Introduction to National Survey of Family Growth (NSFG)

Hsueh-Sheng Wu

CFDR Workshop Series

October 21, 2019

BGSU
Center for Family and Demographic Research

Outline

- Aims of NSFG
- NSFG study design
- 2015-2017 NSFG data files
- Contents in 2015-2017 NSFG public data
- Examples of faculty's work using NSFG
- Navigate NSFG website
  - Questionnaire
  - Interactive codebook
  - Download data
- Sample Stata codes for:
  - Reading NSFG data file into statistical software
  - Linking a female respondent file with a pregnancy file
  - Pooling men and women data from the same cycle
  - Using weight variables in analyses
- Conclusions

Aims of NSFG

- Address Sec 306 of the Public Health Service Act: "NCHS shall collect statistics on… family formation, growth, and dissolution."
  - Provide reliable national data on marriage, divorce, sexual activity, contraception, pregnancy, sterilization, and infertility
  - Describe attitudes of sex, marriage, cohabitation, and parenthood
  - Monitor health risks, sexually transmitted diseases (including HIV), and the health of infants, women, and men
  - Policies can be made to meet the needs of families
  - Service agencies can design health and service programs

NSFG Study Design

Name	Year	Target population: Non-institutionalized people in the U.S. people meeting the criteria			Survey Design (repeated cross-sectional survey)	N
		Gender	Age Range	marital or parental status		
Cycle 1	1973	women	15-44	1. ever married or 2. have offspring in the household	Interviews completed in a year	9,797
Cycle 2	1976	women	15-44	1. ever married or 2. have offspring in the household	Interviews completed in a year	8,611
Cycle 3	1982	women	15-44		Interviews completed in a year	7,969
Cycle 4	1988	women	15-44		Interviews completed in a year	8,450
Cycle 5	1995	women	15-44		Interviews completed in a year	10,847
Cycle 6	2002	women men	15-44		Interviews completed in a year	12,571 W=7,643 M=4,928
2006-2010 NSFG	2006-2010	women men	15-44		Interviews spread over several years	22,682 W=12,279 M=10,403
2011-2013 NSFG	2011-2013	women men	15-44		Interviews spread over several years	10,416 W=5,601 M=5,815
2013-2015 NSFG	2013-2015	women men	15-44		Interviews spread over several years	10,205 W=5,699 M=4,506
2015-2017 NSFG	2015-2017	women men	15-49		Interviews spread over several years	10,094 W=5,554 M=4,540

BGSU Center for Family and Demographic Research

NSFG Study Design (Cont.)

Table 1. Study Designs of NSFG (Continued)

Name	Year	Sampling Method	Over-Sampled groups	Data Collection Methods
Cycle 1	1973	multistage probability sampling	Black women	Personal interview
Cycle 2	1976	multistage probability sampling	Black women	face-to-face interview
Cycle 3	1982	multistage area probability sampling	Black women & Teens	face-to-face interview
Cycle 4	1988	Using Households used in National Health Interview Survey	Black women	face-to-face interview
Cycle 5	1995	Using Households used in National Health Interview Survey	Black and hispanic women	Computer-assisted personal interview (CAPI)
Cycle 6	2002	multistage area probability sampling	Blacks, Hispanics, & Teens	Interviewer administered CAPI self-administered ACAPI
2006-2010 NSFG	2006-2010	multistage area probability sampling	Blacks, Hispanics, & Teens	Interviewer administered CAPI self-administered ACAPI
2011-2013 NSFG	2011-2013	multistage area probability sampling	Blacks, Hispanics, & Teens	Interviewer administered CAPI self-administered ACAPI
2013-2015 NSFG	2013-2015	multistage area probability sampling	Blacks, Hispanics, & Teens	Interviewer administered CAPI self-administered ACAPI
2015-2017 NSFG	2015-2017	multistage area probability sampling	Blacks, Hispanics, & Teens	Interviewer administered CAPI self-administered ACAPI

2015-2017 NSFG Data Files

2015-2017 NSFG consists of public data and restricted data

Public data

- Public data can be directly downloaded form the NSFG website
- Data files include (1) Female Respondent Data File, (2) Female Pregnancy Data File, (3) Male Respondent Data File, and (4) 4- and 6-year weight variables for the periods between 2011-2015, 2013-2017, and 2011-2017, respectively.
- Values of some variables are collapsed, modified, or suppressed to protect the confidentiality of the respondents. For example, the century month data values for marriages, divorces, pregnancies, cohabitations, educational degrees, military services, and selected health services were suppressed.
- New variables were constructed (e.g., inter-pregnancy interval variables) and added to the data files, so that researchers can still address important research questions without the exact information on the century months of the pregnancies.

Restricted data

- Restricted data can only be accessed from research data centers (RDC) of National Center for Health Statistics (NCHS). The locations of these centers can be found at https://www.cdc.gov/rdc/leftbrch/locaterdc.htm and https://www.census.gov/about/adrm/fsrdc/locations.html .
- With restricted data, researchers have the access to some important variables, such as the century months of the pregnancies and the region of residence for the respondents.
- To access restricted data, researchers need to follow the application process that NCHS specifies in https://www.cdc.gov/rdc/b3prosal/PP300.htm.

Contents in 2015-2017 NSFG Public Data

Table 2. The Topics covered in the Female and Male Respondents in the 2015-2017 NSFG

Section	Female Respondent Data	Male Respondent Data
	Respondent ID and Selected Screener Items	Respondent ID and Selected Screener Items
A	Calendar Instructions; Demographic Characteristics; Household Roster; Childhood Background	Demographic Characteristics; Household Roster; Childhood Background; Marital/Cohabiting Status
B	Pregnancy & Birth History; Adoption & Nonbiological Children	Sex Communication, Ever Sex, Number of Sexual Partners, Enumeration and Relationship With Up To 3 Recent (Or Last) Sexual Partner(s)
C	Marital and Relationship History	Current Wife or Cohabiting Partner
D	Sterilizing Operations and Impaired Fecundity	Recent (Or Last) Sexual Partner(s) and First Sexual
E	Contraceptive History and Pregnancy Wantedness	Former Wives and First Cohabiting Partner
F	Family Planning and Medical Services	Other Biological Children, Other Adopted Children, Other Pregnancies
G	Birth Desires and Intentions	Fathering
H	Infertility Services and Reproductive Health	Desires and Intentions for Future Children
I	Insurance; Residence and place of birth; Religion; Past and current work (R and current H/P); Child Care;	Health Conditions and Health Services
J	Audio CASI	Residence and place of birth; Religion; Military service; Past and current work (R and current wife/partner);
K		Audio CASI
	Recode Variables and Imputation Flags	Recode Variables and Imputation Flags
	Weights and Other Variables	Weights and Other Variables
	Date of interview, key reference dates, and fieldwork variables	Date of interview, key reference dates, and fieldwork variables

Contents in 2015-2017 NSFG Public Data (Cont.)

Table 3. The file index of the Pregnancy Data in the 2015-2017 NSFG
Pregnancy Data
Respondent ID and Pregnancy Order
Section B raw variables
Section E raw variables
Section B pregnancy-based recodes
Section E pregnancy-based recodes
Selected respondent file variables
Pregnancy recode imputation flags
Respondent recode imputation flags included on pregnancy file
Date of interview, key reference dates, and fieldwork variables

BGSU Center for Family and Demographic Research

Examples of Faculty's Work Using NSFG

Guzzo, K. B. (2017). Is stepfamily status associated with cohabiting and married women's fertility behaviors? *Demography, 54*, 45-70.

Joyner, K, Peters, H. E., Hynes, K., Sikora, A., Taber, J. R., & Rendall, M. S. (2012). The quality of male fertility data in major U.S. surveys. *Demography*, *49*, 101-122.

Lamidi, E. O., Manning, W. D., Brown, S. L. (2019). Change in the stability of first premarital cohabitation among women in the United States, 1983-2013, *Demography*, *56*, 427-450.

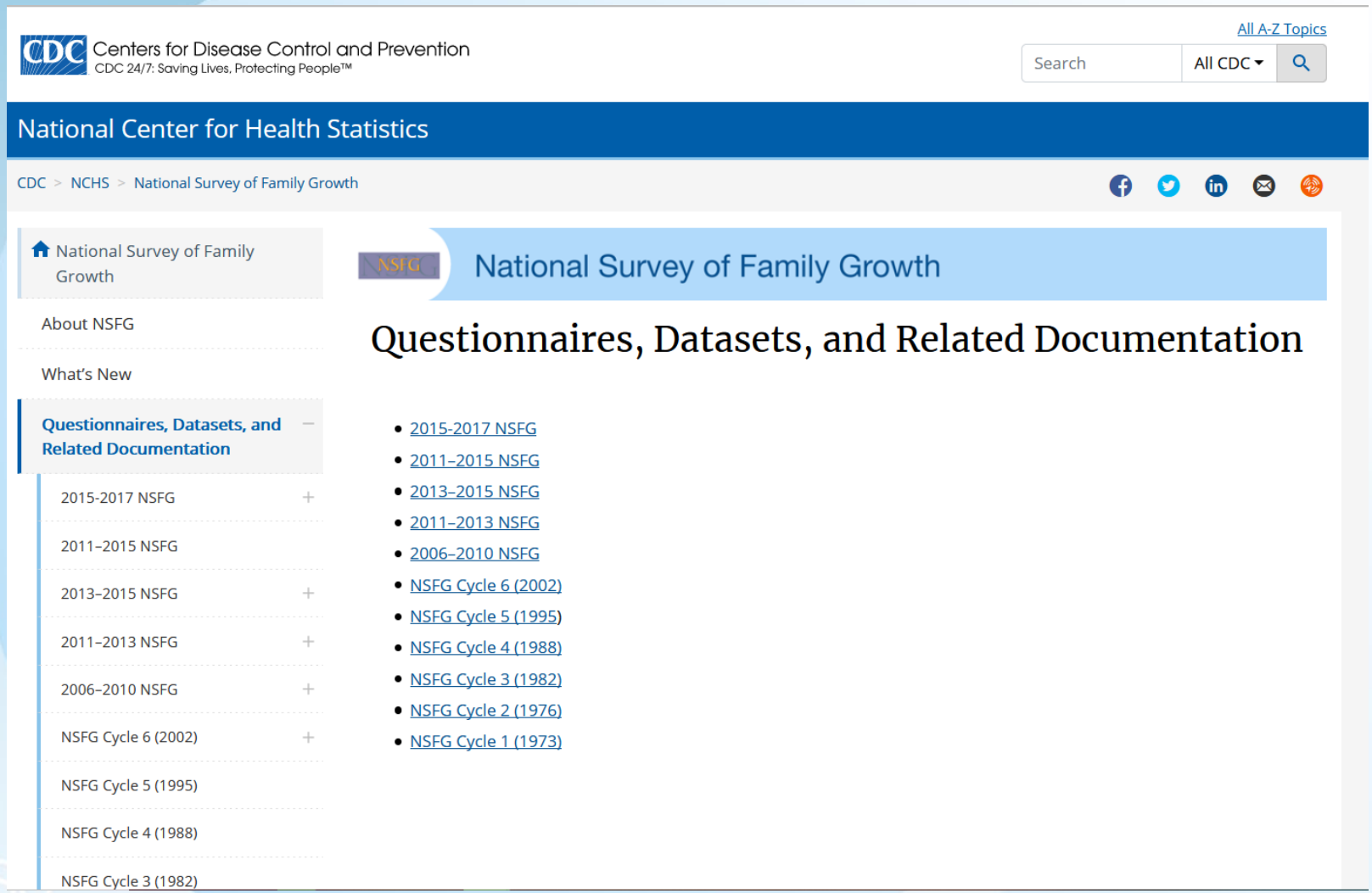
Manning, W. D., & Cohen, J. A. (2012). Premarital cohabitation and marital dissolution: An examination of recent marriages. *Journal of Marriage and Family*, *74*, 377–387.

Manning W. D., & Brown, S. L. (2014). Two decades of stability and change in age at first union formation. *Journal of Marriage and Family, 76*, 247-260.

Stykes, J. B., Manning, W. D., & Brown, S. L. (2013). Nonresident fathers and formal child support: Evidence from the CPS, NSFG, and SIPP. *Demographic Research*, *29*, 1299-1330

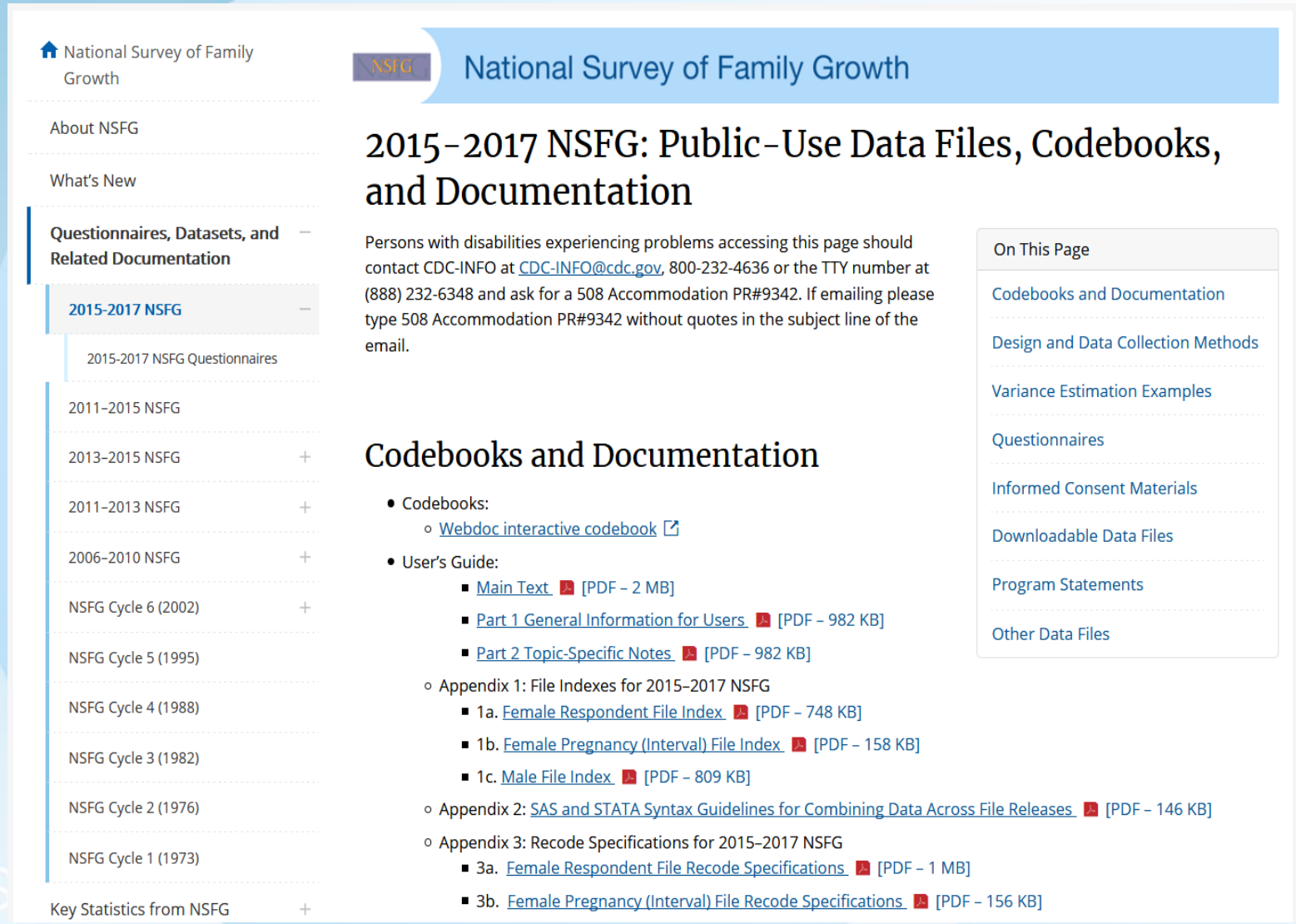
Navigate NSFG Website

https://www.cdc.gov/nchs/nsfg/nsfg_questionnaires.htm



Navigate NSFG Website (Cont.)

https://www.cdc.gov/nchs/nsfg/nsfg_2015_2017_puf.htm



National Survey of Family Growth

About NSFG

What's New

Questionnaires, Datasets, and Related Documentation

2015-2017 NSFG

2015-2017 NSFG Questionnaires

2011–2015 NSFG

2013–2015 NSFG

2011–2013 NSFG

2006–2010 NSFG

NSFG Cycle 6 (2002)

NSFG Cycle 5 (1995)

NSFG Cycle 4 (1988)

NSFG Cycle 3 (1982)

NSFG Cycle 2 (1976)

NSFG Cycle 1 (1973)

Key Statistics from NSFG

NSFG National Survey of Family Growth

2015–2017 NSFG: Public-Use Data Files, Codebooks, and Documentation

Persons with disabilities experiencing problems accessing this page should contact CDC-INFO at CDC-INFO@cdc.gov, 800-232-4636 or the TTY number at (888) 232-6348 and ask for a 508 Accommodation PR#9342. If emailing please type 508 Accommodation PR#9342 without quotes in the subject line of the email.

On This Page

- Codebooks and Documentation
- Design and Data Collection Methods
- Variance Estimation Examples
- Questionnaires
- Informed Consent Materials
- Downloadable Data Files
- Program Statements
- Other Data Files

Codebooks and Documentation

- Codebooks:
  - Webdoc interactive codebook
- User's Guide:
  - Main Text [PDF – 2 MB]
  - Part 1 General Information for Users [PDF – 982 KB]
  - Part 2 Topic-Specific Notes [PDF – 982 KB]
  - Appendix 1: File Indexes for 2015–2017 NSFG
    - 1a. Female Respondent File Index [PDF – 748 KB]
    - 1b. Female Pregnancy (Interval) File Index [PDF – 158 KB]
    - 1c. Male File Index [PDF – 809 KB]
  - Appendix 2: SAS and STATA Syntax Guidelines for Combining Data Across File Releases [PDF – 146 KB]
  - Appendix 3: Recode Specifications for 2015–2017 NSFG
    - 3a. Female Respondent File Recode Specifications [PDF – 1 MB]
    - 3b. Female Pregnancy (Interval) File Recode Specifications [PDF – 156 KB]

2015-2017 NSFG Website

NSFG website provides important information on each NSFG survey

Take 2015-2015 NSFG as an example, the website provides:

- Codebooks and documentation
- Design and data collection methods
- Variance estimation examples
- Questionnaires
- Informed consent materials
- The location of downloadable data files
- Program statements for SAS, Stata, and/or SPSS.
- The availability of other data files

Questionnaire

NSFG questionnaire has two formats:

1. CAPI-Lite format provides abridged representations of the question wording variants and short descriptions of skip patterns through the interview

TIMESMAR

CA-1. (Including your present marriage,) how many times have you been married?

Number _____

{ CA-2, CA-2b, & CA-2c ARE INTENDED TO OBTAIN NAMES OR INITIALS OF HUSBANDS,
{ ONLY FOR PURPOSES OF LOOPING THROUGH CA SERIES.
{ IF R HAS ANSWERED DK/RF FOR # OF TIMES MARRIED, SHE IS LOOPED ONLY ONCE
{ THROUGH CA SERIES.

Questionnaire (Cont.)

2. CAPI Reference Questionnaire (CRQ) provides comprehensive specifications of each question that NSFG staff provided to the computer programmers to created the CAPI instrument

TIMESMAR

CA-1. IF FMARIT = 1 or 4 (married or separated), ASK:
Including your present marriage, how many times have you been married?

ELSE IF FMARIT = 2 or 3 (widowed or divorced), ASK:
How many times have you been married?

- *ENTER number*

[HELP AVAILABLE]

Number _____

UNDERLYING RANGE: 1 to 95
(Looping for only 6 husbands)

FLOW CHECK C-4: IF MARSTAT = 1 (married) AND 1 LE TIMESMAR LE 6, ASK CA-2 HUSBNAMEX FOR EACH FORMER HUSBAND IF ANY (up to TIMESMAR-1) AND THEN ASK CA-2b HSBVERIF TO VERIFY NAME OF CURRENT HUSBAND.

ELSE IF 2 LE CA-1 TIMESMAR LE 6 OR (CA-1 TIMESMAR = 1 AND AD-7b MARSTAT NE 1), ASK CA-2 HUSBNAMEX FOR EACH OF THE [TIMESMAR] FORMER (or currently separated) HUSBANDS.

ELSE IF MARSTAT = 1 (married) AND 1 LE TIMESMAR LE 6 AND CHPNAME = "YOUR HUSBAND", ASK CA-2 HUSBNAMEX FOR EACH FORMER HUSBAND IF ANY (up to TIMESMAR-1) AND THEN GO TO CA-2c CHVERIFY TO FIND OUT NAME OF CURRENT HUSBAND.

ELSE IF TIMESMAR = DK OR RF, GO TO CA-2c CHVERIFY TO FIND OUT NAME OF CURRENT OR MOST RECENT HUSBAND (BASED ON MARSTAT VALUE).

ELSE IF TIMESMAR GT 6 AND MARSTAT = 1 (married), ASK HUSBNAMEX FOR HER FIRST 5 FORMER HUSBANDS AND THEN ASK CA-2c CHVERIFY OR CA-2b HSBVERIF (as appropriate based on CHPNAME value) ABOUT R's CURRENT HUSBAND. *(For currently married women with more than 6 marriages, we will only ask about her 1st 5 husbands and her current husband.)*

ELSE IF TIMESMAR GT 6 AND MARSTAT NE 1 (married), ASK

HUSBNAMEX FOR HER FIRST 6 FORMER HUSBANDS. *(For currently unmarried women with more than 6 marriages, we will only ask about her 1st 6 marriages.)*

Interactive Codebook

BGSU Center for Family and Demographic Research

Interactive Codebook (Cont.)

Search: in All Files

Search variable name only Search (search tips)

NSFG 2015-2017 :: Female Respondent File Codebook :: SECTION C: Marital and Relationship History :: (CA) Number of marriages

TIMESMAR (413-413)

Variable Type : raw

CA-1 : (Including your present marriage, how / How) many times have you been married?

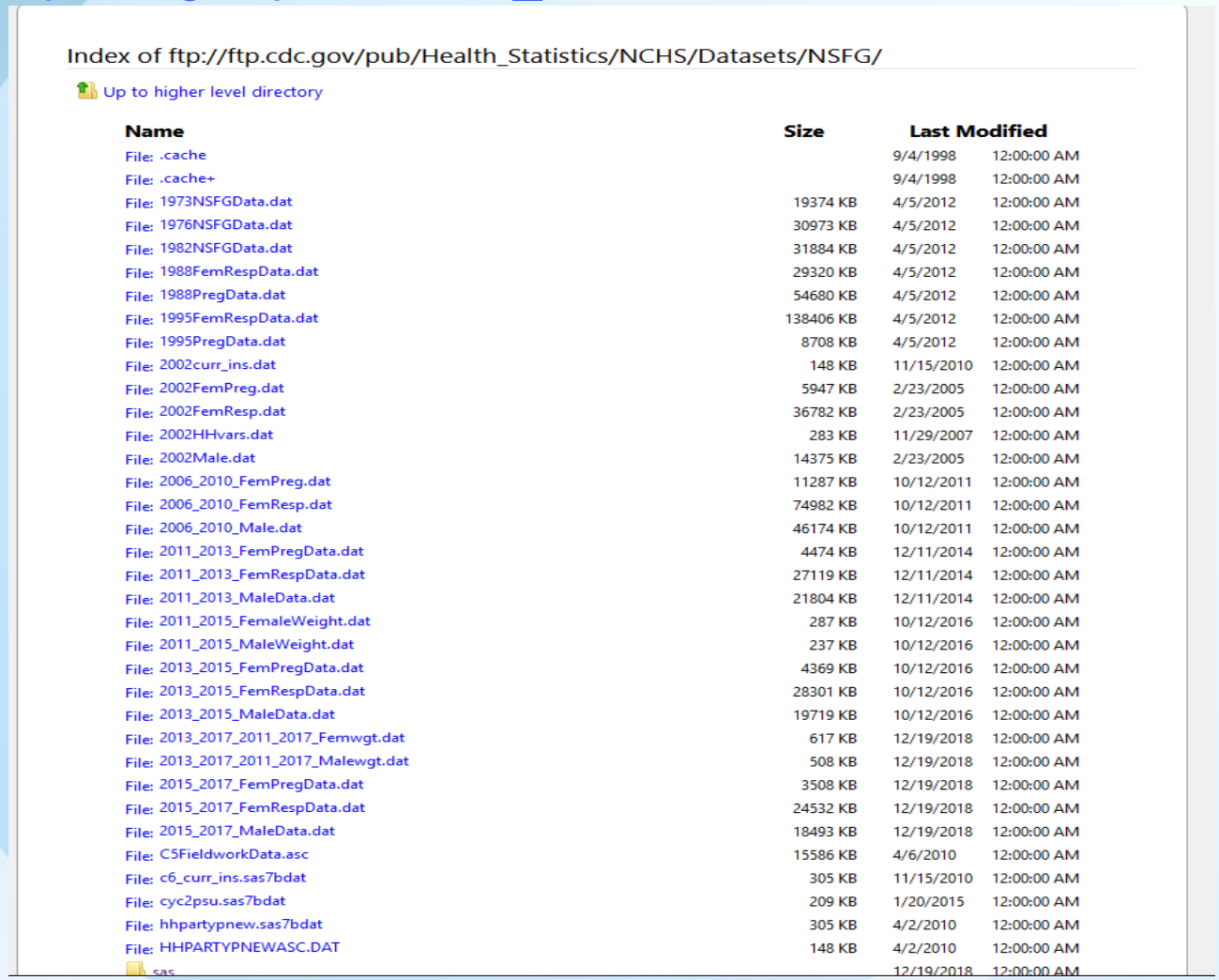
value	label	Total
.	INAPPLICABLE	2972
1	1 TIME	2156
2	2 TIMES	357
3	3 TIMES	55
4	4 TIMES	10
5	5 TIMES	2
6	6 TIMES	2
	Total	**5554**

Universe : Applicable if R has ever been married (fmarit = 1, 2, 3, or 4)

Notes : use recode FMARNO

Download Data

NSFG data can be downloaded from the NSFG FTP site:
ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Datasets/NSFG



Index of ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Datasets/NSFG/

Up to higher level directory

Name	Size	Last Modified
File: .cache		9/4/1998 12:00:00 AM
File: .cache+		9/4/1998 12:00:00 AM
File: 1973NSFGData.dat	19374 KB	4/5/2012 12:00:00 AM
File: 1976NSFGData.dat	30973 KB	4/5/2012 12:00:00 AM
File: 1982NSFGData.dat	31884 KB	4/5/2012 12:00:00 AM
File: 1988FemRespData.dat	29320 KB	4/5/2012 12:00:00 AM
File: 1988PregData.dat	54680 KB	4/5/2012 12:00:00 AM
File: 1995FemRespData.dat	138406 KB	4/5/2012 12:00:00 AM
File: 1995PregData.dat	8708 KB	4/5/2012 12:00:00 AM
File: 2002curr_ins.dat	148 KB	11/15/2010 12:00:00 AM
File: 2002FemPreg.dat	5947 KB	2/23/2005 12:00:00 AM
File: 2002FemResp.dat	36782 KB	2/23/2005 12:00:00 AM
File: 2002HHvars.dat	283 KB	11/29/2007 12:00:00 AM
File: 2002Male.dat	14375 KB	2/23/2005 12:00:00 AM
File: 2006_2010_FemPreg.dat	11287 KB	10/12/2011 12:00:00 AM
File: 2006_2010_FemResp.dat	74982 KB	10/12/2011 12:00:00 AM
File: 2006_2010_Male.dat	46174 KB	10/12/2011 12:00:00 AM
File: 2011_2013_FemPregData.dat	4474 KB	12/11/2014 12:00:00 AM
File: 2011_2013_FemRespData.dat	27119 KB	12/11/2014 12:00:00 AM
File: 2011_2013_MaleData.dat	21804 KB	12/11/2014 12:00:00 AM
File: 2011_2015_FemaleWeight.dat	287 KB	10/12/2016 12:00:00 AM
File: 2011_2015_MaleWeight.dat	237 KB	10/12/2016 12:00:00 AM
File: 2013_2015_FemPregData.dat	4369 KB	10/12/2016 12:00:00 AM
File: 2013_2015_FemRespData.dat	28301 KB	10/12/2016 12:00:00 AM
File: 2013_2015_MaleData.dat	19719 KB	10/12/2016 12:00:00 AM
File: 2013_2017_2011_2017_Femwgt.dat	617 KB	12/19/2018 12:00:00 AM
File: 2013_2017_2011_2017_Malewgt.dat	508 KB	12/19/2018 12:00:00 AM
File: 2015_2017_FemPregData.dat	3508 KB	12/19/2018 12:00:00 AM
File: 2015_2017_FemRespData.dat	24532 KB	12/19/2018 12:00:00 AM
File: 2015_2017_MaleData.dat	18493 KB	12/19/2018 12:00:00 AM
File: C5FieldworkData.asc	15586 KB	4/6/2010 12:00:00 AM
File: c6_curr_ins.sas7bdat	305 KB	11/15/2010 12:00:00 AM
File: cyc2psu.sas7bdat	209 KB	1/20/2015 12:00:00 AM
File: hhpartypnew.sas7bdat	305 KB	4/2/2010 12:00:00 AM
File: HHPARTYPNEWASC.DAT	148 KB	4/2/2010 12:00:00 AM
sas		12/19/2018 12:00:00 AM

BGSU

Center for Family and Demographic Research

Weighting in NSFG Analyses

NSFG analysis needs to be weighted to adjust for:

- The complex survey design (not a simple random sampling)
- Oversample sub-populations
- Screener and interview non-response
- Estimated characteristics of population from the Census Bureau

The name of the weight variables may change across NSFG cycles.

Table 4. Weight Variables for NSFG from 2002-2017

Design variable	2002	2006–2010	2011–2013	2013–2015	2015–2017
Stratum variable	SEST	SEST	SEST	SEST	SEST
Four Cluster/Panel Variable	SECU_R (fem resp) SECU_P (fem preg) SECU (male resp)	SECU	SECU	SECU	SECU
Final post-stratified, fully adjusted case weight	FINALWGT	WGTQ1Q16	WGT2011_2013	WGT2013_2015	WGT2015_2017

Note:

1. There is no weight variables for single year estimate because of insufficeint statistical power
2. The 4-and 6-year weight variables, WGT2011-2015, WGT2013_2017 and WGT2011_2017, should e used when researchers combined NSFG data across the different spans of years.

BGSU Center for Family and Demographic Research

Reading NSFG 2015-2017 Data

Stata : Modify the following lines in the Stata command file from the NSFG website.

```
clear
cd d:\temp\nsfg
local raw_data "2015_2017_FemRespData.dat"
local dict "2015_2017_FemRespSetup.dct"
local outfile " 2015_2017_FemRespData.dta“
.
.
.
save `outfile', replace
```

Linking A Female Respondent File with A Pregnancy File

```
**********************
* Sort the female respondent data by ID
**********************
use 2015_2017_FemRespSetup.dta, clear
sort caseid
save, replace

*****************************************************
* Sort the pregnancy data by ID and pregnancy order
*****************************************************
use 2015_2017_FemPregData.dta, clear
sort caseid pregordr
save, replace

*******************************
* Merge and save data
*******************************
use 2011_2013_FemRespSetup.dta, clear
merge 1:m caseid using 2011_2013_FemPregData.dta

save 2011_2013_fem_pregnancy.dta, replace
```

BGSU Center for Family and Demographic Research

Pooling Men and Women Data from the Same Time Period

```
**************************************
* Select variables from the male file
**************************************
use 2015_2017_MaleSetup.dta, clear
keep caseid rscrninf wgt2015_2017 secu sest
save temp1.dta, replace

**************************************
* Select variables from the female file
**************************************
use 2015_2017_FemRespSetup.dta, clear
keep caseid rscrninf wgt2015_2017 secu sest
save temp2.dta, replace

********************************
* Pool two data sets together
********************************
use temp1.dta, clear
append using temp2.dta

save 2015_2017_male_female.dta, replace
```

Using Weight Variables in the Analysis

NSFG Website provides three examples of incorporating weight variables into the analysis (https://www.cdc.gov/nchs/nsfg/nsfg_2015_2017_puf.htm)

Example 1: Percentage of Women Ages 15-49 Currently Using the Oral Contraceptive Pill, by Age

```
use “EX1.DTA”
svyset [pweight=WGT2015_2017], strata(SEST) psu(SECU)

generate agerx=1 if AGER <=19
replace agerx=2 if AGER >=20 & AGER <=24
replace agerx=3 if AGER >=25 & AGER <=29
replace agerx=4 if AGER >=30 & AGER <=34
replace agerx=5 if AGER >=35 & AGER <=39
replace agerx=6 if AGER >=40

generate cpill=2
replace cpill=1 if CONSTAT1==6

svy: tab agerx cpill, row se percent
```

STATA Output

```
. svy: tab agerx cpill, row se percent
(running tabulate on estimation sample)

Number of strata   =        18                Number of obs      =       5,554
Number of PSUs     =        72                Population size    =  72,218,086
                                              Design df          =          54
```

agerx	cpill: yes	cpill: no	Total
15-19	16.64	83.36	100
	(2.442)	(2.442)	
20-24	22.5	77.5	100
	(2.476)	(2.476)	
25-29	16.68	83.32	100
	(1.738)	(1.738)	
30-34	14.05	85.95	100
	(2.004)	(2.004)	
35-39	7.817	92.18	100
	(1.585)	(1.585)	
40-49	5.145	94.85	100
	(.9642)	(.9642)	
Total	12.59	87.41	100
	(.7102)	(.7102)	

```
Key:  row percentage
      (linearized standard error of row percentage)

Pearson:
  Uncorrected   chi2(5)         =  190.7037
  Design-based  F(4.71, 254.25) =   13.7941     P = 0.0000
```

Using Weight Variables in the Analysis (Cont.)

Example 2: Mean Number of Children Ever Born, by Place of Residence for Women 20-49 Years of Age

```
use “EX2.DTA”
svyset [pweight=WGT2015_2017], strata(SEST) psu(SECU)

create a variable for your subpopulation of ages 20 and older
generate agepop=0
replace agepop=1 if AGER>=20
svy: mean parity, over(agepop metro)
```

STATA Output

```
. svy: mean parity, over(agepop metro)
(running mean on estimation sample)

Survey: Mean estimation

Number of strata =      18        Number of obs    =       5,554
Number of PSUs   =      72        Population size  =  72,218,086
                                  Design df        =          54

        Over: agepop metro
    _subpop_1: yes Principal city of MSA
    _subpop_2: yes Other MSA
    _subpop_3: yes Not MSA
    _subpop_4: 2 Principal city of MSA
    _subpop_5: 2 Other MSA
    _subpop_6: 2 Not MSA
```

Over	Mean	Linearized Std. Err.	[95% Conf.	Interval]
parity				
_subpop_1	1.444937	.0745269	1.295519	1.594354
_subpop_2	1.458681	.0652102	1.327942	1.589419
_subpop_3	1.549797	.0797015	1.390005	1.709589
_subpop_4	.052007	.0149942	.0219455	.0820685
_subpop_5	.0256736	.0111679	.0032833	.0480639
_subpop_6	.0574415	.0272327	.0028432	.1120399

Using Weight Variables in the Analysis (Cont.)

NSFG Website provides three examples of incorporating weight variables into the analysis (https://www.cdc.gov/nchs/nsfg/nsfg_2015_2017_puf.htm)

Example 3: Percentage of Men 20-49 Years of Age Who Have Ever Had One or More Biological Children, by Hispanic Origin and Race

```
use "EX3.DTA"
svyset [pweight=WGT2015_2017], strata(SEST) psu(SECU)

generate biokidsx=0
replace biokidsx=1 if BIOKIDS>0

create a variable for your subpopulation of ages 20 and older
generate agepop=0 replace agepop=1 if ager>=20

svy, subpop(agepop) row percent se: tab hisprace2 biokidsx
```

STATA Output

```
. svy, subpop(agepop) row percent se: tab hisprace2 biokidsx
(running tabulate on estimation sample)

Number of strata   =        18                  Number of obs      =       4,540
Number of PSUs     =        72                  Population size    =  71,617,244
                                                Subpop. no. obs    =       3,654
                                                Subpop. size       =  61,690,581
                                                Design df          =          54
```

RACE AND HISPANIC ORIGIN -- BASED ON 1997 OMB GUIDELINES (NEW FOR CYCLE 7)	biokidsx no	biokidsx yes	Total
Hispanic	40.09	59.91	100
	(2.567)	(2.567)	
Non-Hisp	50.19	49.81	100
	(2.142)	(2.142)	
Non-Hisp	44.9	55.1	100
	(3.28)	(3.28)	
Non-Hisp	57.43	42.57	100
	(4.536)	(4.536)	
Total	48.15	51.85	100
	(1.745)	(1.745)	

```
  Key:  row percentage
        (linearized standard error of row percentage)

  Pearson:
    Uncorrected   chi2(3)         =   45.8912
    Design-based  F(2.84, 153.33) =    6.0172     P = 0.0008
```

Conclusions

- Since 2002, NSFG provide data from a nationally representative sample of men and women aged 15-44. The age range has been expanded to 15-49 in the 2015-2017 NSFG survey.
- NSFG provides a unique opportunity to study life events (marriage, cohabitation, and fertility) and attitudes toward marriage and family. New NSFG data should come out this fall and will be great for new research project.
- Users should always use recoded variables, instead of raw variables because recoded variables had been corrected for possible errors and inconsistencies.
- You should always include weight variables in the analyses to take into account NSFG's complex sample design.
- NSFG uses a repeated cross-sectional study design and respondents were not followed over time. In addition, female respondents and male respondents in the NSFG were selected from different households and are not couples.
- If you have any questions about using NSFG data. Please feel free to contact Hsueh-Sheng Wu @372-3119 or wuh@bgsu.edu