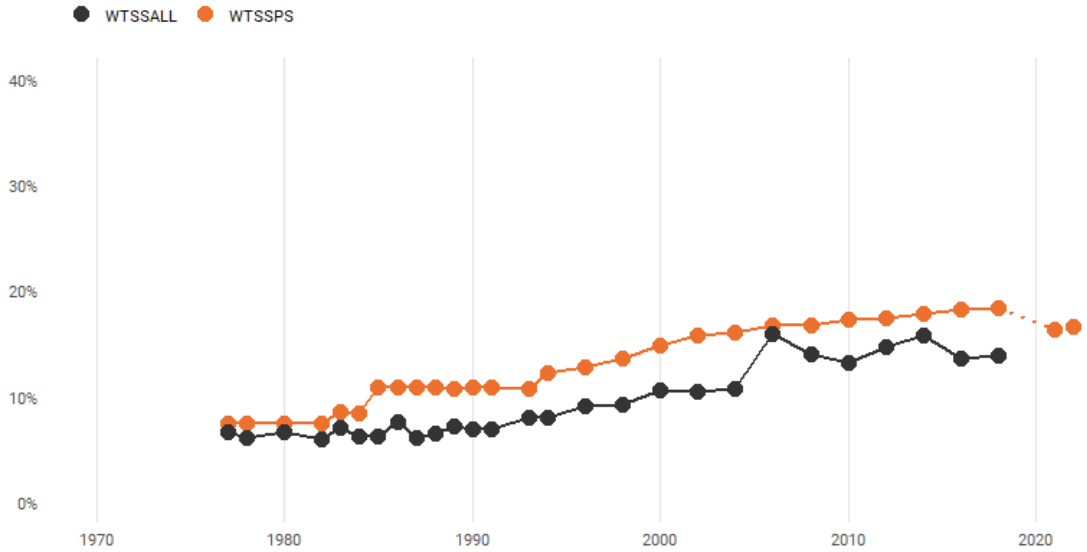
The historical weights resulted in consistently underestimating the proportion of foreign-born respondents with 23 of the 27 years relative to the post-stratified weights. The differences between the estimated proportion of persons born outside of the United States comparing the historical weights and the post-stratification weights range from a one and a half (1982) to over five percentage (2002 and 2004) point difference (see Exhibit 2).

Exhibit 1. Percent of large differences between WTSSALL and WTSSPS among response categories by year



Source: General Social Survey, 1972-2018
Note: 'Large difference' defined as absolute difference-to-standard error ratio greater than 1.96. The longer the black line the more of the large differences for response categories were post-stratification response categories.

Exhibit 2. Estimates for not U.S. born using WTSSALL and WTSSPS

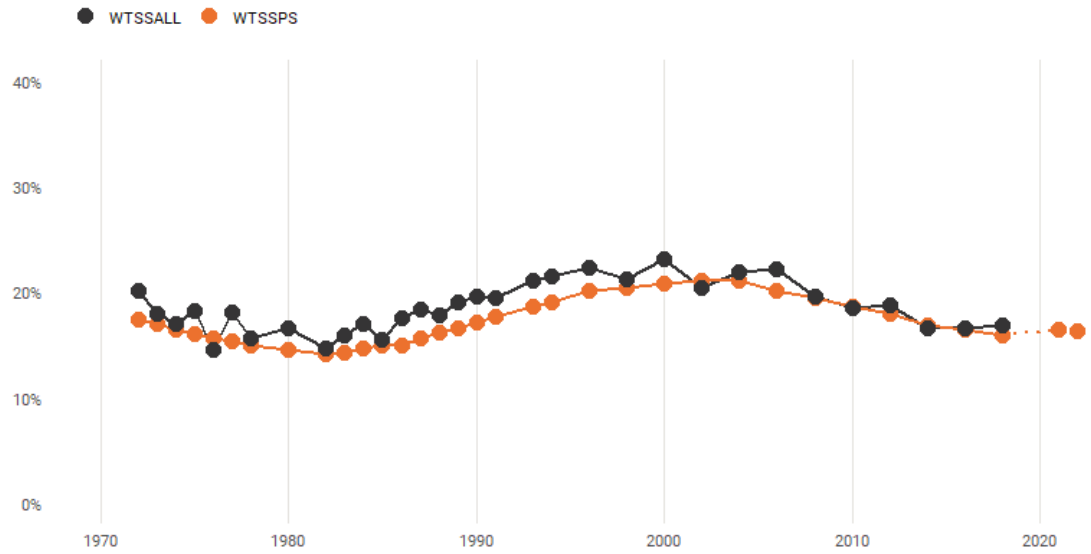


Source: General Social Survey, 1977-2022
Note: Variable BORN, response 2. Estimates exclude reserve codes.

We observed many large differences for respondent age (21 of 32 years examined), with differing categories with large differences across the five age groupings used in post-stratification weighting (i.e., aged 18-29, 30-39, 40-49, 50-64, and 65 or older). Using the most common group seeing differences as an example, those aged 40-49 (see Exhibit 3), we see a longer period of overestimation between the late 1980's through 2000.

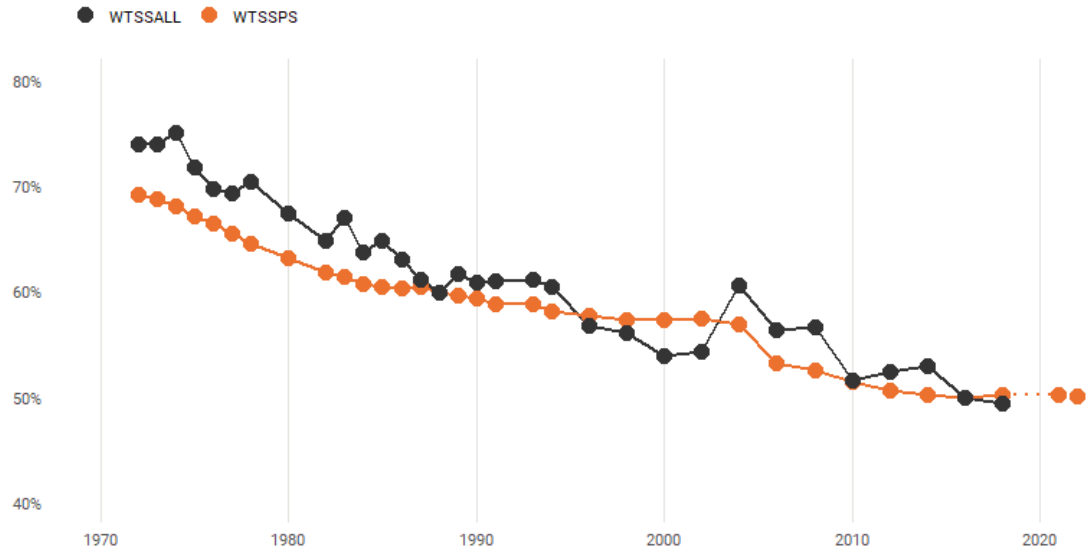
Marital status is the third most common post-stratification variable to see large differences between WTSSALL and WTSSPS estimates. Being married has often been overestimated in the GSS (with exceptions around 2000) with differences from three to seven percentage points (see Exhibit 4).

Exhibit 3. Estimates for aged 40-49 using WTSSALL and WTSSPS



Source: General Social Survey, 1972-2022
Note: Variable AGE, responses 40-49. Estimates exclude reserve codes.

Exhibit 4. Estimates for married respondents using WTSSALL and WTSSPS

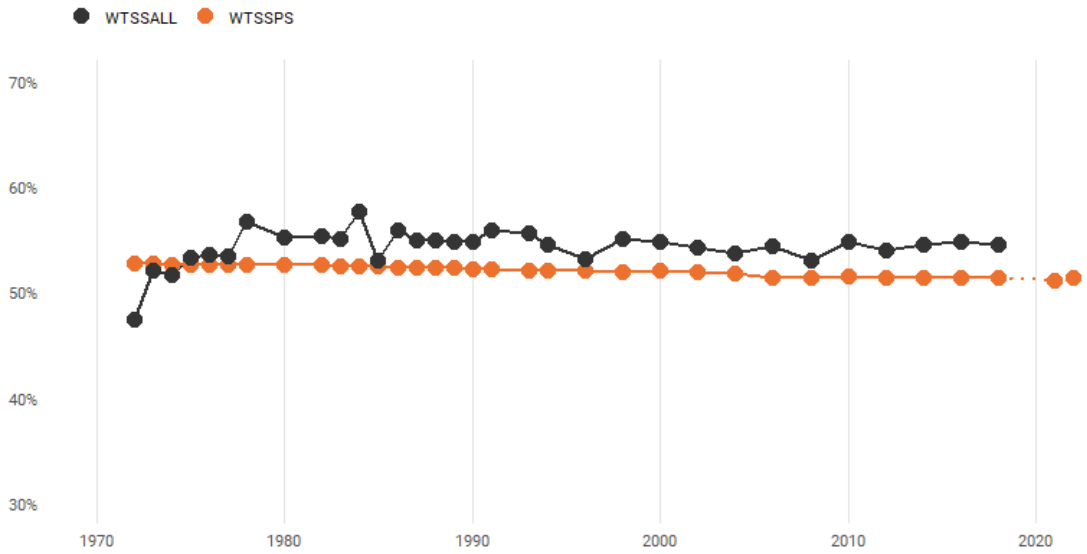


Source: General Social Survey, 1972-2022
Note: Variable MARITAL, response 1. Estimates exclude reserve codes.

Apart from 1972-1974, females have been consistently overestimated in the GSS with 19 of the 32 years showing large differences ranging from two to five percentage points (see Exhibit 5). In addition, the proportion of females has been fairly consistent over the last forty years.

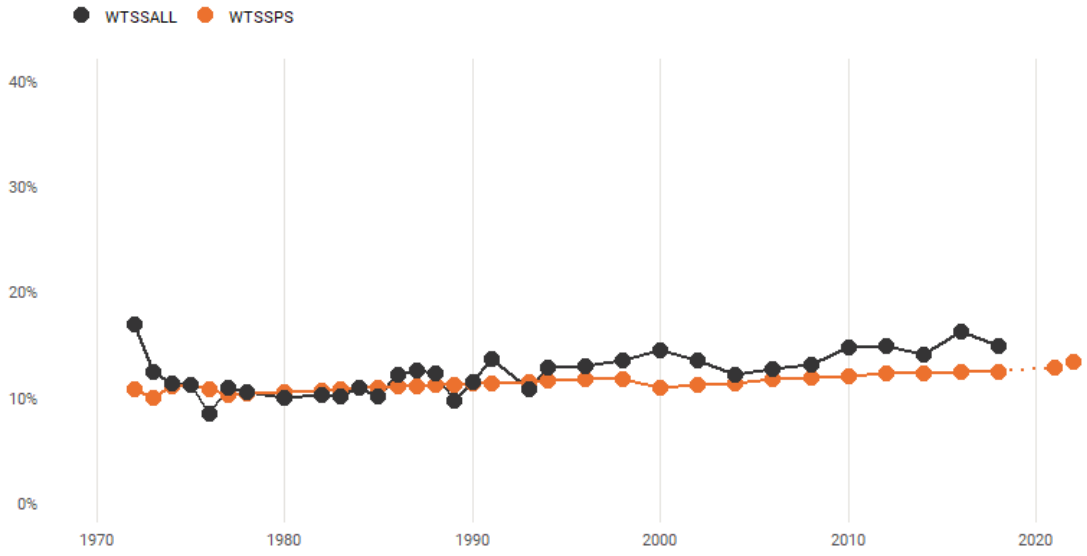
Looking at the simplest form of race reported by the GSS (the variable RACE), we see an overestimation of Blacks particularly over the last two decades (see Exhibit 6). Differences with the "Other" race category see bigger changes, which is related to changing practices with respect to collecting data on racial/ethnic groups. This is particularly evident in 2000 when a big jump occurs in the WTSSPS estimate from around 5 percent up to 10 percent (see Exhibit 7).

Exhibit 5. Estimates for female respondents using WTSSALL and WTSSPS



Source: General Social Survey, 1972-2022
Note: Variable SEX, response 2. Estimates exclude reserve codes.

Exhibit 6. Estimates for Black respondents using WTSSALL and WTSSPS

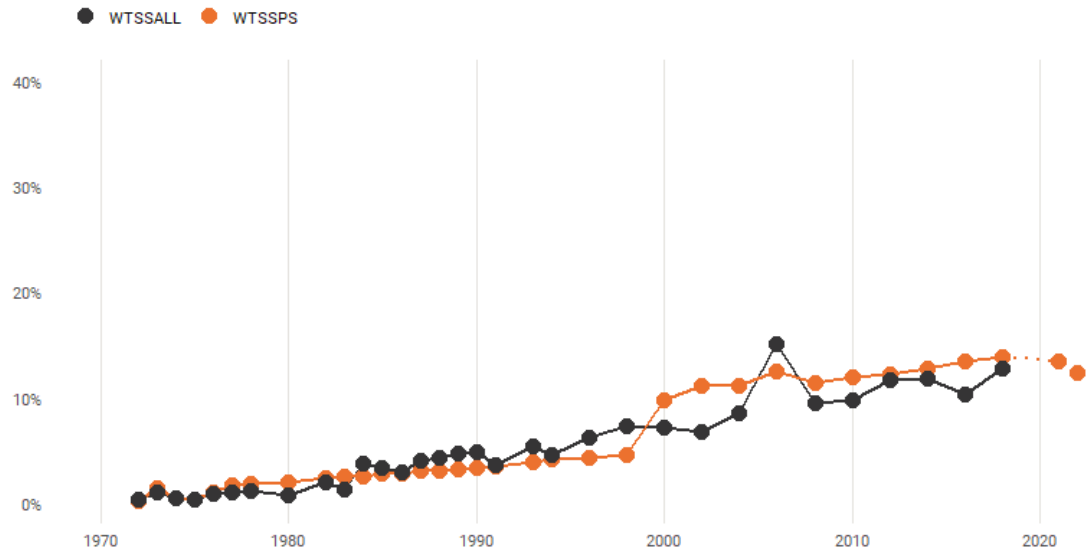


Source: General Social Survey, 1972-2022
Note: Variable RACE, response 2. Estimates exclude reserve codes.

Finally for education, those with less than a high school education as an example (see Exhibit 8), GSS underestimated this nationally from the late 1980's into the mid-2000's.

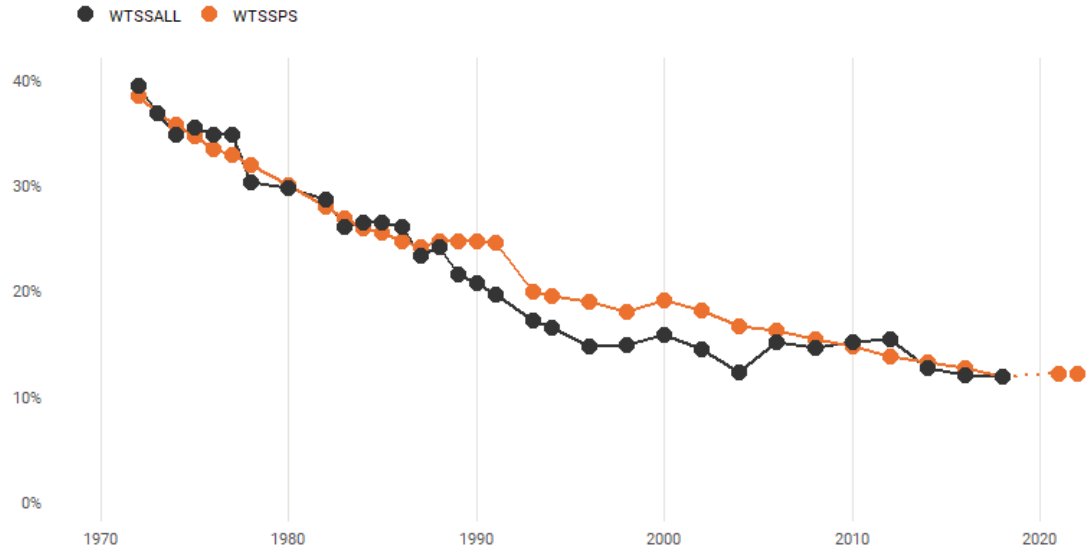
We note that census division (variable REGION) also saw many large differences, heavily concentrated in the 1970's.

Exhibit 7. Estimates for non-white, non-Black respondents using WTSSALL and WTSSPS



Source: General Social Survey, 1972-2022
Note: Variable RACE, response 3. Estimates exclude reserve codes.

Exhibit 8. Estimates for less than high school education using WTSSALL and WTSSPS

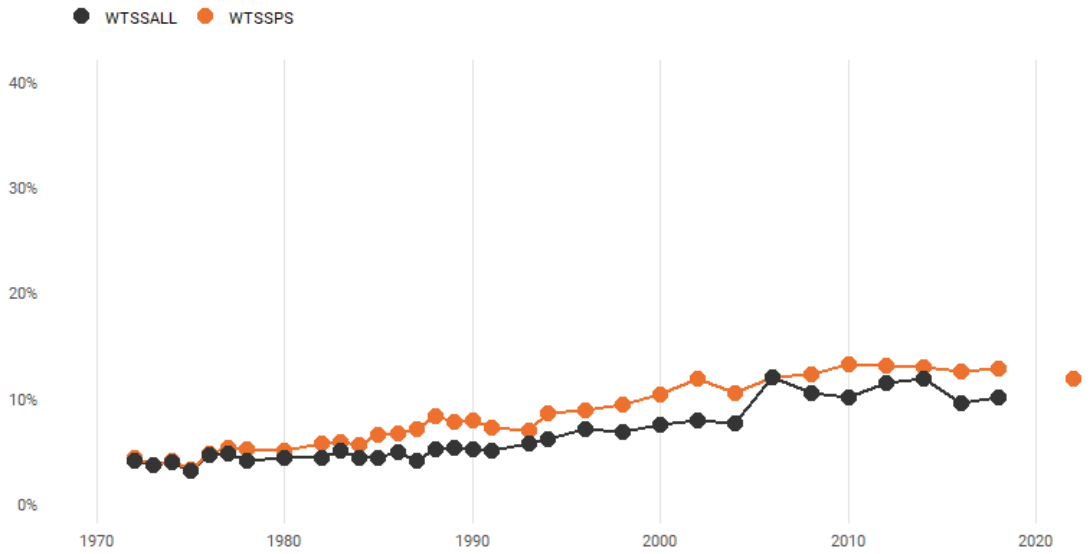


Source: General Social Survey, 1972-2022
Note: Variable DEGREE, response 0. Estimates exclude reserve codes.

Substantive variables

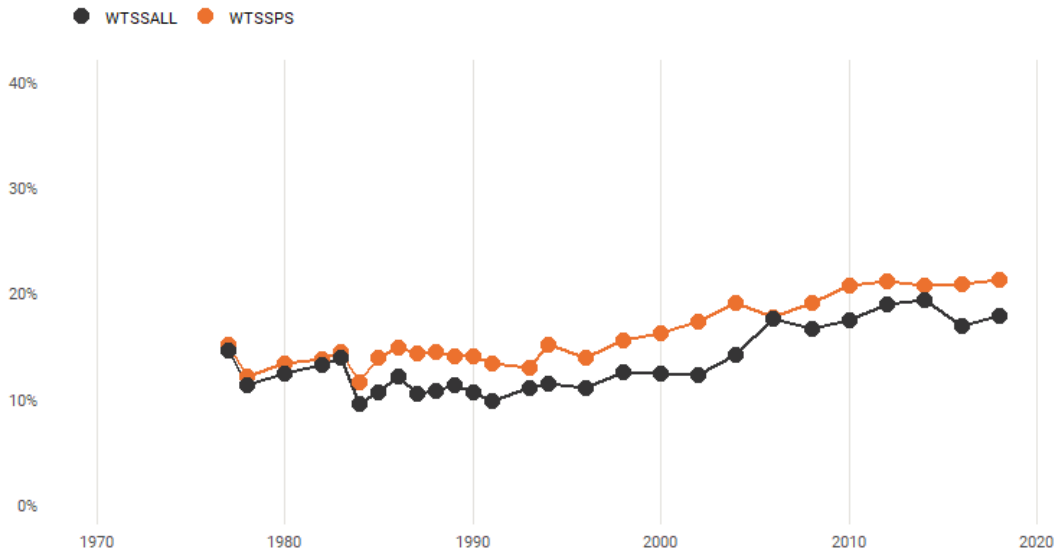
Moving beyond the weighting variables, the variables with large differences were often strongly correlated with U.S. born status including where they lived at age 16 (variable REG16) and whether parents were born in the U.S. (variable PARBORN). Consistent with Exhibit 2, we can see that the historical weights underestimated those living outside of the U.S. at age 16 (Exhibit 9) and having neither parent being born in the U.S. (Exhibit 10).

Exhibit 9. Estimates for those who lived outside of the United States at age 16 using WTSSALL and WTSSPS



Source: General Social Survey, 1972-2018, 2022
Note: Variable REG16, response 0. Corresponding estimate for 2021 is unavailable given a programming error discovered during this analysis. A correction for REG16 is scheduled for Release 3. Estimates exclude reserve codes.

Exhibit 10. Estimates for neither parent born in the United States using WTSSALL and WTSSPS

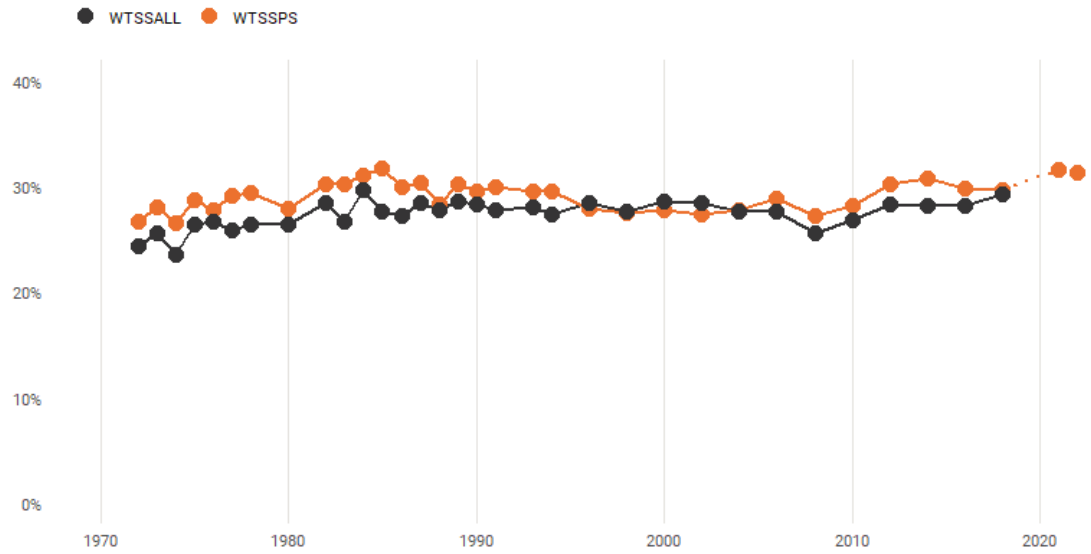


Source: General Social Survey, 1972-2018
Note: Variable PARBORN, response 8. PARBORN was replaced by PABORN and MABORN in 2021 and 2022. Estimates exclude reserve codes.

Beyond those close correlates to U.S. born status, the presence of children in the household is another that sees differences primarily in the early years of the GSS. We see an underestimation of childless households in the first two decades of the GSS ranging from a two to four percentage point difference (see Exhibit 11).

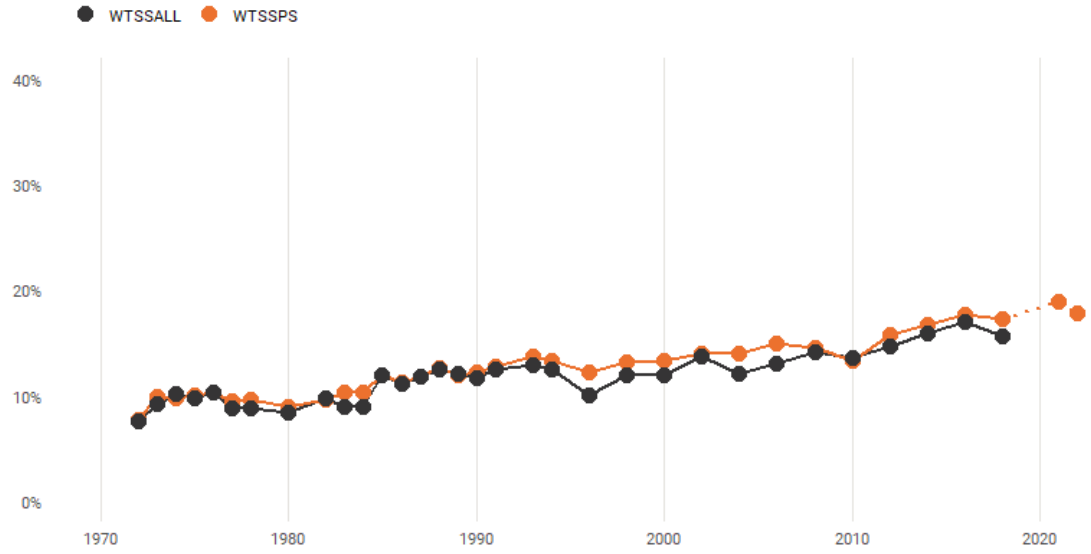
Current work status (variable WRKSTAT) also saw five years where at least one category had a large absolute difference-to-standard error ratio. Most often the historical weights underestimated retired persons (see Exhibit 12). The years where retired does have a large change with the new weight do correspond with changes in the estimates for those aged 65 and older.

Exhibit 11. Estimates for those without children using WTSSALL and WTSSPS



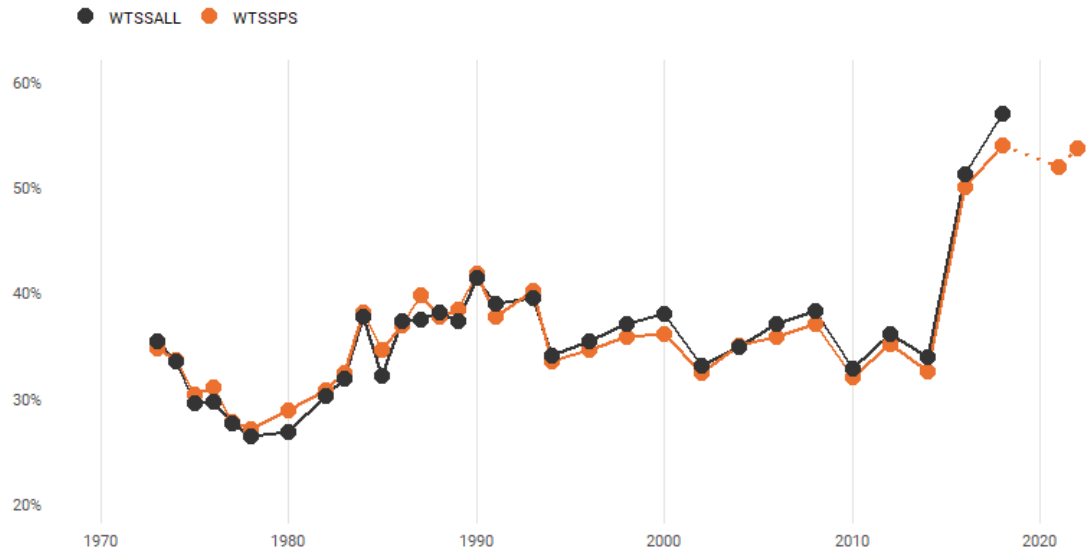
Source: General Social Survey, 1972-2022
Note: Variable CHILDS, response 0. Estimates exclude reserve codes.

Exhibit 12. Estimates for those retired using WTSSALL and WTSSPS



Source: General Social Survey, 1972-2012
Note: Variable WRKSTAT, response 5. Estimates exclude reserve codes.

Exhibit 13. Estimates for too little national spending on improving conditions of Blacks using WTSSALL and WTSSPS



Source: General Social Survey, 1973-2022
Note: Variable NATRACE, response 1. Estimates exclude reserve codes.

The remainder of substantive variables cover a variety of topics but have less than five years' worth of large absolute difference-to-standard error ratios across the years considered, with most only having one year. The variables covered above account for two-thirds of the large response differences we observed. With nearly 70 percent of variables seeing no absolute difference-to-standard error ratios greater than 1.96, we anticipate only minor shifts with the adoption of WTSSPS over WTSSALL (for example, see variable NATRACE in Exhibit 13).

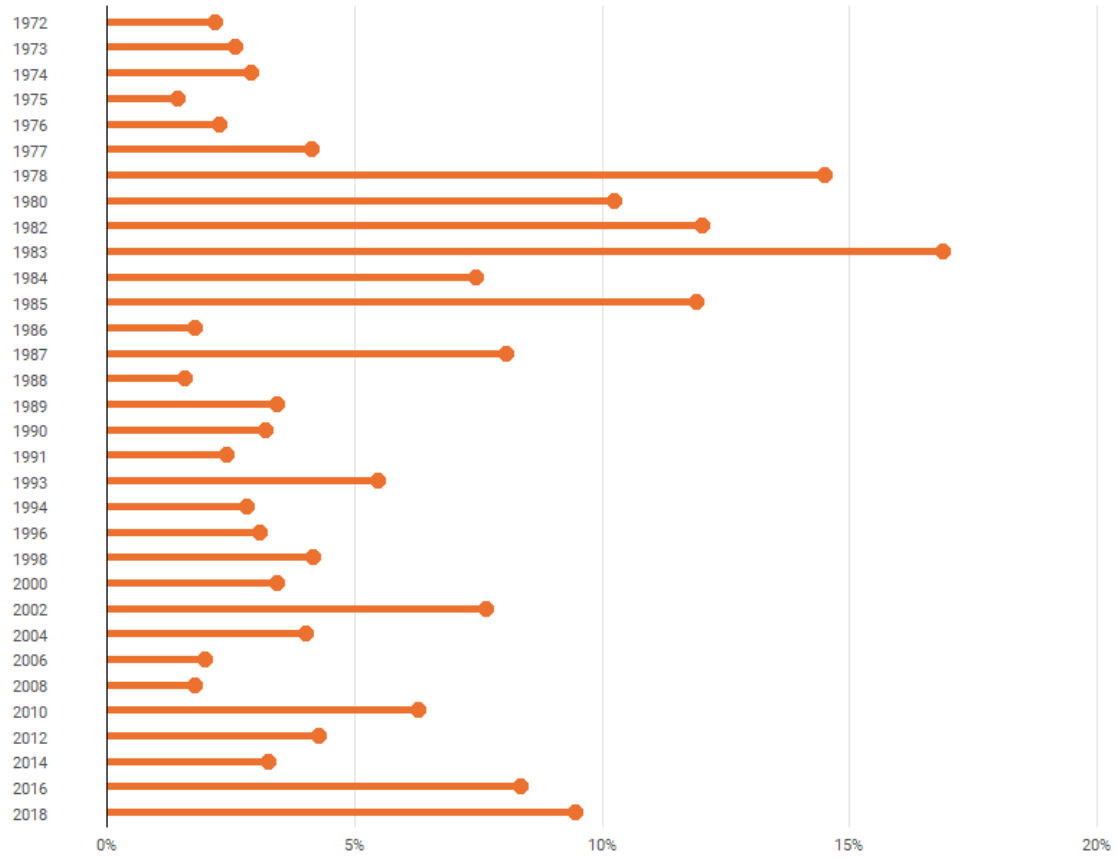
Relative standard error

Finally looking at estimate standard errors, we found that the median relative standard error per year increased by about seven percent on average switching from WTSSALL to WTSSPS (see Exhibit 14). Given this is a median by year, we know that many estimates had reduced standard errors up to 50 or 60 percent. Other estimates experienced standard error increases by double or more, but this was typically related to reserve codes which have extremely small sample sizes.

DISCUSSION AND CONCLUSIONS

In general, the use of post-stratification weights does little to substantively change key historical estimates. In the case of post-stratification dimensions like U.S. born status, age, marital status, and sex, there are marked improvements reducing both over- and underestimation of different variables. Many of the shifting historical univariate trends are strongly tied with variables like U.S. born status or age. While GSS data users may see some increases in estimate standard errors, these increases should be relatively small for most estimates.

Exhibit 14. Median relative change in standard errors from WTSSALL to WTSSPS by GSS year



Source: General Social Survey, 1972-2018

One shift for GSS data users will be the inclusion of the Black oversamples in 1982 and 1987. However, none of the analyses we conducted saw notable shifts for race related variables (see Exhibit 13) which is reassuring for data users.

This analysis has some limitations. Our analytic set of variables is limited and does not represent the full list of possible variables. A future analysis could explore more variables. However, given the post-stratification variables were the primary source of changes and were universally covered in our analysis, we do not anticipate a change in the patterns observed for non-weighting variables. This analysis also does not explore multivariate relationships (e.g., cross-tabulations, subgroup analyses) with substantive variables which may. Future analyses should explore how the change in factors like U.S. born status, age, or sex have on the relationships among other variables. In addition, our analysis does not explore differences between WTSSNR and WTSSNRPS for 2004 to 2018. A future analysis could confirm that these findings are consistent with those weights as well. Given the strong correlation between WTSSALL and WTSSNR (0.96), and WTSSPS and WTSSNRPS (0.98), we feel that similar results would come from such an exploration.

When conducting analyses for long-term GSS trends, GSS data users should note that while the target population for the GSS for all years has been the U.S. 18 and older population, there have been changes in eligibility for the GSS over time. For example, a particularly notable change in eligibility for the GSS is the addition of Spanish speakers in 2006. For all GSS years, post-stratification control totals were developed for the overall 18 and older population to represent the full target population. Nonetheless, changes in eligibility for the GSS may impact studying certain GSS trends.

Based on the results presented in this memo, NORC recommends that data users make use of the new WTSSPS and WTSSNRPS weights for the GSS 1972-2022 Cumulative data file. For analyses conducted from 1972-2022, WTSSPS should be the default weight including for the GSS Data Explorer. We recommend researchers analyzing data from 2004-2022 to use WTSSNRPS to enjoy the benefits of the nonresponse adjustment. The inclusion of all weight adjustments for ballot and form issues and oversamples makes these weights analytically simpler and superior for all types of data users.

ACKNOWLEDGEMENT

The authors wish to acknowledge our reviewers for their contributions to this report: René Bautista, Michael Davern, Jeremy Freese, Pamela Herd, Stephen Morgan, and Susan Paddock. We would also like to thank the other members of the GSS Statistics Team who provided guidance during the initial phases of this analysis: Ned English, Beth Fisher, Colm O'Muircheartaigh, Steven Pedlow, and Jodie Smylie. Finally, we would like to thank the following GSS team members for their contributions to developing the historical post-stratification weights: Erika Archie, Nick Davis, and Taylor Wing. This report was prepared with support from the National Science Foundation grant award 2049169.

The recommended citation for this report is:

Wells, Brian M.; Seeskin, Zachary H.; and Ihde, Amy. "Post-stratification Weights for GSS 1972-2022." General Social Survey Methodological Report No. 137. Chicago, IL: NORC, 2024.

REFERENCES

Davern, M., Bautista, R., Freese, J., Herd, P., & Morgan, S. L. (2024). *General Social Survey, 1972-2022* (Release 2a) [Data set and code book]. NORC. https://gss.norc.org/Documents/codebook/GSS%202022%20Codebook.pdf

Dever, J. A., & Valliant, R. (2010). A comparison of variance estimators for poststratification to estimated control totals. *Survey Methodology, 36*(1), 45-56.

Lohr, S. L. (2021). *Sampling: design and analysis* (3rd edition). CRC press.

National Research Council. (2013). *Nonresponse in Social Science Surveys: A Research Agenda*. The National Academies Press. https://doi.org/10.17226/18293.

SAMPLE CODE FOR ANALYSIS

SAS and Stata sample code for a data analysis example is provided below. For 1972-2022, WTSSPS is the recommended weight for analysis. For 2004-2022, WTSSNRPS is available in addition to WTSSPS, and users may substitute WTSSPS for WTSSNRPS in the code below.

SAS

```
proc surveyfreq data=library.gss7222_r3 missing nosummary;
cluster VPSU;
strata VSTRAT;
weight WTSSPS*YEAR;
table WRKSTAT;
run;
```

Stata

```
svyset [pweight=wtssps], strata(vstrat) psu(vpsu) singleunit(scaled)
svy: tab wrkstat year, percent col format(%3.0f)
```

APPENDIX A

Methodology for Combining Main Sample and Black Oversample and for Rebalanced Weights

The 1982 and 1987 GSS included both a main sample and a Black oversample. OVERSAMP has been available to conduct weighted household-level analyses of the combined main sample and Black Oversample. For these years, we conducted a weighting adjustment to support person-level analyses by making the weights proportional to the product of OVERSAMP and the number of adults in the household. Weights were normalized to have an average value of 1.

For 1987, the resulting weight was then post-stratified as discussed in the section "New Post-Stratification Weight Variables." As the 1982 weights were also subject to form assignment errors, the resulting weight was then rebalanced as discussed in Appendix B.

APPENDIX B

Methodology for Rebalancing Weights for Form and Ballot-and-Form Assignment Errors

Rebalancing was conducted to adjust WTSSALL for form assignment errors in 1978, 1980, and 1982-1985 and to adjust WTSSALL and WTSSNR for ballot-and-form assignment errors in 2002, 2010, 2012, 2016, and 2018.

Rebalancing was conducted within all values of form for the earlier years and within all combinations of ballot and form in the later years. Adjustments were conducted for potentially affected households based on the number of adults and the number of forms or combinations of ballot and form. For 1978, 1982, 1983, and 1985, there were three forms, and households with either 3 or 6+ adults were potentially affected. For 1980 and 1984, there were two forms, and households either 2, 4, or 6+ adults were potentially affected. For 2002, 2010, 2012, 2016, and 2018, there were two forms and three ballots, and household with either 2, 3, 4, or 6+ adults were potentially affected.

We computed weighted totals for the following variables and raking dimensions across the full sample of affected households[8] using the input GSS weight:

- Marital status (3-level)
- Sex (2-level)
- Age group (5-level)
- Working status (3-level)

Cases with missing values for these variables were imputed with hot deck imputation prior to calculation of the weighted totals. Then within each value of form or combination of ballot and form, the weights for affected households were raked to align with the calculated full sample weighted totals. After raking, the weights for affected households were ratio-adjusted to match the total of the input weights for affected households across all forms and ballots. The resulting weight was then post-stratified as discussed in the section "New Post-Stratification Weight Variables."

[8] For 1983, the GSS sample was drawn from multiple sampling frames and the entire rebalancing process was conducted within each sampling frame for these years.

APPENDIX C

Post-Stratification Dimensions by Year and Source

Post-Stratification Dimension	1972	1973	1974	1975	1976	1977	1978	1980	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1993	1994	1996	1998	2000	2002	2004	2006	2008	2010	2012	2014	2016	2018	2021	2022
Census Division	CPS ASEC																						ACS											
Education[1]	CPS ASEC														Decennial Census					CPS ASEC			ACS											
Marital Status	CPS ASEC																						ACS											
Sex	CPS ASEC																						ACS											
Age Group	CPS ASEC																						ACS											
Race[2]	CPS ASEC																						ACS											
U.S. Born Status[3]	*Not available*									Decennial Census											CPS ASEC		ACS											
Hispanic	*Not available*																						ACS											

Notes: CPS ASEC: Current Population Survey Annual Social and Economic Supplement estimates (single year). Data obtained from *IPUMS CPS, University of Minnesota,* www.ipums.org. ACS: American Community Survey estimates. For 2000-2018, ACS 1-year estimates were obtained from *IPUMS USA, University of Minnesota,* www.ipums.org. For 2021 and 2022, the 2019 5-year ACS Public Use Microdata Sample was used.

[1] There were differences between CPS ASEC education coding and GSS coding prior to 1993 due to the CPS ASEC use of "highest grade attended." 1990 Census numbers were used for 1988 to 1991 GSS post-stratification. In addition, earlier decennial censuses did not have education coding aligning with GSS coding. For 1972 to 1987, NORC determined that the CPS ASEC was the best source despite the CPS ASEC use of "highest grade attended," and NORC aligned the CPS ASEC and GSS coding for post-stratification.

[2] As described in the section "New Post-Stratification Weight Variables", three race categories were used for post-stratification the 1972-2018 GSS. For the 2021 and 2022 GSS, six race categories were used.

[3] U.S. born status was not included in the CPS ASEC until 1994. For GSS post-stratification prior to 1994, the decennial census was used. The decennial census year was chosen to align with the census used for the GSS sampling frame. The 1970 census was used for U.S. born status post-stratification for the 1977-1982 GSS. The 1980 census was used for the 1983-1991 GSS. The 1990 census was used for the 1993 GSS.