

Introduction to National Survey of Family Growth (NSFG)

Hsueh-Sheng Wu
CFDR Workshop Series
May 23, 2022

BGSU

 Center for Family and Demographic Research

Outline

- Aims of NSFG
- NSFG study design
- 2017-2019 NSFG data files
- Contents in 2017-2019 NSFG Public Data
- Examples of faculty's work using NSFG
- Navigate NSFG website
 - Questionnaire
 - Codebook
 - Data
- Sample Stata code for:
 - Reading NSFG data file into Stata
 - Linking a female respondent file with a pregnancy file
 - Pooling men and women data from the same cycle
 - Pooling data from multiple years
 - Using weight variables in the analysis
- Conclusions

Aims of NSFG

- Address Sec 306 of the Public Health Service Act: “National Center for Health Statistics (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
2017-2019 NSFG	2017-2019	women men	15-49		Interviews spread over several years	11,347 W=6,141 M=5,206

NSFG Study Design (Continued)

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
2015-2019 NSFG	2015-2019	multistage area probability sampling	Blacks, Hispanics, & Teens	Interviewer administered CAPI self-administered ACAPI

2017-2019 NSFG Data Files

The 2017-2019 NSFG survey have public data and restricted data

Public-use data

- Public-used data can be directly downloaded from the NSFG website
- The data files are (1) Female respondent data file, (2) Female pregnancy data file, (3) Male respondent data file
- The 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 are suppressed.
- Some variables are constructed (e.g., inter-pregnancy interval variables), so researchers can still address important research questions without the exact information on the century months of the pregnancies.

Restricted data

- Restricted-use 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-used 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 the restricted data, researchers need to follow the application process that NCHS specifies in <https://www.cdc.gov/rdc/b3prosal/pp300.htm>

Contents in 2017-2019 NSFG Public Data

Table 2. The Topics Covered in the Male and Female Respondent Data files in 2011

Section	Female Respondent data	Male Respondent Data
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	Ever Sex, Sex Communication and Education, Vasectomy and Physical Ability to Father Children, 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 Partner
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	Desires and Intentions for Future Births	Fathering
H	Infertility Services and Reproductive Health	Desires and Intentions for Future Biological Children
I	Insurance; Residence and place of birth; Religion; Past and current work (R and current H/P); Attitudes	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); Attitudes
K		Audio CASI
R	Recode Variables and Imputation Flags	Recode Variables and Imputation Flags
W	Weights and Other Variables	Weights and Other Variables

Contents in 2017-2019 NSFG Public Data (Continued)

Table 3. The file index of the Pregnancy Data in the 2017-2019 NSFG

- | |
|--|
| 1. Respondent ID and Pregnancy Order |
| 2. Section B raw variables |
| 3. Section E raw variables |
| 4. Section B and E pregnancy-based recodes |
| 5. Selected respondent file variables |
| 6. Pregnancy recode imputation flags |
| 7. Respondent recode imputation flags included on pregnancy file |
| 8. Weights and related variables |
| 9. Date of interview, key reference dates, and fieldwork variables |

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.

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/index.htm>

CDC Centers for Disease Control and Prevention
CDC 24/7: Saving Lives. Protecting People™

[A-Z Index](#)

Search Search NCHS

[Advanced Search](#)

National Center for Health Statistics

CDC > NCHS

[f](#) [t](#) [in](#) [rss](#)

[National Survey of Family Growth](#)

About NSFG

What's New

Questionnaires, Datasets, and Related Documentation +

Key Statistics from NSFG +

Publications and Information Products

Bibliography

Research Conferences

Survey Participants

NSFG National Survey of Family Growth

National Survey of Family Growth

The National Survey of Family Growth (NSFG) gathers information on family life, marriage and divorce, pregnancy, infertility, use of contraception, and men's and women's health. The survey results are used by the U.S. Department of Health and Human Services and others to plan health services and health education programs, and to do statistical studies of families, fertility, and health. Links to some of those studies are included on this web site, under "[Publications and Information Products.](#)"

What's New

Data Releases

- [2017-2019 NSFG Public Use Data Files](#) (10/2020)

Publications

- [2015-2019 NSFG Key Statistics](#) (11/2021)
- [Current Contraceptive Status Among Women Aged 15-49: United States, 2017-2019](#) (10/2020)
- [Trends and Patterns in Menarche in the United States:1995 through 2013-2017](#) [PDF - 366 KB] (9/2020)
- [Sexual Activity and Contraceptive Use Among Teenagers Aged 15-19 in the United States, 2015-2017](#) (5/2020)

BGSU

Navigate NSFG Website (Continued)

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

The screenshot shows the National Center for Health Statistics website for the National Survey of Family Growth. The page features a navigation menu on the left with options like 'National Survey of Family Growth', 'About NSFG', and 'What's New'. The main content area is titled 'National Survey of Family Growth' and 'Questionnaires, Datasets, and Related Documentation'. A list of links is provided, including '2011-2019 Combined Data Files', '2017-2019 NSFG', '2015-2017 NSFG', '2013-2015 NSFG', '2011-2013 NSFG', '2006-2010 NSFG', and 'NSFG Cycle 6 (2002)'. A secondary list of links includes '2011-2019 Combined Data Files', '2017-2019 NSFG', '2015-2017 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)', and 'NSFG Cycle 1 (1973)'.

National Center for Health Statistics

CDC > NCHS > National Survey of Family Growth

National Survey of Family Growth

Questionnaires, Datasets, and Related Documentation

- [2011-2019 Combined Data Files](#)
- [2017-2019 NSFG](#)
- [2015-2017 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 \(1973\)](#)

BGSU

Navigate NSFG Website (Continued)

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

The screenshot shows the CDC website for the National Survey of Family Growth (NSFG). The page is titled "2017-2019 NSFG: Public-Use Data Files, Codebooks, and Documentation". The left sidebar contains a navigation menu with the following items: "National Survey of Family Growth", "About NSFG", "What's New", "Questionnaires, Datasets, and Related Documentation" (selected), "2011-2019 Combined Data Files", "2017-2019 NSFG" (selected), "2017-2019 NSFG Codebooks", "2017-2019 NSFG Questionnaires", "2015-2017 NSFG", "2013-2015 NSFG", "2011-2013 NSFG", "2006-2010 NSFG", and "NSFG Cycle 6 (2002)". The main content area features a blue header with the NSFG logo and the title "National Survey of Family Growth". Below this is the main heading "2017-2019 NSFG: Public-Use Data Files, Codebooks, and Documentation". A paragraph of text provides contact information for persons with disabilities: "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." Below this is a section titled "Codebooks and Documentation" with a sub-section for "Codebooks:" containing a bullet point: "Webdoc, the NSFG's interactive online codebook, was deactivated as of 12/31/20. Public-use file indexes (Appendix 1a, 1b, and 1c linked below) can be searched to find variable names or identify relevant variables based on key words in variable labels. The section letters noted in the file indexes correspond to the sections of the codebook, so the file indexes indicate which codebook PDF will contain your variables of interest." On the right side, there is a "On This Page" section with a list of links: "Codebooks and Documentation", "Design and Data Collection Methods", "Variance Estimation Examples", "Design and Weight Variables for Each 2-Year NSFG Public-Use File 2011-2019", "Questionnaires", "Informed Consent Materials", "Downloadable Data Files", and "Program Statements".

2017-2019 NSFG Webpage

NSFG website provides important information on each NSFG survey.

Take 2017-2019 NSFG as an example, the website (https://www.cdc.gov/nchs/nsfg/nsfg_2017_2019_puf.htm) provides:

- Codebooks and Documentation
- Design and Data Collection Methods
- Variance Estimation Examples
- Design and Weight Variables
- Questionnaires
- Informed Consent Materials
- Downloadable Data Files
- Program Statements for SAS, Stata, and/or SPSS.
- Other Data Files

Questionnaire

NSFG Questionnaire provides comprehensive specifications of each question that NSFG staff provide to the computer programmers to create the CAPI instrument.

```
C_INTRO1
CA-0. The next questions are about your marriages and other relationships.

    • ENTER [1] to continue

{ Asked if R has ever been married
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
    (But we want to allow space for only 6 husbands in looping)

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 HSEVERIF 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 HSEVERIF (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.)
```

Codebook

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

TIMESMAR(231-232)

Variable Type : raw

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

value	label	Total
.	INAPPLICABLE	3419
1	1 TIME	2288
2	2 TIMES	358
3	3 TIMES	61
4	4 TIMES	12
98	Refused	2
99	Don't know	1
	Total	6141

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 Website site under the section “Downloadable Data Files.”

Downloadable Data Files

- [Female Respondent Data File \(2017_2019_FemRespData.dat\)](#)
- [Female Pregnancy Data File \(2017_2019_FemPregData.dat\)](#)
- [Male Respondent Data File \(2017_2019_MaleData.dat\)](#)

- Additional Weight variables for data from multiple cycles are available at https://www.cdc.gov/nchs/nsfg/nsfg_2011_2019_combined_files.htm.

2011–2019 Combined Files: Selected Data and Documentation

This page is intended to provide information for users who wish to combine NSFG data from multiple 2-year file releases from 2011-2019 for their research.

Within this page are also descriptions of special supplementary data files for the periods 2011-2015 and 2015-2019, featuring restricted-use Interviewer Observations data and Paradata files.

- [Combining Data across NSFG File Releases from 2011-2019](#)

Weighting in NSFG Analyses

NSFG analysis needs to be weighted to adjust for:

- Complex survey design (not a simple random sampling)
- Oversample sub-populations
- Screener and interview non-response
- Estimate 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 2002-20119

Design variable	2002	2006–2010	2011–2013	2013–2015	2015–2017	2017-2019
Stratum variable	SEST	SEST	SEST	SEST	SEST	SEST
Four Cluster/Panel Variable	SECU_R (fem resp) SECU_P (fem preg) SECU (male resp)	SECU	SECU	SECU	SECU	SECU
Final post-stratified, fully adjusted case weight	FINALWGT	WGTQ1Q16	WGT2011_2013	WGT2013_2015	WGT2015_2017	WGT2017_2019

Note:

1. There is no weight variables for single year estimate because of insufficient statistical power
2. The 4-year weight variables (WGT2011_2015, WGT2013_2017, WGT2015_2019), 6-year weight variables (WGT2011_2017 and WGT2013_2019) and 8-year weight variables (WGT2011_2019) should be used when researchers combined NSFG data across the different spans of years (see https://www.cdc.gov/nchs/nsfg/nsfg_combining_data.htm)

Reading NSFG 2017-2019 Data

Stata : Add one line of codes and Modify three other lines in the Stata command file download to the from the NSFG website.

- Insert this line of codes before the local macro variables and it defines where the data and dictionary files are and where to save the final outcome file

```
cd "C:\nsfg\"
```

- Input the names of data, dictionary, and the output files for the three Stata local macro variables

```
local raw_data "2017_2019_FemPregData.dat"  
local dict "2017_2019_FemPregSetup"  
local outfile "2017_2019_FemPregData.dta"
```

Linking A Female Respondent File with A Pregnancy File

* Sort the female respondent data by ID

```
use 2017_2019_FemRespSetup.dta, clear
sort caseid
save, replace
```

* Sort the pregnancy data by ID and pregnancy order

```
use 2017_2019_FemPregData.dta, clear
sort caseid pregodr
save, replace
```

* Merge and save data

```
use 2017_2019_FemRespSetup.dta, clear
merge 1:m caseid using 2011_2013_FemPregData.dta
```

```
save 2017_2019_fem_pregnancy.dta, replace
```

Pooling Men and Women Data from the Same Time Period

```
*****  
* Select variables from the male file  
*****  
use 2017_2019_MaleSetup.dta, clear  
keep caseid rscrninf wgt2017_2019 secu sest  
save temp1.dta, replace  
  
*****  
* Select variables from the female file  
*****  
use 2017_2019_FemRespSetup.dta, clear  
keep caseid rscrninf wgt2017_2019 secu sest  
save temp2.dta, replace  
  
*****  
* Pool two data sets together  
*****  
use temp1.dta, clear  
append using temp2.dta  
  
save 2017_2019_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_2017_2019_puf.htm)

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

* Read in the data
use "EX1.DTA"

* Specify how data should be weighted
svyset [pweight=WGT2017_2019], strata(sest) psu(secu)

* Generate the variable of age groups
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

* Add value labels
label define agerx 1 "15-19" 2 "20-24" 3 "25-29" 4 "30-34" 5 "35-39" 6 "40-49"
label value agerx agerx

- Generate the variable describing whether respondents currently use the oral contraceptive pill

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

* Conduct the analysis
svy: tab agerx cpill, row se percent

Using Weight Variables in the Analysis (Continued)

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

Number of strata   =      18          Number of obs   =      6,141
Number of PSUs    =      72          Population size  = 72,671,926
                                   Design df         =          54
```

agerx	cpill		Total
	1	2	
15-19	19.51 (1.867)	80.49 (1.867)	100
20-24	23.78 (2.171)	76.22 (2.171)	100
25-29	19.69 (2.277)	80.31 (2.277)	100
30-34	15.28 (1.755)	84.72 (1.755)	100
35-39	6.497 (.9895)	93.5 (.9895)	100
40-49	6.522 (1.003)	93.48 (1.003)	100
Total	13.95 (.829)	86.05 (.829)	100

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

Pearson:
Uncorrected chi2(5) = 236.7424
Design-based F(4.39, 236.89) = 23.0013 P = 0.0000

Using Weight Variables in the Analysis (Continued)

Example 2: Mean Number of Children Ever Born, by Hispanic Origin and Race for Women 15-49 Years of Age

* Read in the data
use "EX2.DTA"

* Specify how data should be weighted
svyset [pweight=WGT2017_2019], strata(sest) psu(secu)

* Conduct the analysis
svy: mean parity, over(HISPRACE2)

Using Weight Variables in the Analysis (Continued)

```
. svy: mean parity, over(HISPRACE2)
(running mean on estimation sample)

Survey: Mean estimation

Number of strata =      18      Number of obs   =      6,141
Number of PSUs   =      72      Population size = 72,671,926
                                   Design df        =           54

    Hispanic: HISPRACE2 = Hispanic
    _subpop_2: HISPRACE2 = Non-Hispanic White, Single Race
    _subpop_3: HISPRACE2 = Non-Hispanic Black, Single Race
    _subpop_4: HISPRACE2 = Non-Hispanic Other or Multiple R
```

Over	Mean	Linearized Std. Err.	[95% Conf. Interval]	
parity				
Hispanic	1.544184	.0825869	1.378607	1.709761
_subpop_2	1.077374	.0379518	1.001285	1.153462
_subpop_3	1.400032	.0704658	1.258757	1.541307
_subpop_4	1.067237	.0964947	.8737773	1.260698

Using Weight Variables in the Analysis (Continued)

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

* Read in the data
use "EX3.DTA"

* Specify how data should be weighted
svyset [pweight=WGT2017_2019], strata(sect) psu(secu)

* Generate a variable showing if respondents have biological kid(s).
generate biokidsx=0
replace biokidsx=1 if biokids>0

* create a dichotomous variable for the subpopulation of ages 20 and older.

generate agepop=0
replace agepop=1 if ager>=20

* Conduct the analysis
svy, subpop(agepop) row percent se: tab HISPRACE2 biokidsx

Using Weight Variables in the Analysis (Continued)

```
. svy, subpop(agepop) row percent se: tab HISPRACE2 biokidsx
(running tabulate on estimation sample)
```

Number of strata	=	18	Number of obs	=	5,206
Number of PSUs	=	72	Population size	=	72,221,885
			Subpop. no. obs	=	4,174
			Subpop. size	=	62,331,053
			Design df	=	54

Race & Hispanic origin of respondent - 1997 OMB standards (RECODE)	biokidsx		Total
	0	1	
Hispanic	41.75 (2.41)	58.25 (2.41)	100
Non-Hisp	50.05 (2.535)	49.95 (2.535)	100
Non-Hisp	45.98 (3.226)	54.02 (3.226)	100
Non-Hisp	58.29 (4.598)	41.71 (4.598)	100
Total	48.68 (2.116)	51.32 (2.116)	100


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

```
Pearson:
Uncorrected chi2(3) = 45.7307
Design-based F(2.89, 156.14) = 4.9479 P = 0.0030
```

Conclusions

- Since 2002, NSFG provides 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 that is currently in the field.
- 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.
- Users 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