



GSS POST-STRATIFICATION WEIGHTS MEMO

METHODOLOGICAL REPORT #135

SUMMARY: This memo describes a new set of post-stratification weights available for users of the 2000–2018 GSS cross-sectional surveys. The weight derivation follows the approach applied to the previously released [2021 GSS Cross-section](#), for which post-stratification weights were developed to improve nonresponse bias adjustment, given the impact of the COVID-19 pandemic on survey operations and response rate. Use of these weights results in weighted totals of each GSS cross-sectional sample that equal marginal control totals from the U.S. Census Bureau estimates for education, sex, marital status, age, region of the country, race, Hispanic origin, and U.S. born status. These weights also correct for the ballot and ballot-and-form assignment errors reported in [GSS Methodological Report 134](#) for 2002, 2010, 2012, 2016, and 2018. 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, given the global trend of declining response rates over the past several years.

NEW WEIGHT VARIABLES

The new weights, WTSSPS and WTSSNRPS, are the new post-stratified weights. WTSSPS is available for years 2000–2018, and WTSSNRPS is available for 2004–2021.¹

In fall of 2022, two updated weights, BALLOTFORMWT and BALLOTFORMWTNR, will also be included in Release 3 of the 2021 GSS cumulative file to adjust for the form and ballot assignment errors in 2002, 2010, 2012, 2016, and 2018. These weights are the assignment error-adjusted versions of WTSSALL and WTSSNR,² respectively. The difference between these weights is that WTSSNR includes a national frame area (NFA)-level nonresponse adjustment. The full definitions of WTSSALL and WTSSNR are described in the [GSS Codebook, Appendix A](#).

¹ WTSSNR was introduced in 2004 and, along with WTSSNRPS, is not available for prior years.

² WTSSNR was introduced in 2004 and thus is not available for 2002.

WTSSPS and WTSSNRPS were derived by raking WTSSALL (or BALLOTFORMWT in 2002, 2010, 2012, 2016, and 2018) and WTSSNR (or BALLOTFORMWTNR in 2002, 2010, 2012, 2016, and 2018), respectively, to U.S. Census control totals for the following characteristics:

- Census division (9-level)
- Marital status (2-level)
- Hispanic origin (2-level)
- Education (3-level)
- U.S.-born status (2-level)
- Gender (2-level)
- Race (3-level: White, Black/African American, and Other/Multiple)
- Age group (5-level)

These characteristics are the same as those used for post-stratification of the 2021 GSS Cross-section except here a three-level rather than a six-level race variable was used for raking.

Moving forward, WTSSNRPS will be the preferred weight for cross-sectional analyses of GSS 2004 and onward and WTSSPS the preferred weight for 2000–2002.

SAMPLE CODE FOR ANALYSIS

The same guidance previously provided for how to use these weights holds; specifically, SAS and Stata sample code for a data analysis example is as follows. For analyses of GSS 2004 through 2021:

SAS

```
proc surveyfreq data=library.gss7221_r1 missing nosummary;
cluster VPSU;
strata VSTRAT;
weight WTSSNRPS;
table RELIG;
run;
```

STATA

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

Analysts can use the same code as above for the analysis of GSS 2000–2002 with the post-stratified weights except with WTSSNRPS replaced by WTSSPS.