The General Social Survey

GSS

2016-2020 GSS
PANEL

CODEBOOK
Release 1

NORC at the University of Chicago

CONTENTS

Citation of this document iv
Copyright 2021 NORC iv
INTRODUCTION 1
2020 Overview 1
Panel Overview 1
Figure 1: Configuration of 2016-2020 GSS Panel 2
Table 1. Key Aspects of the 2016-2020 GSS Panel 3
MAJOR CHANGES TO THE GSS IN 2020 4
Changes Related to Mode of Data Collection 4
Don't Know and No Answer Responses 4
Volunteered Responses 5
SEX 5
Optimization of Items for Web Mode 6
Screenshots of changes 6
Figure 2: Visual display of a survey question in the 2016-2020 GSS Panel Wave 2 6
Changes Related to Shifts to Panel Survey Methodology 7
Changes in Occupation and Industry 7
DATA FILE STRUCTURE 9
Table 2. Panel Status for all Respondents to Wave 1 of the GSS 2016-2020 Panel 9
Table 3. Data File Structure 9
New Reserve Codes 10
2016-2020 GSS PANEL WAVE 2 SPECIFIC NOTES 11
GSS Layout & Questionnaires 11
Figure 3: Module Layout for 2016-2020 GSS Panel 11
New Experiments in the 2016-2020 GSS Panel Wave 2 12
Table 4: Experiments and Variables 12
GSS-ANES Collaboration 13
Figure 4: Configuration of the 2016-2020 GSS Panel and 2016-2020 GSS-ANES Panel 14
Future Releases 15
FIELDING 16
Figure 5: Recuritment Protocol for Mail 16
SAMPLE DESIGN 18
Sampling 18
The 2016 GSS Cross-Sectional Sample 18
The 2018 GSS Cross-Sectional Sample 19
The 2016-2020 GSS Panel Wave 2 Sample 19

Retention Rates 19
Weights 20
Table 5: NFA-Level Response Rates 21
HOW TO READ GSS CODEBOOK TABLES 22
Table 6: Example Codebook Table - NATPARK 23
APPENDIX A: GSS PANEL FREQUENCIES 25
APPENDIX B: WEIGHT SUMMARIES 4584
Table 7 B1: Weight Summary 4584
Table 7 B2: Weight Summary for Completes Only 4584
Table 7 B3: Variance 4585
APPENDIX C: SAMPLE CODE 4586
SAS 4586
STATA 4586

INTRODUCTION

2020 Overview

The General Social Survey (GSS), a national representative survey of the attitudes and behaviors of adults in the United States, has taken place annually or biennially since 1972.

Due to the onset of the COVID-19 pandemic in the United States in the early months of 2020, GSS staff redesigned the 2020 GSS in several ways to ensure the safety and wellbeing of all people who participate in and administer the study.

In the year of 2020, we conducted the GSS as two studies; namely, (1) a panel re-interview of past respondents from the 2016 and 2018 cross sectional GSS studies (referred to as the 2016-2020 GSS Panel), and (2) an independent fresh cross-sectional address-based sampling push to web study (referred to as 2020 cross-sectional survey). This codebook provides details of the first study; namely, the study empaneling former 2016 and 2018 GSS respondents to answer a GSS questionnaire in 2020 (i.e., the 2016-2020 GSS panel). Documentation for the second study (cross-sectional study) will be provided separately. GSS staff redesigned both the panel and the cross-section study in 2020 to be administered primarily via web self-administered questionnaire, instead of face-to-face interviews, with phone interviews as a secondary mode. In this codebook, the 2020 follow-up survey of 2016 or 2018 GSS cases is referred to as the 2016-2020 GSS Panel Wave 2 (where Wave 1 is the pool of GSS respondents who participated in either the 2016 or 2018 GSS cross sectional surveys).

Each of these major changes had several ramifications for sampling, fielding, questionnaire design, data cleaning, response rates, and weights.

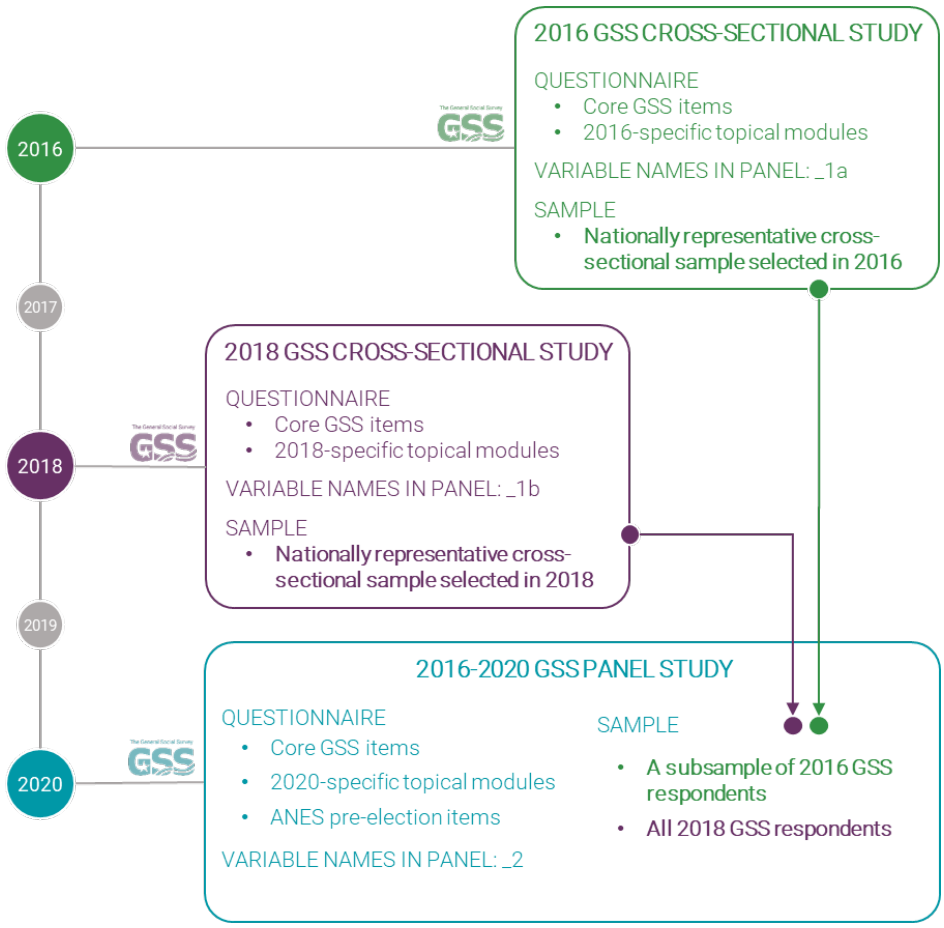
Panel Overview

This codebook focuses on Wave 2 of the 2016-2020 GSS Panel – i.e. the panel reinterviews with 2018 GSS respondents and a randomly selected subset of 2016 GSS respondents. The GSS has used a panel format previously, as parts of the 2006-2014 GSS. The codebook for those panels is available here.

In the 2016-2020 GSS Panel, variables only contain data from one of the three years. To differentiate between versions of each variable, they have been appended with suffixes. Variables from 2016 (Wave 1a) have _1a appended, variables from 2018 (Wave 1b) have _1b appended, and variables from 2020 (Wave 2) have _2 appended. Users can also track cases from 2016 and 2018, and reinterviews from 2020 with the variable SAMPTYPE. Figure 1 shows the relationship of 2016, 2018, and 2020 rounds of data collection for the 2016-2020 GSS Panel.

Wave 1 – First Interview	_1a variables contain data from the 2016 GSS _1b variables contain data from the 2018 GSS
Wave 2 – Follow-up Interview	_2 variables contain data from the 2020 GSS reinterviews

Figure 1: Configuration of 2016-2020 GSS Panel



The 2016 to 2020 GSS Panel Wave 2 interviews included no new respondents—meaning that GSS staff convened the panel in response to the COVID-19 pandemic. Therefore, the 2016 to 2020 GSS Panel represents an ad hoc panel in which respondents from previous rounds (2016 and 2018) were contacted to be administered a questionnaire relevant for 2020. Some of those questions come from the American National Election Studies (ANES) –this is the first time the GSS and ANES have formally collaborated to include ANES questions into a GSS questionnaire. Respondents were located based on prior round contact information, and invited to complete a shortened version of the traditional GSS questionnaire (see Major Changes to the GSS in 2020 for details), either via web-based self-administered interview or phone interview with a trained interviewer using a computer-assisted telephone interviewing (CATI) system. Eligible respondents, based on citizenship status, also received a module for the GSS-ANES collaboration (see GSS-ANES Collaboration), and those who completed that module were later invited to complete the post-election 2020 ANES questionnaire. Table 1 lists key aspects of the 2016-2020 GSS Panel Wave 2.

DATA FILE STRUCTURE

In addition to the panel re-interviews, the 2016-2020 GSS Panel data file contains all information from the 2016 and 2018 GSS. With a few exceptions, variables exist for each year and are distinguished by the variable name suffix (e.g., WRKSTAT_1a, WRKSTAT_1b, WRKSTAT_2). Please see section, How to Read GSS Codebook Tables, for further detail. Respondents who were not re-interviewed in 2020 are still included in their original year, as are the variables in topical modules that were asked in only one year and not replicated elsewhere.

Two variables have been included to help describe the sample of the 2020 GSS Panel. The first, SAMPTYPE, allows users to differentiate between respondents first interviewed in 2016 and those first interviewed in 2018. The second, PANSTAT, allows users to track whether respondents were re-interviewed in 2020. Third, the case ID number found in YEARID can also be used to identify the original year and ID number.

Table 2 provides a full breakdown of PANSTAT.

Table 2. Panel Status for all Respondents to Wave 1 of the GSS 2016-2020 Panel

Panel Cases Status (PANSTAT)	2016	2018	Total
Selected, eligible, and re-interviewed	809	1,014	1,823
Not selected	721	0	721
Selected, but not re-interviewed	1,261	1,267	2,528
Selected, but not eligible and not re-interviewed because R was deceased	61	45	106
Selected, but not eligible and not re-interviewed because R was permanently incapacitated, outside the U.S., or otherwise out of scope	15	22	37
Total	2,867	2,348	5,215

Two variables have been included for weighting purposes: 1) WTSSALL_2 accounts for the original probability of selection in 2016 or 2018, household size, and subsampling design. 2) WTSSNR_2 also includes an additional adjustment for nonresponse at the National Frame Area (NFA) level, giving more weight to respondents who are in NFAs with lower response rates. Please see section Weights for details. The GSS staff recommends using WTSSNR_2 as the standard weight for data analysis in the 2016-2020 GSS Panel Wave 2.

Table 3 provides a preview of the data file structure.

Table 3. Data File Structure

FILEVERSION	PANSTAT	WTSSALL_2	WTSSNR_2	VSTRAT_2	VPSU_2	WRKSTAT_1A	WRKSTAT_1B	WRKSTAT_2
GSS 2020 Panel Release 1 (May 2021)	Selected, eligible, and reinterviewed	1.4318711	1.3578916	3201	1	Working full time	Inapplicable (IAP)	Working full time
GSS 2020 Panel Release 1 (May 2021)	Selected, eligible, and reinterviewed	0.7159355	0.6740571	3201	1	Working full time	IAP	Working part time

GSS 2020 Panel Release 1 (May 2021)	Not selected	I	I	IAP	IAP	Retired	IAP	IAP
GSS 2020 Panel Release 1 (May 2021)	Selected, eligible, and reinterviewed	2.4347396	3.8002517	3301	1	IAP	With a job, but not at work because of temporary illness, vacation, strike	Working full time
GSS 2020 Panel Release 1 (May 2021)	Selected, eligible, and reinterviewed	0.9738958	1.0183642	3301	1	IAP	Retired	Retired
GSS 2020 Panel Release 1 (May 2021)	Selected, eligible, and reinterviewed	0.9738958	0.9339024	3301	1	IAP	Working full time	Working full time

New Reserve Codes

The 2016-2020 GSS Panel now uses a standardized set of reserve codes for all variables:

.d – Don't know (DK). This reserve code is used when respondents indicate that they do not know the answer to a question.

.i – Inapplicable (IAP). This reserve code is used when a respondent does not see a question, either due to structural factors (e.g. for a respondent on BALLOT C, they will be marked inapplicable for a question exclusive to BALLOTs A and B), or to personal factors (e.g. a respondent who is not married will be marked inapplicable for a spouse-related item).

.n – No Answer (NA). This reserve code is used when a respondent refuses to answer a question. In 2020, this code is exclusive to phone mode cases. If a respondent gives a "No Answer" response to a screener question, all follow-up questions are marked as "No Answer."

.s – Skipped on Web. This reserve code is used when a respondent skips a question in the web mode. If a respondent skips a screener question, follow-up questions are marked as "Skipped on Web."

.y – Not available for this year. In previous rounds of the GSS, .i would be used to mark variables which were not asked in that round. However, this lead to confusion, as it meant that .i could be used for three separate meanings. For the 2020 GSS panel, _1a variables can be assigned .y to indicate that they were not asked in 2016 (e.g COVID-related items), _1b variables can be assigned .y to indicate that they were not asked in 2018 (e.g COVID-related items), and _2 variables can be assigned .y to indicate that they were not asked in 2020 (e.g items related to the 2018 International Social Survey Program (ISSP) topical module on religion).

.x – Not available in this release. Variables that were asked in the 2016-2020 GSS Panel Wave 2, but have proven difficult to clean (mostly due to mode changes) have been omitted from the first data release. Instead of dropping the variable entirely, all cases are marked with .x, indicating that the data will be made available in a subsequent release.

2016-2020 GSS PANEL WAVE 2 SPECIFIC NOTES

GSS Layout & Questionnaires

While the GSS is fielded as a single survey, it is actually comprised of three separate ballots. This rotating structure is discussed in greater detail in GSS Methodological Report 52, and current distribution of questions is available in The GSS Replicating Core. The questions that are repeated in each round are known as the Replicating Core. Other sections (which may reoccur, but not every year) are called topical modules. Figure 3 shows the distribution of topical modules in the 2016-2020 GSS Panel.

Figure 3: Module Layout for 2016-2020 GSS Panel

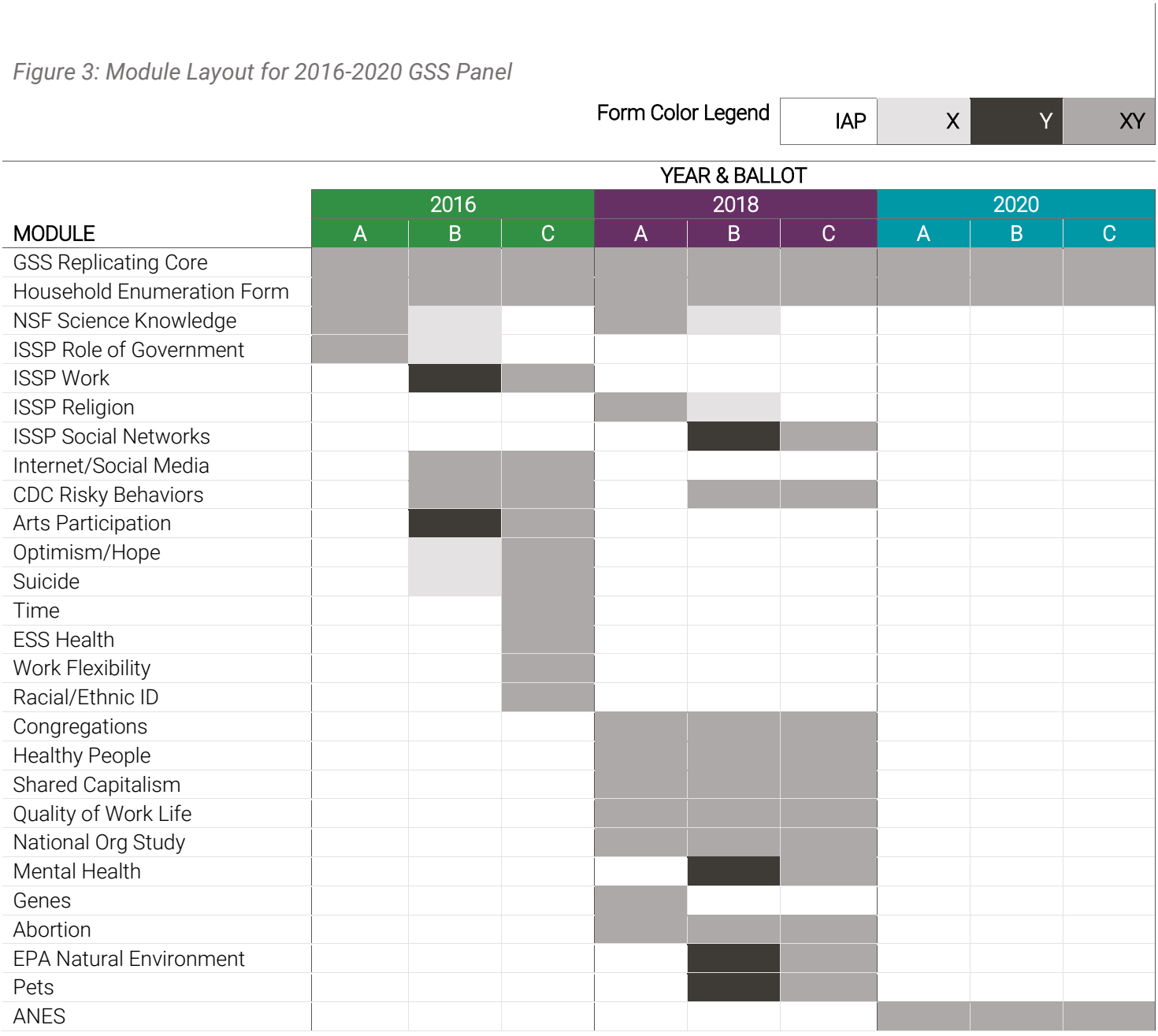


Figure 3: Module Layout for 2016-2020 GSS Panel (Cont'd)

Form Color Legend: IAP | X | Y | XY

MODULE	YEAR & BALLOT 2016 A	2016 B	2016 C	2018 A	2018 B	2018 C	2020 A	2020 B	2020 C
Spirituality									
GSS Board Topical Items									
Contact									
Interviewer Paradata									

The GSS includes slightly different wordings between web self-administered and phone interviewer-administered modes. All questionnaires are available in both English and Spanish. For the 2016-2020 GSS Panel Wave 2, this amounts to 12 different questionnaires: three ballots, two modes, two languages.

The questionnaires can be found https://gss.norc.org/Get-Documentation/questionnaires.

New Experiments in the 2016-2020 GSS Panel Wave 2

The 2016-2020 GSS Panel Wave 2 features two experiments in the Replicating Core that may alter long-term trends for core variables. The first experiment, discussed under Changes Related to Mode of Data Collection, is the volunteered-responses experiment. The second experiment looks at gender-neutral wording for certain items on the GSS: those that refer to men as a default.

For all of these variables, the GSS includes a new version, which is identified with a "Y" at the end. For example, an item using gender-neutral wording displays "police officer" instead of a gendered "policeman." Both sets of experimental variables should be compared to their nonexperimental versions that do not end in "Y" (e.g., LIBMIL vs. LIBMILY, FAIR vs. FAIRY). Each respondent sees only one version of the variable and is recorded as inapplicable on the other version, based on the variable FORM. For example, a respondent with FORM=Y answers SPKRACY and is marked .i on SPKRAC, the FORM=X version of the question.

Table 4 lists the variables affected by both of these experiments.

Table 4: Experiments and Variables

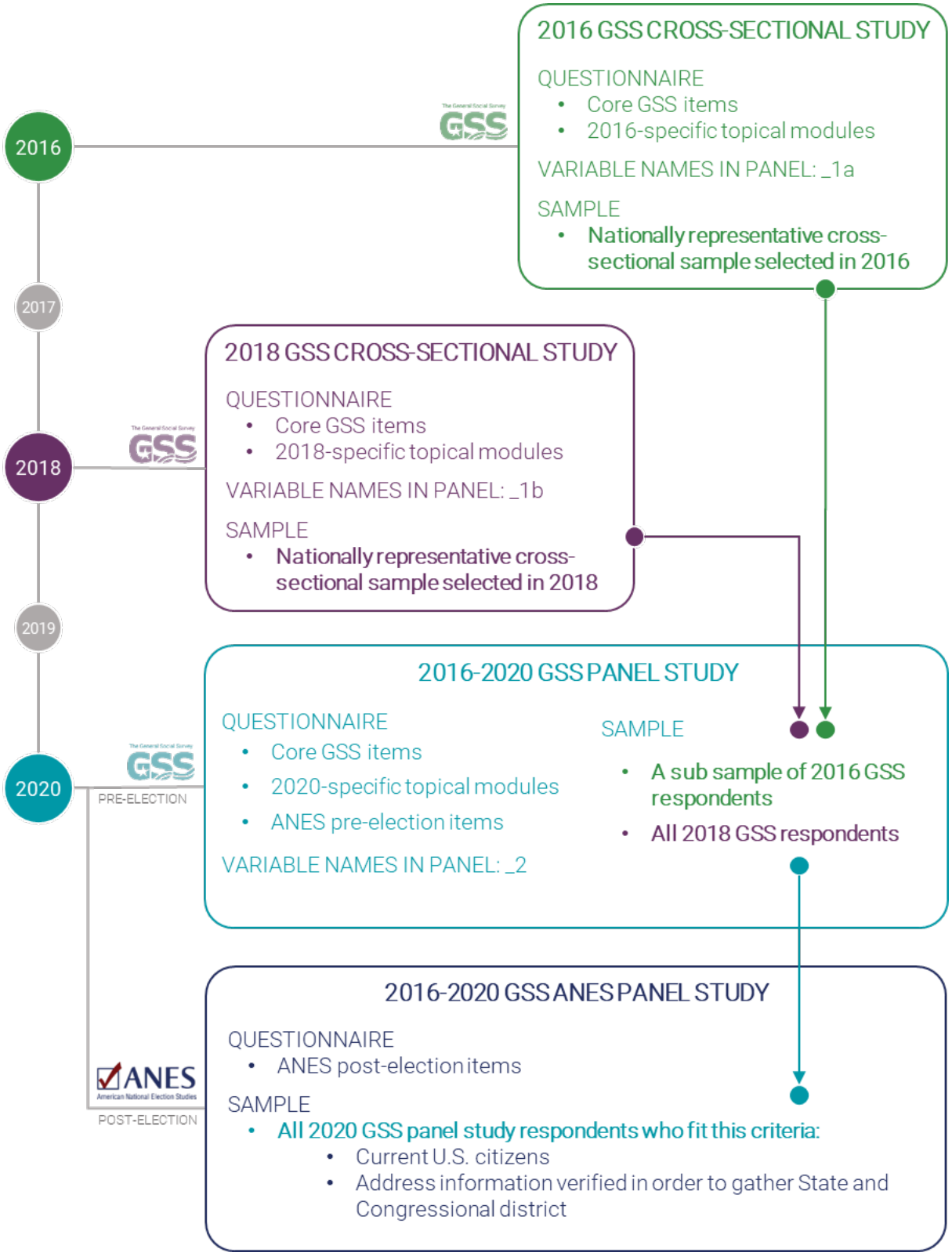
Gender Neutral Wording		
SPKATHY	SPKMILY	POLHITOKY
LIBATHY	LIBMILY	POLABUSEY
SPKRACY	SPKHOMOY	POLATTAKY
LIBRACY	LIBHOMOY	RACOPENY
SPKCOMY	SPKMSLMY	
COLCOMY	LIBMSLMY	
LIBCOMY	LETDIE1Y	
Volunteered Responses		
USWARY	DIVLAW	BIBLE
PRAYER	HELPFUL	POSTLIFE

COURTS	TRUST	KIDSSOL
FEPOL	FAIR	USCITZN
DISCAFFW	AGED	FUCITZN
RACOPEN	GRASS	
GETAHEAD	RELITEN	

GSS-ANES Collaboration

The 2016-2020 GSS Panel Wave 2 includes a collaboration between the GSS and ANES, which was planned before the COVID-19 pandemic and was extensively modified in light of delays in data collection. The 2016-2020 GSS Panel Wave 2 interview contained a module of items drawn from the ANES, including attitudes on several issues, feelings thermometers for the presidential candidates, and plans for voting in the 2020 presidential election. Following the completion of the 2016-2020 GSS Panel Wave 2 in September 2020, all respondents who: 1) were U.S. citizens, 2) could be located within U.S. voting districts, and 3) completed the ANES module were eligible and were contacted to complete the 2020 ANES postelection survey. Any eligible citizen who completed both surveys (i.e., Wave 2 GSS survey and the ANES follow-up questionnaire) can be linked with a key, which will be available later in 2021. For more information, please see the ANES/GSS collaboration codebook, which will be available later in 2021. Figure 4 shows the relationship between the 2016-2020 GSS Panel and the GSS-ANES collaboration.

Figure 4: Configuration of the 2016-2020 GSS Panel and 2016-2020 GSS-ANES Panel



Since the 2016-2020 GSS Panel Wave 2 attempted to re-interview only the respondents from the 2016 and 2018 GSS rounds, the full panel response rate is the initial round response rate multiplied by the panel retention rate. The full-panel weighted response rate for the 2016 GSS round is the cross-sectional weighted response rate of 61.3 percent multiplied by the 2016-2020 GSS Panel Wave 2 retention rate of 38.8 percent = 23.8 percent. The full-panel weighted response rate for the 2018 GSS round is the cross-sectional weighted response rate of 59.5 percent multiplied by the 2016-2020 GSS Panel Wave 2 retention rate of 44.9 percent = 26.7 percent. We do not attempt to combine the two rounds for the full panel response rate.

Weights

The data presented here are intended to be used with weights that adjust for non-response that may have occurred at the National Frame Area level from 2016 and 2018 to 2020. The 2016-2020 GSS Panel Wave 2 uses as its frame the 2016 and 2018 survey completes, thus is not intended as an explicitly representative cross-section sample of adults in the U.S. in 2020. A separate cross sectional survey was conducted in 2020 to be representative of the 2020 population. We recommend using the 2016-2020 GSS Panel to study person-level change resulting from: 1) the passage of time, 2) the COVID-19 pandemic, and 3) the shift from face-to-face (F2F) to web self-administered questionnaire (SAQ) mode.

The 2016-2020 GSS Panel contains several different variables that are useable for weighting and variance estimation, all of which have counterparts in the GSS Cumulative Cross-section. The data include the standard variance stratum (VSTRAT) and variance primary sampling unit (VPSU), which are used to protect against potential inferential disclosure risks. The standard GSS weights, WTSSALL, and WTSSNR, have been modified to work with the 2016-2020 GSS Panel, and are included as WTSSALL_2 and WTSSNR_2. WTSSALL_2 accounts for the sub-sampling of nonrespondents as well as the number of adults in households. WTSSNR_2 also includes an area non-response adjustment based on NFAs. For most users, we recommend WTSSNR_2.These weights were built in several steps, detailed here, with some additional tables in Appendix B:

Step 1. We started with the final weight from each cross-sectional round. The variable is WTSSNR in the 2016 and 2018 cross-sectional weight files. There were 2,867 completes in 2016 and 2,348 completes in 2018. These variables were normalized to sum to their sample sizes, so the sum of the 2,867 completes in 2016 is 2,867, and the sum of the 2,348 completes in 2018 is 2,348. Please note that not all 2,867 completes from 2016 were released for the panel.

Step 2. Only some of the 2016 cases were released for fielding in 2020. We excluded 200 of the completed interviews from 2016 because they were used as a pilot for locating. We split the remaining 2,667 completes into five batches and released only the first four for the 2020 panel. This means that 2,146 out of 2,867 (74.85 percent) of the 2016 completes were released for the 2020 panel. Therefore, we applied a subsampling weight adjustment of 2,867/2,146. For the 2018 completes, all 2,348 were released; therefore, their subsampling weight adjustment is 1.

Step 3. A few panelists were out of the country or deceased, so at this step, these few ineligible cases were given a zero weight. All other cases kept their Step 2 weight.

Step 4. GSS uses an NFA-level nonresponse adjustment. Our first-stage clusters are referred to as National Frame Areas (NFAs), which are counties in rural areas or entire metropolitan statistical areas. We calculated the weighted response rate for each NFA using the calculation in Step 2. The NFA response rates range from 12.84 percent (47 eligibles) to 64.51 percent (38 eligibles).

Importantly, the number of eligible cases ranges from 22 to 238, but the two extremes seem to have enough cases. The Step 2 calculation is divided by the NFA response rate.

Step 5. GSS normalizes the calculation in Step 4 so that the sum of the weights is equal to the number of respondents. However, we believe that analyses of the 2020 panel should be a 50 percent mix of each cohort (2016 and 2018). Therefore, the average normalized weight for the smaller cohort (2016) will be slightly greater than the average normalized weight for the larger cohort (2018). Each cohort has its weight multiplied by a constant so that the sum for the cohort is 911.5 (half of 1,823). These normalized weights are WTSSNR_2 and WTSSALL_2.

Table 5 contains the distribution for the NFA-level response rates among the 4,351 panel eligibles.

Table 5: NFA-Level Response Rates

N	MEAN	STANDARD DEVIATION	MINIMUM	10TH PERCENTILE	25TH Q1	MEDIAN	75TH Q3	90TH	MAXIMUM
4,351	41.5%	10.1%	12.8%	28.8%	34.0%	42.0%	47.8%	54.7%	64.5%

HOW TO READ GSS CODEBOOK TABLES

This 2016-2020 GSS Panel codebook presents all data included in the 2016-2020 GSS Panel Release 1, in alphabetical order.

Variables in the 2016-2020 GSS Panel have been appended with either _1a, _1b, or _2. In this panel, Wave 1 is 2016 (_1a) or 2018 (_1b) or wave 2 (2020).

_1a variables contain data from the 2016 GSS

_1b variables contain data from the 2018 GSS

_2 variables contain data from the 2020 GSS

No respondent will have valid responses in all three variables, since panel respondents will have participated in either the 2016 GSS or the 2018 GSS, but not both. Respondents will have valid responses in either _1a and _2, or _1b and _2, but never both _1a and _1b.

Table 6 shows an example of crosstabs found in Appendix A.

Note from Professor Baird:

When doing analysis in Stata (you would do the R version of this as well,) you have to use their weights. This is the command that must be run, everytime you open your data, before you do any analysis.

```
svyset [pweight=wtssnr_2], strata(vstrat) psu(vpsu) singleunit(scaled)
```

R

```
gss_all <- read_stata("data/GSS7218_R1.DTA")

wt_vars <- c("vpsu",
             "vstrat",
             "wtssnr_2")
```

Table 6: Example Codebook Table - NATPARK

YEAR	NATPARK	Unwgt Freq	Wgt Freq	SE of Wgt Freq	Wgt Perc	SE Wgt Perc
2016	*Don't Know*	48	42.2	6.0	1.5	0.2
	No Answer	2	1.7	1.2	0.1	0.0
	Skipped on Web	0	.	.	.	.
	Too Little	952	963.6	40.3	33.6	1.0
	About Right	1692	1686.0	50.2	58.8	1.0
	Too Much	173	173.5	16.4	6.1	0.5
	Total	2867	2867.0	77.5	100.0	
2018	*Don't Know*	60	58.5	8.3	2.5	0.4
	No Answer	0	.	.	.	.
	Skipped on Web	0	.	.	.	.
	Too Little	839	805.2	30.3	34.3	1.1
	About Right	1346	1383.4	39.7	58.9	1.3
	Too Much	103	100.9	13.0	4.3	0.5
	Total	2348	2348.0	44.9	100.0	
2020	*Don't Know*	11	11.1	3.1	0.6	0.2
	No Answer	1	0.8	0.8	0.0	0.0
	Skipped on Web	7	6.6	2.7	0.4	0.1
	Too Little	655	649.5	28.0	35.6	1.4
	About Right	1081	1087.5	41.8	59.7	1.4
	Too Much	68	67.4	10.2	3.7	0.6
	Total	1823	1823.0	47.4	100.0	
Total	*Don't Know*	119	111.8	10.4		
	No Answer	3	2.5	1.5		
	Skipped on Web	7	6.6	2.7		
	Too Little	2446	2418.3	60.9		
	About Right	4119	4157.0	89.9		
	Too Much	344	341.7	25.7		
	Total	7038	7038.0	121.8		

Table 6 contains seven columns, each with unique information. Column one, Year, shows the year of data collection for the corresponding section of the table. The GSS Codebook uses Year in the place of the _1a, _1b, and _2 labels for codebook tables. Column two shows all response options for the item as labels (and with numeric values for unlabeled responses). Column three shows the unweighted frequency for each response. Column four shows the weighted frequency for each response. Column five shows the standard error of the weighted frequency for each response. Column six shows the weighted percentage for each response. Finally, column seven shows the standard error of the weighted percentage for each response.

Each subsection of the table contains responses from one year of the GSS 2016-2020 Panel: Initial Interviews in 2016 and 2018, and follow-up interviews in 2020. The “Total” section of the table sums to more responses than are found in the dataset, due to double-counting 2020 respondents.

APPENDIX B: WEIGHT SUMMARIES

Table 7 B1 summarizes the weights through their creation:

TABLE 7 B1: WEIGHT SUMMARY

YEAR	Eligible for Panel	Sum WTSSNR	Released cases in 2020	Sum Step 1	Sum Step 2	2020 Panel Completes	Sum Step 4
2016	2,867	2867	2,146	2140.37	2859.48	809	2557.53
2018	2,348	2348	2,348	2348.00	2348.00	1,014	2498.82
TOTAL	5,215	5215	4,496	4488.37	5207.48	1,823	5056.36

Table 7 B2 includes only the 1,823 completes.

TABLE 7 B2: WEIGHT SUMMARY for Completes Only

Year	Eligible for Panel	Sum WTSSNR	Released Cases in 2020	Sum Step 1	Sum Step 2	2020 Panel Completes	Sum Step 4
2016	809	802.78	809	802.78	1072.49	809	2557.53
2018	1,014	1029.01	1,014	1029.01	1029.01	1,014	2498.82
Total	1,823	1831.79	1,823	1831.79	2101.50	1,823	5056.36

Table 7 B3 show tracking the variance across the 2016-2020 GSS Panel steps.

TABLE 7 B3: VARIANCE

	WTSSNR	Step 1	Step 2	Step 4
2016 ONLY	n = 2,867	n = 2,146	n = 2070	n = 809
MEAN WEIGHT	1	1.332	1.336	3.161
STANDARD DEVIATION (SD) OF WEIGHTS	0.538	0.704	0.703	1.846
ESTIMATED DESIGN EFFECT DUE TO WEIGHTING	1.289	1.279	1.277	1.341
2018 ONLY	n = 2,348	n = 2,348	n = 2281	n = 1,014
MEAN WEIGHT	1	1	1.004	2.464
SD OF WEIGHTS	0.645	0.645	0.650	1.610
ESTIMATED DESIGN EFFECT DUE TO WEIGHTING	1.416	1.416	1.419	1.427
2016 & 2018	n = 5,215	n = 4,496	n = 4351	n =1,823
MEAN WEIGHT	1	1.159	1.162	2.774
SD OF WEIGHTS	0.588	0.694	0.696	1.753
ESTIMATED DESIGN EFFECT DUE TO WEIGHTING	1.346	1.358	1.358	1.399

The estimated design effect due to weights is 1 + CV(weights)^2, where the CV(weights) = (standard deviation of weights)/(mean weight)

APPENDIX C: SAMPLE CODE

SAS

```
proc surveyfreq data=library.gss2020panel_r1 missing nosummary;
      cluster VPSU_2;
      strata VSTRAT_2;
      weight WTSSNR_2;
      table wrkstat_2;
run;
```

STATA

```
svyset [pweight=wtssnr_2], strata(vstrat) psu(vpsu) singleunit(scaled)
svy: tab wrkstat_2, percent col format(%3.0f)
```