

GENERAL SOCIAL SURVEYS, 1972-2018:

CUMULATIVE CODEBOOK

March 2019

Conducted for
The National Data Program for the Social Sciences

at

NORC
University of Chicago

NORC

Founded in 1941, the National Opinion Research Center (NORC) is the oldest national survey research facility that is not-for-profit and university affiliated. Through its Survey Operations Group, NORC provides consultation and services on the full range of survey research tasks--study design and analysis, sampling, questionnaire design, data collection, and data processing--for a wide range of clients in the research community. NORC maintains a national probability sample, which is fully revised to accord with 2000 Census information. Many NORC surveys are archived at the Roper Center for Public Opinion Research and in NORC's own data archive.

NORC

University of Chicago

1155 East 60th Street

Chicago, Illinois 60637

(773) 256-6288

EMAIL: smitht@norc.uchicago.edu

GENERAL SOCIAL SURVEYS, 1972-2018 CUMULATIVE CODEBOOK

(Codebook for the Machine-Readable Data File
General Social Surveys, 1972-2018)

Principal Investigator
Co-Principal Investigator
Co-Principal Investigator
Co-Principal Investigator

Tom W. Smith
Michael Davern
Jeremy Freese
Stephen L. Morgan

Research Associates

Jaesok Son
Benjamin Schapiro

Research Assistants

Arnav Chatterjee

NORC Edition

Produced by
NORC
University of Chicago

as part of
The National Data Program for the Social Sciences
2019

This project was supported by
the National Science Foundation
(National Data Program for the Social Science Series, No. 25)

CATALOGING-IN-SOURCE

(Machine-readable data file plus codebook)

Smith, Tom W., Davern, Michael, Freese, Jeremy, and Morgan, Stephen L., *General Social Surveys, 1972-2018* [machine-readable data file] /Principal Investigator, Smith, Tom W.; Co-Principal Investigators, Michael Davern, Jeremy Freese and Stephen L. Morgan; Sponsored by National Science Foundation. --NORC ed.-- Chicago: NORC, 2019.

1 data file (64,814 logical records) + 1 codebook (3,758 pp.). -- (National Data Program for the Social Sciences, no. 25).

Summary: This cumulative data file merges all 32 General Social Surveys (1972-1978, 1980, 1982-91, 1993, 1994, 1996, 1998, 2000, 2002, 2004, 2006, 2008, 2010, 2012, 2014, 2016, 2018) into a single file with each year or survey acting as a subfile. This arrangement of the data facilitates trend analysis on repeated questions over the 46-year period.

+++++

CATALOGING-IN-SOURCE

(Printed codebook only)

Smith, Tom W., Davern, Michael, Freese, Jeremy, and Stephen L. Morgan. *General Social Surveys, 1972-2018: cumulative codebook* / Principal Investigator, Tom W. Smith; Co-Principal Investigators, Michael Davern, Jeremy Freese and Stephen L. Morgan. -- Chicago: NORC, 2019. 3,758 pp., 28cm. -- (National Data Program for the Social Sciences Series, no. 25).

This codebook is to be used in conjunction with the machine-readable data file titled General Social Survey, 1972-2018. 1. Social surveys - U.S. 2. Public Opinion - U.S. I. Smith, Tom W. II. NORC. III. Title. IV. Series.

+++++

BIBLIOGRAPHIC REFERENCE

Manuscripts using this data file should contain the following citation:

Smith, Tom W., Davern, Michael, Freese, Jeremy, and Stephen L. Morgan. *General Social Surveys, 1972-2018*. [machine-readable data file]. Principal Investigator, Tom W. Smith; Co-Principal Investigators, Michael Davern, Jeremy Freese and Stephen L. Morgan, NORC ed. Chicago: NORC, 2019. 1 data file (64,814 logical records) and 1 codebook (3,758 pp).

+++++

ISSN 0161-3340

(Codebook only)

ISBN 978-0-932132-61-1

+++++

This publication may be copied in whole or in part by any copying or duplicating process, provided it is not sold except to recover duplicating costs.

*Cataloging was prepared with the assistance of Sue A. Dodd, Social Science Data Library, Institute for Research in Social Science, University of North Carolina, Chapel Hill

TABLE OF CONTENTS

Introduction.....	viii
Changes in Design.....	x
Abbreviations and Data Identification Numbers.....	xi
Index to Data Set.....	1
1972-2018 Surveys Core Variables and Related.....	122
General Social Survey Topical Modules	
1985, 2004 Social Networks	572
1987, 1994 Socio-Political Participation	612
1988, 2000 Religion	642
1990, 2000 Intergroup Relations	696
1991 Work Organizations	737
1993, 2002 Culture	783
1994 Family Mobility	824
1994 Multiculturalism	904
1996 Gender	930
1996 Emotions	953
1996 Markets	993
1996 Markets and Exchange – Giving and Volunteering	1032
1996, 2006, 2018 Mental Health I.....	1070
1998, 2004 Religion and Culture	1184
1998 Jobs.....	1208
1998 Mental Health II	1223
2000 Multi-Ethnic United States	1274
2000-06, 2010+ Information Society	1346
2000 Freedom	1468
2000 Health Status	1478
2002 Prejudice	1498
2002 Doctors and Patients	1504
2002, 06, 10, 14, 18 Quality of Worklife	1514
2002, 06, 10, 18 Employee Compensation (see Shared Capitalism)	1551
2002, 2004 Altruism	1562
2002 Adult Transitions	1575
2002 Child Stigma	1584
2000 Information Seeking Online	1613
2004, 2012 Workplace Conflict	1619
2004 Immigration (3 card)	1630
2004 Genes and the Environment	1636
2004 Altruistic Love	1653
2004 Guns	1655
2004 Spirituality (Religion II).....	1657
2004 Women and Alcohol	1682
2004 Voluntary Associations/Social Networks	1688
2006-18 Science, Knowledge about and Attitudes towards	1773
2006, 2014, 2018 Shared Capitalism (see Employee Compensation)	1898
2006 Mental Health III	1908
2006 Number Known	1911
2006 Gun Control	1976
2006 Boards/Trends	1979
2006 Disabled	1981
2006 Mental Health IV.....	1984
2008 Self-Employment.....	2030

2008	Jewish Identity	2035
2008	Terrorism Preparedness	2048
2008	Global Economics	2063

International Social Survey Programme (ISSP) Modules (Years listed refer to ISSP years)

1985, 1990, 1996, 2006, 2016	Role of Government	2089
1986	Social Networks & Support Systems	2170
2001	Social Relations & Support Systems	2209
2018	Social Relations & Support Systems	2241
1988, 2000	Social Inequality	2265
1988, 1994, 2002, 2012	Family & Changing Gender Roles	2319
1989, 1997, 2007, 2015	Work Orientation	2391
1991, 1998, 2008, 2018	Religion	2474
1993, 2000, 2010	Environment	2528
1996, 2003, 2013	National Identity	2578
2004, 2014	Citizenship	2631

Additional GSS Modules

2012	Single Variable Modules	2667
2006	Personality Traits	2668
1989	Civil Liberties	2673
1989, 1990	Victims and AIDS	2676
2016	Suicide	2706
2016	Social Media Usage	2715
2016	Hope and Optimism	2728
2016	Arts and Culture	2735
2016	Racial/Ethnic Identity	2759
2016	Flexible Work	2763
2016	ESS Health	2768
2018	Everyday Discrimination	2770
2018	Healthy People	2773
2018	Abortion	2778
2018	Nature	2783
2018	Pets	2787
1988+	(Self-Administered) High Risk Behaviors	2797
2008	(Self-Administered) Sexual Orientation	2816
2008	(Self-Administered) Clergy/Sex	2830
1990	Birth Order	2841
1990, 1996, 2018	Genes	2842
1990	Children	2850
1991, 2004	Negative Life Events	2865
1972-1993	Income	2901
1994-1998	National Security	2908
1972+	Ethnicity	2933
1974+	Military Service and Work	2956
1982+	Residence	2959
1974+	Vocabulary Test	2961
1996	Religious Attendance and Identity	2965
1972+	Household Members and Structure	2976
1975+	Date of Interview	3071
1972+	International Occupation	3072
1972+	War and International	3093
1986	Welfare	3100
2014	Work and Relationships	3103

1972+	Cohort and Birth	3105
2004+	Interviewer Characteristics	3107
2004	Shared Memberships	3111
1996+	Paradata	3146
2008, 2012	Form Y Questions not listed elsewhere	3158
1972+	Sampling and Weights	3162

Appendices

A.	Sampling Design and Weighting	3171
B.	Field Work and Interviewer Specifications	3190
C.	General Coding Instructions	3239
D.	Recodes	3241
E.	Age and Cohort Distributions	3284
F.	Occupational Classification Distributions	3286
G.	Prestige Scores Distributions	3320
H.	Industrial Classifications Distributions	3327
I.	International Standard of Classification of Occupation	3345
J.	DOT Variables	3405
K.	Protestant Denominations Distributions	3406
L.	Hours Worked Distributions	3411
M.	Abortion and ERA Distributions	3413
N.	Changes in Question Wording, Response Categories, and Format	3417
O.	Previous Usage	3435
P.	Experimental Form	3437
Q.	Rotation Design	3451
R.	Crossnational and Topical Modules	3454
S.	Supplemental and Related Data	3488
T.	General Social Survey Papers	3491
U.	Variables by Year	3513
V.	Subject Index to Questions	3678

INTRODUCTION

The National Data Program for the Social Sciences is designed as a data diffusion project and a program of social indicator research. The data come from the General Social Surveys, interviews administered to NORC national samples using a standard questionnaire. Toward the major goal of functioning as a social indicator program, items which have appeared on previous national surveys starting in 1937 have been replicated here. The search for trend items led us to published reports from Gallup, Harris, the Detroit Area Study, SRC (Michigan) studies, NORC files, and Federal Commissions such as those on Violence and Pornography.

By retaining the exact wording, we hope to facilitate time trend studies as well as replications of earlier findings. For the base line items in the initial 1972 survey, some 105 sociologists and social scientists reviewed drafts of the questionnaire, suggested revisions and additions, and expressed their question preference by vote. Their serious assistance was extremely helpful in putting together a final version of the questionnaire which would represent the varied interests of social scientists. Topic and question selection continues to be monitored by leading social scientists who serve as a Board of Overseers: Deborah Carr, Jeremy Freese, Bridget Goosby, Darrick Hamilton, Vincent Hutchings, Frauke Kreuter, Taeku Lee, Jason Schnittker, Judith Seltzer, Florenica Torche, Rob Warren, and Rebeca Wong.

The items appearing on the surveys are one of three types: Permanent questions that occur on each survey, rotating questions that appear on two out of every three surveys (1973, 1974, and 1976, or 1973, 1975, and 1976), and a few occasional questions such as split ballot experiments that occur in a single survey. Starting in 1988, items were no longer rotated across years but appeared on two-thirds of the cases every year. This design is discussed in Appendix Q. A detailed layout of the appearance of questions can be found right before the index to this codebook.

A second objective is the prompt distribution of fresh, interesting, and high-quality data to a variety of users who are not affiliated with large research centers. The initial survey, 1972, was supported by grants from the Russell Sage Foundation and the National Science Foundation. NSF has provided support for the 1973 through 1978, 1980, and 1982 through 2018 surveys. We welcome your participation in this program. While it is not necessary to request permission from NORC before publishing analyses of these data, we do ask that NORC be cited as the source of your data. We also request that copies of reports which utilize the data be sent to the General Social Survey, NORC, 1155 East 60th Street, Chicago, IL 60637.

DATA

The General Social Surveys have been conducted in 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, and 2018. There are a total of 64,814 completed interviews (1,613 in 1972, 1,504 in 1973, 1,484 in 1974, 1,490 in 1975, 1,499 in 1976, 1,530 in 1977, 1,532 in 1978, 1,468 in 1980, 1,506 in 1982, 354 in 1982 black oversample, 1,599 in 1983, 1,473 in 1984, 1,534 in 1985, 1,470 in 1986, 1466 in 1987, 353 in 1987 black oversample, 1481 in 1988, 1,537 in 1989, 1372 in 1990, 1,517 in 1991, 1,606 in 1993, 2,904 in 1996, 2,832 in 1998, 2,817 in 2000, 2,765 in 2002, 2,812 in 2004, 4510 in 2006, 2023 in 2008, 2,044 in 2010, 1,974 in 2012, 2,538 in 2014, 2,867 in 2016, and 2,348 in 2018). The median length of the interview has been about one and a half hours. Each survey from 1972 to 2004 was an independently drawn sample of English-speaking persons 18 years of age or over, living in non-institutional arrangements within the United States. Starting in 2006 Spanish-speakers were added to the target population. Block quota sampling was used in 1972, 1973, and 1974 surveys and for half of the 1975 and 1976 surveys. Full probability sampling was employed in half of the 1975 and 1976 surveys and the 1977, 1978, 1980, 1982-1991, 1993-1998, 2000, 2002, 2004, 2006, 2008, 2010, 2012, 2014, 2016, and 2018 surveys. Also, the 2004, 2006, 2008, 2010, 2012, 2014, 2016, and 2018 surveys had sub-sampled non-respondents (see Appendix A for a detailed description of the sample design).

The data from the interviews were processed according to standard NORC procedures.

Cleaning procedures--utilizing a combination of the coding specifications and the interviewer instructions--were used to check for inconsistent or illegitimate codes (see Appendix B for interviewer instructions and Appendix C for general coding instructions). Some variables--age, occupation, and occupational prestige--are coded so that the first digit of the two- or three-digit codes may be used separately.

This cumulative data set merges all 32 surveys into a single file with each year or survey acting as a subfile. This greatly simplifies the use of the General Social Surveys for both trend analysis and pooling. In addition, this cumulative data set contains newly created variables (e.g. a poverty line code). Finally, the cumulative file contains certain items never before available.

To facilitate the use of the codebook, several terms must be explained. The abbreviation "R," which appears throughout the text and appendices, stands for "respondent." The format which we have used in the text of the codebook is as follows:

110. Do you think the use of marijuana should be made legal or not?

[VAR: GRASS]

RESPONSE	PUNCH	YEAR														
		1972-82	1982B	1983-87	1987B	1988-91	1993-98	2000-04	2006	2008	2010	2012	2014	2016	2018	ALL
Should	1	1803	0	1156	63	668	1708	1195	672	496	603	586	870	1126	938	11884
Should not	2	5413	0	4654	277	3124	4815	2238	1156	751	656	648	704	717	509	25662
Don't know	8	242	0	181	12	136	361	245	161	116	117	100	111	110	112	2004
No answer	9	35	0	17	1	20	20	6	0	2	1	4	2	2	4	114
Not applicable	0	6133	354	1534	0	1959	3430	4710	2521	658	667	636	851	912	785	25150

REMARKS: GRASS was asked on both Form 1 and Form 2 at different points in the interview in 1978. See Appendix B for locations in the questionnaire.

The format includes the question exactly as it appears in the questionnaire. For those few questions that were recoded, the symbol [RECODE] appears immediately after the question. For the original question wording, the user must turn to Appendix D: Recodes. Question numbering as it appeared on the actual questionnaire is given in Appendix B.

"[VAR: GRASS]" refers to the variable name. A mnemonic was assigned to each question to promote standardization in the use of General Social Survey variable names and also to meet the eight character limitation imposed by some computer software systems (e.g., SPSS).

Under the heading "RESPONSE," all possible answers to the questions are listed. The questionnaire contains three alternate forms of response as follows: (1) the answers were read to the respondent (if they were included in the question); (2) answers were presented to the respondent on a card (indicated by interviewer instructions); or (3) answers were marked by the interviewer to best correspond to the answer of the respondent (also indicated by interviewer instructions).

The term "PUNCH" represents the code or numerical value which was assigned to each response. These are the numbers that the user will find punched in the columns. The frequency of occurrence of each of the punch values appears in the next four columns. The combined marginals across the surveys are in the last column headed "ALL."

In most cases, the marginal distributions for all punches are given in the text. For a small number of variables--the two-or-more-column variables--frequencies or marginal distributions appear in the appendices. Responses are mutually exclusive (i.e., only one code can appear for each respondent for each question).

The first column under "YEAR," 1972-1982, gives the combined totals for the 1972-1982 cross-sections. In the second column, 1982B, the counts for the 1982 black oversample appear. Blacks who were part of the regular 1982 sample are not part of these figures. The third column, 1983-1987, gives the combined totals for 1983-1987. The fourth column, 1987B, contains the counts for the 1987 black oversample. The fifth column, 1988-91, gives the combined totals for 1988-1991. The sixth column, 1993-98, gives the combined totals for 1993-98. The seventh column, 2000-04, gives the combined totals for 2000-2004. The eighth column, 2006, contains the counts for the 2006 survey. The ninth column, 2008, contains the counts for the 2008 survey. The tenth column, 2010, contains the counts for the 2010 survey. The eleventh column, 2012, contains the counts for the 2012 survey. The twelfth column, 2014, contains the counts for the 2014 survey. The thirteenth column, 2016, contains the counts for the 2016 survey. The fourteenth column, 2018, contains the counts for the 2018 survey. Lastly, the fifteenth column, ALL, contains the total for the preceding fifteen columns. For a discussion of the use of the black oversample see Appendix A. For the individual yearly totals for 1972-1982 consult the *General Social Surveys, 1972-1982: Cumulative Codebook*; for 1983-1987 consult the *General Social Surveys, 1972-1987: Cumulative Codebook*; and for 1988-1991, consult *General Social Surveys, 1972-1991: Cumulative Codebook*. To determine in what years or on which ballots a variable appeared in see Appendix U.

CHANGES IN DESIGN

With NSF's renewal of the GSS for 1993-1997, major changes in design occurred. The 1993 GSS was the last survey conducted under the old design. In 1994 two major innovations were introduced to the GSS.

First, the traditional core is substantially reduced to allow for the creation of modules (i.e. blocks devoted to some combination of small- to medium-sized supplements). The module space gives us greater flexibility to incorporate innovations and to include important items proposed by the social science community.

Second, a new biennial, split-sample design is used. The sample consists of two parallel sub-samples of approximately 1,500 cases each. The two sub-samples both contain the identical core. The A sample also contains a standard, topical module, the mini-modules, and an ISSP module (on women, work, and the family). The B sample has a second topical module, mini-modules, and an ISSP module (on the environment). In effect, one can think of the A sample as representing a traditional GSS for 1994 and the B sample as representing a traditional GSS for 1995. Rather than being fielded separately in two different years they are fielded together.

While we generally field separate topical and ISSP modules on the A and B samples, we have the option of including some items on both samples if a larger sample size is needed.

In 1996 and in subsequent even numbered years the same design described for 1994 has been repeated. In addition, in 1994 only there is a transitional design to calibrate any impact of deletions from the core. On Sample A, the old core was administered to respondents receiving Version 1 (X) and the new reduced core was given on Version 2 (Y). See Appendix U for further information about specific items.

In 2002 the GSS underwent a change in survey mode. Between 1972 & 2000, the survey was administered in a paper-and-pencil format, but starting in 2002, the GSS was conducted by computer-assisted personal interviewing (CAPI). See Appendix B for more information about field mode.

In 2004 there were two major changes. First, a new sample frame was adopted based on the 2000 US Census. This is explained in Appendix A. Second, a non-respondent, sub-sampling design was utilized. This is explained in Appendix A. The adoption of the non-respondent, sub-sampling design means that a weight variable must be employed (see Appendix A). These two features will continue for the foreseeable future.

In 2002 and 2004 the sample size for rotating items in the GSS core was reduced by about 50% due to lower funding levels.

In 2006 there were two additional changes. First, Spanish interviews were conducted so the target population is now adults living in households in the US who speak English or Spanish. Second, a third sample was added to accommodate more supplements. This led to a total sample size in 2006 of 4510.

In 2008 the GSS started a transition from a replicating cross-sectional design to a design that uses rotating panels. There were two components: a new 2008 cross-section with 2,023 cases and the first reinterviews with 1,536 respondents from the 2006 GSS. In 2010 the new design was fully implemented. There was a new cross-section of 2,044 cases, the first reinterviews of the 2008 GSS respondents, and the second and final reinterviews of the 2006 GSS respondents. This three-wave, rolling panel design was also used in 2012. Each GSS in 2006-2010 thus started a new 4-year/3-wave panel, which was completed four years later (2006-08-10, 2008-10-12, 2010-12-14). However, in 2012 no new three wave panel was started.

Abbreviations:

The following abbreviations are used throughout the text and appendices:

AIPO	American Institute of Public Opinion (Gallup Poll)
Col(s).	Column(s)
IISR	International Institute for Social Research
ISSP	International Social Survey Program
GO	Gallup Organization
N	Number
IAP	Inapplicable
NORC/SRS	National Opinion Research Center/Survey Research Service
n.e.c.	Not elsewhere classified
ORCO	Opinion Research Corporation
POS	Public Opinion Survey (Gallup)
PSU	Primary Sampling Unit
Q(s).	Question(s)
R	Respondent, except in Appendix C: General Coding Instructions, where R stands for blank.
Roper	Roper Public Opinion Research Center, University of Connecticut
ICPSR	Inter-University Consortium for Political and Social Research, University of Michigan
SRC	Survey Research Center, University of Michigan
Var.	Variable
Vol.	Volunteered
ZUMA	Zentrum fuer Umfragen, Methoden, and Analysen, Germany

Data Identification Numbers:

Identification numbers and locations are as follows:

N = 64,814

- Year appears in col. 1-4
- Respondent identification number in cols. 5-8