

Release Notes for the GSS 2008 Panel Cumulative File (Release 6)

Wave Affected	Variable	Issue in Release 5	Fix in Release 6
1	LETIN1	Contained values for LETIN1A	LETIN1 Dropped
1	LETIN1A	Contained values for LETIN1	Recoded to LETIN1A
All	ISCO88	Based on old OCC10 codes	Recalculated to new OCC10 Codes
All	SPISCO88	Based on old OCC10 codes	Recalculated to new OCC10 Codes
All	PAISCO88	Based on old OCC10 codes	Recalculated to new OCC10 Codes
All	MAISCO88	Based on old OCC10 codes	Recalculated to new OCC10 Codes
All	ISCO08	Based on old OCC10 codes	Recalculated to new OCC10 Codes
All	SPISCO08	Based on old OCC10 codes	Recalculated to new OCC10 Codes
All	PAISCO08	Based on old OCC10 codes	Recalculated to new OCC10 Codes
All	MAISCO08	Based on old OCC10 codes	Recalculated to new OCC10 Codes
All	COISCO08	Based on old OCC10 codes	Recalculated to new OCC10 Codes
All	PRESTG10	Based on old OCC10 codes	Recalculated to new OCC10 Codes
All	SPPRES10	Based on old OCC10 codes	Recalculated to new OCC10 Codes
All	PAPRES10	Based on old OCC10 codes	Recalculated to new OCC10 Codes
All	MAPRES10	Based on old OCC10 codes	Recalculated to new OCC10 Codes
All	COPRESTG10	Based on old OCC10 codes	Recalculated to new OCC10 Codes
All	PRESTG105PLUS	Based on old OCC10 codes	Recalculated to new OCC10 Codes
All	SPPRES105PLUS	Based on old OCC10 codes	Recalculated to new OCC10 Codes
All	PAPRES105PLUS	Based on old OCC10 codes	Recalculated to new OCC10 Codes
All	MAPRES105PLUS	Based on old OCC10 codes	Recalculated to new OCC10 Codes
All	COPRES105PLUS	Based on old OCC10 codes	Recalculated to new OCC10 Codes
All	SEI10	Based on old OCC10 codes	Recalculated to new OCC10 Codes
All	SPSEI10	Based on old OCC10 codes	Recalculated to new OCC10 Codes
All	PASEI10	Based on old OCC10 codes	Recalculated to new OCC10 Codes
All	MASEI10	Based on old OCC10 codes	Recalculated to new OCC10 Codes
All	COSEI10	Based on old OCC10 codes	Recalculated to new OCC10 Codes
All	SEI10EDUC	Based on old OCC10 codes	Recalculated to new OCC10 Codes
All	SPSEI10EDUC	Based on old OCC10 codes	Recalculated to new OCC10 Codes
All	PASEI10EDUC	Based on old OCC10 codes	Recalculated to new OCC10 Codes
All	MASEI10EDUC	Based on old OCC10 codes	Recalculated to new OCC10 Codes
All	COSEI10EDUC	Based on old OCC10 codes	Recalculated to new OCC10 Codes

Wave Affected	Variable	Issue in Release 5	Fix in Release 6
All	SEI10INC	Based on old OCC10 codes	Recalculated to new OCC10 Codes
All	SPSEI10INC	Based on old OCC10 codes	Recalculated to new OCC10 Codes
All	PASEI10INC	Based on old OCC10 codes	Recalculated to new OCC10 Codes
All	MASEI10INC	Based on old OCC10 codes	Recalculated to new OCC10 Codes
All	COSEI10INC	Based on old OCC10 codes	Recalculated to new OCC10 Codes

Release Notes for the GSS 2008 Panel (Release 5)

Sample Affected	Variable	Issue in Release 7	Fix in Release 8
2008 Panel	MAEDUC	1 case in wave 3 incorrectly coded NA ¹	Recoded to 7
2008 Panel	RACTIRED	3 cases in wave 3 incorrectly coded NA ¹	Recoded to IAP ²
2008 Panel	ETHVSRAC	3 cases in wave 3 incorrectly coded NA ¹	Recoded to IAP ²
2008 Panel	NORACISM	3 cases in wave 3 incorrectly coded NA ¹	Recoded to IAP ²
2008 Panel	RACRESNT	1 case in wave 2 incorrectly coded IAP ²	Recoded to NA ¹
2008 Panel	RACEXCUS	1 case in wave 2 incorrectly coded IAP ²	Recoded to NA ¹
2008 Panel	DONATE	4 cases in wave 2 incorrectly coded NA ¹	Recoded to IAP ²
2008 Panel	REFAPPLY	1 case in wave 3 incorrectly coded IAP ²	Recoded to NA ¹
2008 Panel	CONSENT	Wave 2 and 3 cases incorrectly coded NA ¹	Recoded to IAP ²
2008 Panel	COHORT	26 cases in waves 1-3 miscalculated	Recoded to correct
2008 Panel	SPRACE2	All cases with code 16 or code 99 were given co	Recoded to separate values
2008 Panel	OCC10	1 case coded 996 instead of 9999	Recoded to 9999
2008 Panel	INDUS10	1 case coded 996 instead of 9999	Recoded to 9999
2008 Panel	PAIND10	1 case coded 996 instead of 9999	Recoded to 9999
2008 Panel	PAOCC10	1 case coded 996 instead of 9999	Recoded to 9999
2008 Panel	SPIND10	1 case coded 996 instead of 9999	Recoded to 9999
2008 Panel	SPOCC80	5 cases in wave 2 had OCC & IND transposed	Recoded OCC with IND value
2008 Panel	SPIND80	5 cases in wave 2 had OCC & IND transposed	Recoded IND with OCC value
2008 Panel	VETYEAR2	8 cases in wave 2 coded 0 instead of 99	Recoded to 99
2008 Panel	RACDIFY	RACDIFY codes based on open ended response; different responses produced different initial codes in the 2008 and 2010 panels	Standardized response labels

¹ NA stands for No Answer. Respondents who gave NA responses were eligible for the given question, but did not answer it. Reasons for not answering could include refusing the question, giving a garbled answer, or declining the remaining questions in a given module.

² IAP stands for Inapplicable. Respondents who were IAP were not eligible for the question. Reasons for being ineligible include being on the wrong ballot for a module, giving a disqualifying answer on a screener question, or not having a requisite characteristic, such as being male for a question only asked of females.

Release Notes for the GSS 2008-Sample Panel R4 (June 21, 2017)

1. This release updates the cross-section data up-to-date with the GSS 1972-2016 cumulative file Release 1b.
2. In the panel data, some cases in 2012 (Wave 3) were incorrectly assigned to military occupation/industry codes which are now fixed to civilian codes.

Release Notes for the GSS 2008 Panel R3 (October 24, 2016)

1. This release updates the cross-section data up-to-date with the GSS 1972-2014 cumulative data file Release 6b.
2. RACDIFY_1 to RACDIFY_3 are added, which were omitted in the previous release.

Release Notes (July 27, 2016)

This release has updated the cross-section data up to date. Most notably, the new occupation and industry codes as well as new prestige and SEI scores are added. For more information, please refer to the previous release notes for the cumulative data for a history of fixes.

GSS 2008 Sample Panel Wave 3, Release 1

August 6, 2013

I. Overview

This GSS panel dataset has three waves of interviews: originally sampled and interviewed in 2008, for the second wave in 2010, and the third interview in 2012. Among the 2,023 cases newly interviewed in 2008, we ended up re-interviewing 1,581 cases in 2010, and 1,295 in 2012 (see Table 1). This data file contains those 2,023 respondents and those variables that were asked in any of the three waves.

<Table 1> GSS Design Features: Cross-Sectional and Panel Components

	GSS Year		
	2008	2010	2012
1 st wave	2023	2044	1974
2 nd wave	1536	1581	1551
3 rd wave		1276	1295
Combined N	3559	4901	4820

II. Data File Organization

1. The released data file is in the wide format: cases in rows and variables of each wave in columns.
2. To denote waves, we have added a suffix “_1” or “_2” to the existing GSS variable names. For example, EDUC_1 is the years of education in the first wave (2008), and EDUC_2 is education in the second wave (2010).
3. The values of the following variables do not change over waves so they are included as single variables (without _1 or _2): BALLOT, FORM, FORMWT, OVERSAMP, SAMCODE, and SAMPLE.
4. YEAR_1 is the GSS year of the first wave while YEAR_2 and YEAR_3 are GSS year of the second and third wave.
5. ID_1 is the identification number used in the GSS 2008 data, ID_2 in 2010, and ID_3 in 2012. ID generally changes across years.
6. PANSTAT_2 indicates panel selection status. Users can identify those cases that were: (1) selected, eligible for re-interviews, and actually re-interviewed; (2) selected, eligible, but

not re-interviewed; and (3) selected, but not eligible and not re-interviewed. If we have more information about why the selected cases were not eligible, we used codes 31 through 33 instead of 3 in the data set (codes labeled).

7. For those cases that were not re-interviewed in the second or the third wave, values in all variables are coded to “Inapplicable (IAP)” (actual codes vary by variables).
8. The variables related to respondents’ household members (e.g. OLD₁ to OLD₁₄, GENDER₁ to GENDER₁₄) do not necessarily indicate the same persons over waves. For example, GENDER_{3_1} and GENDER_{3_2} do not necessarily show the gender of the same household member.
9. Interviewers’ ID numbers (INTID) were newly assigned in each wave. Thus, INTID₁=56 and INTID₂=56 do not indicate they are the same interviewer.
10. COHORT reflects year of birth for respondents age 18-89 on AGE. Respondents older than 89 are coded as 89 on AGE and for them COHORT does not match year of birth, but a somewhat more recent year due to the top coding of AGE at 89. Re-interview cases that are older of than 89 are coded to reflect their COHORT at Wave 1 since the top coding of age at 89 prevents their aging from showing up in the data. For example, a respondent who was 90 in the first wave (and top coded as 89 on AGE₁) would have COHORT₁ as 1917 in GSS 2006. If he was re-interviewed in the second and the third wave and reported ages as respective 92 and 94, he would have been top coded as 89 on AGE₂ and AGE₃ and his COHORT₂ and COHORT₃ would be 1917 in the second and third waves.

III. Weights

Three wave panel data include four different weights: WTPAN₁₂, WTPAN₁₂₃, WTPANNR₁₂, and WTPANNR₁₂₃. The four different weight variables are differentiated by whether they include NR in the variable name or not, or if they use 12 or 123 at the end of variable name. First, weight variables with NR indicate if the weight variables considered a nonresponse adjustment in addition to selection. Second, while weight variables with 12 indicate two-wave panel (2008-10), weight variables with 123 indicate the three-wave panel, which are 2008-10-12 panel data. Below is the description of four weight variables and methodology used to calculate these weights.

<Table 2> GSS panel data three wave weight variables

Variable name	Description
Weight Variable Name without "NR"	This weight is assigned to the cases originated from 2008. It accounts for all four stages of selection for the 2008 samples (NFA, segment, HU, and respondent) and also for the selection of the segment and the case into the panel sample.
wtpan12	If you want to analyze only 2008-10 panel data, you need to use this weight variable
wtpan123	For an analysis of all three waves, you need to use this weight variable.
Weight Variable Name with "NR"	This weight has the same case base as WEIGHTpanel2008 and also includes all stages of selection, <i>but also includes a nonresponse adjustment.</i>
wtpannr12	If you want to analyze only 2008-10 panel data, you need to use this weight variable
wtpannr123	For an analysis of all three waves, you need to use this weight variable.

Selection of respondents from the 2008 round for the panel was done in three phases: first, we selected segments; second, we selected cases within those segments; and third, we selected the completes. To calculate weights for the panel cases, we simply adjusted the 2008 design weights $W_{3NR_{2008XSec}}$ to account for these additional stages of selection. Only completes from 2010 were fielded in 2012. So, with the extra step of nonresponse adjustment, $WA_{panel2008}$ (WTPAN123) is simply:

$$WA_{panel2008} = \frac{W_{3NR_{2008XSec}}}{(\pi_{segment}^{panel2008} \cdot \pi_{case}^{panel2008})} \cdot \frac{1}{mean_h(\overline{RP}_{2008 \text{ panel in 2012}})}$$

The final weight WTPAN123 is just these WA variables, rescaled to sum to the number of completed 2008-sample cases in 2012. Cases originated from 2008 would sum to 2008 totals.

To adjust WTPAN123 for nonresponse in 2012 (WTPANNR123), we use logistic regressions to predict a set of response propensity scores for the 2008 panel cases. The independent variables in this regression are factual (not attitudinal) responses collected in 2010. The variables used were: born outside the U.S., living alone, gender, race (white / nonwhite) and Census division. We used

the predicted response propensities from this model to divide the responding and nonresponding cases into five equal size adjustment cells; and, within each cell, we inflated the weights of the responding cases by the inverse of the mean response propensity in that cell.

$$WA_{panel} = \frac{WA_{2008}}{(\pi_{segment}^{panel} \cdot \pi_{case}^{panel})} \cdot \frac{1}{mean_h(\widehat{RP}_{2008 \text{ panel in 2012}})}$$

where h is the nonresponse adjustment cell. This weight was then, again, scaled to sum to the number of completed panel cases.