2017-2019 National Survey of Family Growth (NSFG): Summary of Design and Data Collection Methods

1. Introduction

This report provides key methodological information to users and analysts of the 2017-2019 National Survey of Family Growth (NSFG), beyond the information included in the User's Guide that accompanied the release of these public-use files in October 2020. The 2017-2019 NSFG includes two years (eight quarters) of data from the continuous NSFG. This two-year period covers the 25th through 32nd quarters of an overall eight-year (32 quarters) period of fieldwork. It follows three prior releases of data from the previous 24 quarters of interviews that took place between September 2011 and September 2017. See "2011-2013 National Survey of Family Growth (NSFG): Summary of Design and Data Collection Methods," "2013-2015 National Survey of Family Growth (NSFG): Summary of Design and Data Collection Methods," and "2015-2017 National Survey of Family Growth (NSFG): Summary of Design and Data Collection Methods" for analogous reports for the first, second, and third data releases.

This web-based report and related, detailed reports are intended to replace the Series 1 and Series 2 reports formerly published by the National Center for Health Statistics (NCHS), but still permit the timely release of essential information on the sample design and data collection methods for the NSFG.

The NSFG moved from a periodically conducted survey design as conducted by NCHS six times from 1973 to 2002, to a continuous survey design in 2006. This transition and new design have been described in prior reports. The three reports listed below document the significant changes made for the 2006-2010 NSFG, as well as providing details on how the survey was planned and designed:

- **"Planning and Development of the Continuous National Survey of Family Growth"**: describes planning for and implementation of the transition from a periodic to a continuous survey, prior to the release of the first data from continuous interviewing.
- **"The 2006-2010 National Survey of Family Growth: Sample Design and Analysis of a Continuous Survey":** describes the sample design and weighting and variance estimation procedures under the continuous design, prior to the release of the first data from continuous interviewing.
- **"Responsive Design, Weighting, and Variance Estimation in the 2006-2010 National Survey of Family Growth"**: presents fieldwork results and weighting, imputation, and variance estimation procedures corresponding to the first release of data (2006-2010) under the continuous design.

The current report builds upon the information available in these prior reports, notes design differences from the 2006-2010 survey period made in the 2011-2019 survey period, summarizes updated production outcomes for 2017-2019, and describes the weighting, variance estimation, imputation, and disclosure risk review and operations that took place to produce the public-use datasets for analysis. This report and three detailed reports for 2017-2019 mirror the same set of reports accompanying the 2011-2013, 2013-2015 and 2015-2017 datasets. Thus, each of the three sets of reports is intended to be "stand-alone," containing the same level of comprehensiveness, with this set of reports including information specific to the 2017-2019 dataset. Where relevant, specific outcomes from the 2011-2013 NSFG, 2013-2015 NSFG and 2015-2017 NSFG are provided alongside those for 2017-2019 for comparison purposes.

2. Background on the National Survey of Family Growth

For background information on the purpose, content, and sponsorship of the NSFG, please see the main NSFG webpage, specifically the "About NSFG" section and the User's Guide for 2017-2019.

As with NSFG surveys in 2002, 2006-2010, 2011-2013, 2013-2015, and 2015-2017, sample design and fieldwork for the 2017-2019 NSFG were conducted by the University of Michigan's Institute for Social Research (ISR) under a contract with NCHS, the latest of which covered 2010-2020. Interviewing for the 2017-2019 survey began in late September 2017 and continued through early September 2019 yielding data files based on two years (or eight quarters) of interviews. Interviews were conducted with a national probability sample of women and men 15-49 years of age living in households in the United States. The 11,347 completed interviews (6,141 with women and 5,206 with men) were administered in person by trained female interviewers using laptop computers, a procedure called computer-assisted personal interviewing (CAPI). A subset of the more sensitive questions was administered using audio-computer assisted self-interviewing (Audio-CASI or ACASI). In this mode, respondents answered the questions on the laptop computer, either by reading them on the screen or listening to the pre-recorded questions read over headphones, and entered their answers directly into the computer. The interviews for women averaged 76.4 minutes, and the interviews for men averaged 51.9 minutes, remaining within the Office of Management and Budget (OMB)-approved lengths of 80 minutes for women and 60 minutes for men. About 5% (583 out of 11,347 interviews) were completed in Spanish, which is the only other language accommodated in the NSFG protocol.

3. Sample Design

This document provides a brief overview of the NSFG sample design. A more detailed description of the sampling procedures can be found in "2017-2019 National Survey of family Growth (NSFG): Sample Design Documentation."

The NSFG sample was designed to meet a number of key objectives including:
1) minimizing the overall design effects for women and men
2) controlling the costs of both screening and interviewing
3) obtaining overall sample size of at least 5,000 interviews per year
4) providing for oversamples of teenagers aged 15-19, non-Hispanic black persons of all ages, and Hispanic persons of all ages

Further, the continuous fieldwork design was planned to provide annual, nationally representative samples, permitting data to be cumulated over multiple years of continuous interviewing. However, the weights provided are based on a minimum of two years' worth of interviews, due to the limited sample sizes in single years of interviewing. For the last two quarters of 2017-2019, larger samples were selected with the goal of increasing the number of completed interviews.

To obtain sample sizes comparable to or higher than prior NSFG data releases, data from the 2011-2013, 2013-2015, 2015-2017, and 2017-2019 NSFG can be combined into four-year, six-year, and eight-year files. The NSFG webpage for Combined Files contains technical guidance for combining data from 2011-2013, 2013-2015, 2015-2017, and 2017-2019 into four-, six-, and eight-year files.

3.1 Sample universe

The survey population, or population of inference, for the 2017-2019 NSFG consists of all non-institutionalized women and men aged 15-49 years as of first contact for the survey, whose usual place of residence is the 50 United States or the District of Columbia. Excluded from the survey population are those in institutions, such as prisons, homes for juvenile delinquents, homes for the intellectually disabled, long-term psychiatric hospitals, and those living on military bases. Included in the sample are age-eligible persons living in non-institutional group quarters (e.g., dormitories, sororities), college students sampled through their parents' or guardians' households, and women and men who are in the military but living off base.

3.2 Sample selection

The NSFG is based on a stratified multi-stage area probability sample, using probability proportionate to size (PPS) selection within each of four key domains, as shown below in Table 1. There are five stages of sample selection:
  1) selection of primary sampling units (PSUs)
  2) selection of secondary sampling units (SSUs)
  3) listing and selection of housing units within SSUs
  4) selection of one of the eligible persons within each sampled household
  5) two-phase sampling for nonresponse

These five stages are briefly outlined below. More details on sample selection are provided in "2017-2019 National Survey of family Growth (NSFG): Sample Design Documentation." Data from the 2010 decennial census were used as the sampling frame for the first two stages of selection.

> 1) Selection of Primary Sampling Units

The **first stage** involved the selection of Primary Sampling Units (PSUs). PSUs are Metropolitan Statistical Areas (MSAs), counties or groups of counties. The United States was divided into 2,149 PSUs on the sampling frame. Of these, 366 were MSAs and 1,783 were non-MSA PSUs that include one or more counties. The PSUs were stratified according to attributes such as Census Division, MSA status, and size. One or two PSUs were selected with probability proportionate to size (PPS) from each stratum. The PPS selection method assigns higher probabilities to PSUs with larger populations. The first stage selection probabilities were inversely related to the probabilities of selection at the second and third stages of selection such that sampling rates were approximately equal for all households within a sampling domain (defined below). Across the eight years of data collection (2011-2019) there were a total of 21 "self-representing" (SR) PSUs, defined as PSUs that were automatically included in national probability samples due to their large population, and an additional 192 non-self-representing (NSR) PSUs, defined as PSUs selected into the NSFG sample that represent not only themselves but other non-self-representing PSUs, for a total of 213 PSUs, plus two for Alaska and Hawaii. A subset of these 215 PSUs was selected for each two-year sampling period. For 2017-2019, there were 65 PSUs: 17 SR and 48 NSR PSUs.

In order to facilitate the NSFG's oversample of subgroups defined by race and Hispanic origin, the measures of size for the PSUs were a weighted combination of household counts. **All Census Block groups were classified into four sampling "domains" shown in Table 1.** Households in domains 2, 3,

and 4 were given a higher probability of selection than those in domain 1. These weighted measures of size were then used in both the first and second stages of selection.

Table 1. Domain definitions and characteristics

Domain	Definition	Total Households	Estimated. Proportion Black	Estimated Proportion Hispanic
1	<10% HH Black, <10% HH Hispanic	65,009,685	0.018	0.022
2	>=10% HH Black, <10% HH Hispanic	19,871,976	0.426	0.029
3	<10% HH Black, >=10% HH Hispanic	20,270,438	0.026	0.380
4	>=10% HH Black, >=10% HH Hispanic	11,564,193	0.301	0.299

NOTE: HH stands for "household."

2) Selection of Secondary Sampling Units

In the **second stage** of selection, Secondary Sampling Units (SSUs or segments) were selected within PSUs. These were composed of one or more Census blocks with a minimum measure of size equal to 50 housing units (HUs). SSUs in domains 2, 3, and 4 had relatively higher combined PSU, SSU, and HU selection rates. These weighted measures of size and sampling rates were set such that interviews with black and Hispanic respondents each constituted about 20% of all interviews. Each PSU was assigned one or two ISR interviewers based on its relative size. For each interviewer, 12 SSUs were selected each year. These SSUs were then randomly divided into four groups, with one group of three SSUs assigned to each quarter. In the last two quarters (Quarters 31 and 32), relatively larger samples were selected with the goal of increasing the number of interviews. In Quarter 32, in order to accommodate these larger sample sizes, two additional area segments were released.

3) Listing and Selection of Housing Units within SSU's

For the **third stage** of selection, interviewers updated commercially-available lists (based on the U.S. Postal Service's Delivery Sequence File (DSF)) of housing units for SSUs. Once these lists were updated, a sample of housing units was selected systematically from geographically-sorted lists of housing units, beginning from a random start.

Beginning in Quarter 13 (2013), a sample design change was implemented with the goal of increasing the percentage of screened households that contain an age-eligible person. This was accomplished by stratifying housing units based on a prediction of whether the unit contained someone age-eligible. The model was selected and estimated using data from previous quarters where the binary eligibility outcome was measured. Key predictors in this model included commercial data that estimate whether an eligible person is in the household. The predicted probability of an eligible person being in the household was used to create strata and then oversample the stratum or strata with higher expected eligibility. In the last two quarters, larger samples of housing units were selected with the goal of increasing the number of completed interviews.

The selected units were then contacted by ISR interviewers to determine if any members of the household were eligible (persons age 15-49 at the time of the screening interview). A full household roster was obtained during the screening interview to identify eligible household members.

4) Selection of Eligible Persons

In households with eligible persons, a **fourth stage** of selection involved selecting one of the eligible persons. The within-household selection rates were set so that about 18.2% of all interviews were with teens aged 15-19 and 55% of all interviews were with females.

5) Two-Phase Sampling for Nonresponse

As was done in NSFG surveys for 2006-2010, 2011-2013, 2013-2015, and 2015-2017, the 2017-2019 NSFG also used a two-phase sampling approach as a **fifth stage** of selection. Each quarter, during week 10, a subsample of active cases was selected for continued follow-up. In weeks 11 and 12, this subsample received a special mailed incentive and the interviewers focused their effort on the fewer cases left in the subsample. Details of this two-phase design are described in Lepkowski et al. (2013) and further below.

4. Responsive Design and Management of Fieldwork

The NSFG sample selection and the fieldwork procedures were designed around an interviewer labor model of 38 "workloads" each quarter, with an expectation of each interviewer working at least 30 hours a week for four quarters. A "workload" refers to the average person-time that each interviewer is expected and budgeted to work. One way to accomplish it, implemented initially and through 2017, is by 38 interviewers each working for 30 hours per week and meeting production goals during the 12-week period. This can also be accomplished with more than 38 interviewers who may work fewer hours per week, or fewer interviewers working more hours per week. The number of sample lines was adjusted each quarter, based on predicted interviewer-level efficiency, to ensure that each interviewer had sufficient sample to support a workload.

The strategy of hiring and training interviewers for at least one year of work (in the NSR PSUs) and guaranteeing them a workload of at least 30 hours a week was intended to minimize attrition and did result in a more stable interviewer workforce over the course of the 2011-2017 fieldwork. Starting in September 2017, in response to an increasingly competitive labor market and increases in attrition among interviewers, ISR made a change to the labor model. In areas where hiring was needed, two interviewers were hired and trained. Further, in order to make it easier to recruit interviewers, the expected weekly time commitment was reduced from 30 hours to 20 in these areas. In PSUs staffed with two interviewers, each interviewer was guaranteed a workload of at least 20 hours a week. This staffing model continued through 2019. In August 2017, a total of 34 interviewers were trained, resulting in 20 PSUs with at least two staff. In August 2018, a total of 38 interviewers were trained, resulting in a total of 24 PSUs with two or more staff. An additional training for 17 interviewers was conducted in March 2019, and in May 2019, five interviewers who had previously worked on NSFG were given refresher training and recertified as interviewers available to travel (i.e., filling in gaps where needed).

The NSFG utilized a responsive design each quarter. The overall goals of the responsive design approach were to balance response rates across key subgroups (defined by sex, race/Hispanic origin and age) and manage the costs of data collection. Details of the responsive design approach are provided elsewhere. Key elements of the responsive design approach include:

- Quarterly data collection with replicates (random subsamples of each annual sample)
- Two phases of data collection each quarter
- Sample design around interviewer workloads to maximize efficiency
- Daily monitoring of key fieldwork indicators
- Planned interventions to direct interviewer effort at specific points in space and time

As noted above, Phase 1 data collection comprised the first 10 weeks of data collection in each quarter. In that time, all sample cases were made available to interviewers, who were directed to focus attention on cases not yet screened or cases already screened and ready for the main interview, depending on what the fieldwork indicators showed. In week 10 of each quarter a subsample of about 1/3 of cases was selected for continued effort in Phase 2 (weeks 11 and 12), using criteria that included response propensity (See "2017-2019 National Survey of Family Growth (NSFG): Sample Design Documentation" for more details on the subsample selected in week 10). Interviewer assignments were reduced so that interviewers could concentrate effort on a smaller number of housing units and selected persons for the final two weeks of data collection while maintaining their overall number of hours worked. This two-phase subsampling design was critical for improving final response rates and controlling costs.

5. Data Collection Activities

This section describes the fieldwork protocols used for NSFG data collection. Interviewer training is first briefly described, followed by a description of the computer equipment used for NSFG. The respondent recruitment or fieldwork protocol is then described. Many of the details of the process are the same as, or similar to, those used for the 2006-2010 NSFG fieldwork (see Groves et al., 2009) and are the same as those used in the 2011-2013, 2013-2015, and the 2015-2017 NSFG.

5.1 Interviewer training

Under the 2010-2020 contract covering continuous NSFG data collection for the period of 2011-2019, interviewer training was generally conducted in September each year, as that is the start of the interviewing period for each year. Interviewers were trained in a centralized location near the contractor's home office in Ann Arbor, MI, where the full ISR staff of NSFG was available to assist. NCHS NSFG staff also attended and monitored most training sessions and if not in-person, participated in "question and answer" sessions remotely. In 2011, the first year of fieldwork under the new contract, all 50 interviewers attended training. In subsequent years (2012-2014), only 17-20 new interviewers attended training. The year 5 training, held in September 2015, marked the half-way point of the data collection contract with four of the eight years of data collection completed. In addition, a relatively large number of questionnaire changes and additions were introduced in 2015. As such, in 2015, 22 continuing interviewers attended portions of the training for a refresher, along with 24 new interviewers who attended the full training session. Additionally, the experienced interviewers were also able to contribute to training sessions that covered strategizing on difficult cases. In 2016, the previous year's training model was utilized with 19 new interviewers attending the fall training session. Attrition trainings were held each year as well. Beginning in 2017, the annual interviewer training was conducted

in late August, rather than in September, allowing the new interviewers to list segments in advance of the quarter and sample selection to take place prior to the start of the quarter. Thus, the sample release and start date was the same in all areas. This approach took place in both 2017 and 2018. The labor model was updated as described in the previous section: 4. Responsive Design and Management of Fieldwork.

Interviewer training consisted of a home-study portion completed prior to attending training, a 1.5 day general interviewing techniques (GIT) training for all newly hired interviewers, five days of study-specific training and a certification interview that all interviewers were required to successfully complete before beginning fieldwork. Bilingual interviewers also completed an additional half-day training session. The home study portion of training included on-line videos and sections of the study manual to review, and an assessment to complete prior to attending in-person training. Study-specific training was conducted primarily in smaller groups and focused on the following topics:
- Reviewing and practicing household listing and screening
- Administering the male and female questionnaires using carefully planned "mock" interview scripts and hands-on practice with the interviewer aids such as the show card booklet and life history calendar
- Learning and practicing various study protocols such as addressing common respondent concerns to "avert refusals," and obtaining informed consent (adults) and assent (minors) for participating in the survey
- Learning other administrative tasks such as entering their time and expense reports.

5.2 Computer hardware, software, and related supplies

The computers used for the 2017-2019 NSFG were Fujitsu Lifebook tablet computers. Computer supplies included an AC adaptor, a car adaptor, an extra laptop battery and headphones for use during the ACASI portion of the interview. Interviewers were also provided with a locking laptop case and shredders for secure disposal of any paper materials bearing confidential information. New laptops were provided to all NSFG interviewers in September 2015, and were utilized for the remainder of the four-year data collection period under this contract.

As in all NSFG surveys since 1995, both screener and main interviews for 2017-2019 NSFG were conducted using computer assisted personal interviewing (CAPI), with ACASI for the most sensitive questions. The entire interview was programmed in the Blaise software (version 4.8) developed by Statistics Netherlands. A change introduced in 2011 involved the use of text-to-speech (TTS, or computer-generated) voice files for the ACASI instruments, rather than a recorded human voice, as used in prior years of NSFG (Couper et al., 2016). Beginning in 2015, interviewers gained the ability to use the laptop computer in tablet mode to complete the household screening. This feature was programmed in Blaise 4.8 as well.

5.3 Fieldwork protocol

The fieldwork protocol utilized in 2017-2019 was essentially the same as the protocol used in 2006-2010 (see Groves et al. 2009), 2011-2013, 2013-2015 and 2015-2017. These procedures were reviewed and approved by the Ethics Review Board (ERB) at NCHS and also reviewed by the University of Michigan Institutional Review Board. The key steps in the recruitment protocol were:

1. An [advance household letter](#) and an [NSFG Question-and-Answer Brochure](#) were mailed to all selected housing units prior to initiating in-person contact. The letter was printed in English on one side and Spanish on the other.

2. If no one was home on the first visit, a "Sorry I Missed You" card was left at the household indicating that the interviewer had stopped by and would return at another time. Return visits were made to households during a different time of day or different day of the week than the initial contact. If a household member was not willing to complete the screening interview at that time, the interviewer answered any questions regarding the survey and the process and offered to return at a more convenient time.

3. When contact was made with a sampled household, the field interviewer introduced herself to the household member by displaying her identification badge and identifying that she was an interviewer from the University of Michigan contacting the household on behalf of the National Survey of Family Growth. The [advance household letter](#) was referenced and the [letter of authorization](#) was shown if necessary.

4. Once establishing that the household member was an adult 18 or older and willing to participate in the brief (less than 5 minutes) household screener, the interviewer conducted the household screener to determine if any household member was age-eligible for the survey. If more than one age-eligible household member was identified, the pre-programmed survey selection algorithm selected one person to be interviewed. If no one in the household was eligible, no further contact was made with the household. Age was the primary basis for ineligibility, however in some cases an age-eligible household member may have been ruled out on the basis of language or other factors. Due to resource and sample size constraints, the NSFG interview could only be conducted in English or Spanish.

5. Once selected to participate, **adult** respondents were provided with [a respondent letter](#) explaining that they had been selected for the survey and a copy of the [informed consent form](#), covering all required elements of informed consent. The interviewer asked if they were willing to participate in the survey. If the respondent agreed, they were provided a $40 token of appreciation in advance of completing the interview and asked to provide an electronic signature acknowledging receipt of the token of appreciation. Before the data collection year beginning in September 2015, a**dult** respondents were required to sign the electronic consent form (adult permission to interview form), but for all data collection from September 2015 through September 2019, the NSFG was granted a waiver of documentation of informed consent by the NCHS ERB.

6. In the case where the selected respondent was a minor, defined as ages 15-17 in most states, signed informed consent and permission was first requested of a parent of the minor respondent prior to talking with him or her. Once parental consent and permission were obtained, the minor was provided with a letter explaining that they had been selected for the survey. The minor was provided a copy of the [minor assent form](#), asked to provide an electronic signature acknowledging their assent, and then provided a $40 token of appreciation in advance of completing the interview. Unlike the case with adult respondents, a signature from the minor's parent on the [parent permission form](#) and a signature from the minor respondent on the minor assent form were both required in order to proceed with the minor's interview.

7. The main interview was conducted in a private setting with the interviewer reading the questions and entering the responses in the laptop. A private setting was defined as having no one over the age of four years within hearing range of the interviewer and respondent. Various aids were used throughout the interview: show cards that the respondent referred to for response categories; question-by-question guidance ("help screens") for the interviewer to read to the respondent if additional information was needed on a particular question; and the Life History Calendar used only for female respondents as a tool to aid in recalling dates and detailed events. All respondents were offered headphones to complete the ACASI section of the interview, but they could choose not to use the headphones and read the questions onscreen if they preferred.

8. At the end of the ACASI section, the respondent was prompted to lock the interview data before returning the computer to the interviewer. This locking made it impossible for the interviewer to back up and view any of the respondents' answers to ACASI, nor could the interviewer back up and alter any prior responses to questions she administered before ACASI. Before leaving the household, the interviewer turned off and further locked the computer and thanked the respondent for his/her participation.

5.4 Use of incentives

As noted above, respondents in Phase 1 of data collection were offered a $40 token of appreciation, paid in cash. Those adult respondents screened in Phase 1 and selected into Phase 2 for a main interview were offered an additional $40 (for a total of $80) as a prepaid token of appreciation for completion of the survey. Households selected into Phase 2 that were not yet screened in Phase 1 were also sent a $5 prepaid token of appreciation for completion of the screener. This protocol was based on earlier research on incentives in the NSFG (see Lepkowski et al. 2013, Appendix II).

5.5 Mailed Screener Experiment

During the final three years of data collection for the NSFG, an experiment was conducted (during quarters 24-27 and quarters 31 and 32) to determine the impact of using a mailed screener compared to the default protocol of face-to-face-only screeners. This experiment tested whether mailed screeners maintained the quality of the data and increased efficiency, since a response to a mailed screener meant the interviewer did not have to go to the household to conduct the screener. (Face-to-face screeners were conducted only as a follow-up to nonresponse to the mailed version.) Specifically, the experiment aimed to determine whether costs could be reduced without harming response rates. Results from quarters 24 through 27 were inconclusive. Some aspects of the experiment were modified for quarters 31 and 32 and the results suggested overall modest cost savings of mailed screeners relative to the control condition, and without a reduction in response rates. For the full report on the household screener experiment, please contact the NSFG team at nsfg@cdc.gov.

6. Production Outcomes

The following series of tables shows key production statistics from the 2017-2019 NSFG, with comparable numbers from the 2006-2010, 2011-2013, 2013-2015 and 2015-2017 NSFG, where appropriate for comparative purposes.

Table 2 provides key summary counts for the overall NSFG sample, as well as averages per quarter of data collection.

Table 2. Total number of sampled addresses, screened eligible households and main interviews; and average number per quarter, 2006-2010, 2011-2013, 2013-2015, 2015-2017, and 2017-2019 NSFG.

	2006-2010	2011-2013	2013-2015	2015-2017	2017-2019
Sampled addresses[a]					
Total	78,082	39,494	40,598	38,890	43,808
Average per quarter	4,880	4,937	5,075	4,861	5,476
Screened eligible households[b]					
Total	32,134	15,287	15,239	15,797	17,879
Average per quarter	2,008	1,911	1,905	1,975	2,235
Main interviews[c]					
Total	22,682	10,416	10,178	10,094	11,347
Average per quarter	1,418	1,302	1,276	1,262	1,418

[a]Sampled addresses are the number of addresses selected into the screener sample.

[b]Screened eligible households are successfully screened addresses containing one or more age-eligible persons.

[c]Main interviews refer to completed NSFG interviews with selected respondents (including partial interviews, which are those where the respondent at least reached the last applicable question before ACASI).

Table 3 presents key indicators of eligibility. Due to the changing demographic composition of the U.S., specifically the age composition, the percent of occupied housing units with an age-eligible person (15-44) declined from 2006-2010 (52.3%) to 2011-2013 (48.8%). To mitigate this decline, a procedure ("housing unit stratification," whereby strata with units likely to contain an eligible person were oversampled) was implemented starting in 2013 to increase eligibility rates (see "2013-2015 National Survey of Family Growth (NSFG): Sample Design Documentation"). This slowed the decline in the percent of occupied housing units with an eligible person: the percent in 2013-2015 (47.7%) would have been even lower in expectation without this procedure. This housing unit stratification introduced in 2013, combined with the expansion of the eligible age range from 15-44 to 15-49 in 2015, mentioned earlier, led to an eligibility rate in the most recent period of 2015-2017 (57.9%) that was higher than the previous data collection years shown in Table 3. These procedures continued in 2017-2019 and led to a very similar eligibility rate (58.1%).

The yield of main interview cases in Table 2 above did not rise appreciably in 2015-2017, despite the increased eligibility rate. However, with the change in the labor model for 2017-2019 (which included a larger number of interviewers – See Section 4 "Responsive Design and Management of Fieldwork"), the yield did go up.

Also shown in Table 3, the percent of housing units with access impediments increased slightly from 2011-2013 to 2013-2015, went down in 2015-2017, but went back up again in 2017-2019.

Table 3. Weighted percent of housing units that were occupied, percent of occupied housing units with an age-eligible person, and percent of occupied housing units with access impediments by data collection release, 2006-2010, 2011-2013, 2013-2015, 2015-2017, and 2017-2019 NSFG.

	2006-2010	2011-2013	2013-2015	2015-2017	2017-2019
Percent of all housing units that were occupied	85.6%	84.4%	86.3%	87.0%	87.6%
Percent of all occupied households with an age-eligible person 15-44 (2006-2015) or 15-49 (2015-2017, 2017-2019)	52.3%	48.8%	47.7%	57.9%	58.1%
Percent of occupied housing units with access impediments*	14.1%	13.6%	15.8%	13.5%	14.8%

NOTE: Results are based on removal of screener and main lines not selected for the second-phase sample.
*Examples of access impediments include locked apartment building doors and gated communities with guards.

Table 4 shows selected indicators of fieldwork effort. The table shows some evidence of increased effort required to complete main interviews in 2013-2015, with an increase of about 0.9 interviewer hours per completed interview (representing about a 9% increase in effort) relative to 2011-2013. These levels of effort remained stable in 2015-2017. The hours per completed interview went up for 2017-2019 to 11.5.

Table 4. Average number of calls (in-person visits) to obtain a screener, a main interview, and the total, and average number of hours of interviewer labor to complete an interview, 2006-2010, 2011-2013, 2013-2015, 2015-2017, and 2017-2019 NSFG.

	2006-2010	2011-2013	2013-2015	2015-2017	2017-2019
Number of screener calls to obtain screening interview	3.3	3.3	3.6	3.6	3.6
Number of main interview calls to obtain main interview[a]	4.0	4.3	4.3	4.1	4.3
Number of total calls to achieve main interview[b]	7.2[c]	7.4[c]	7.8	7.8	7.9
Hours of Interviewer labor per completed interview	9.1	9.8	10.7	10.6	11.5

[a]Mean number of calls per main interview is the average number of main calls on the cases with completed interviews.
[b]Mean number of total calls on a case to achieve main interview is the average number of main and screener calls on the cases with completed interviews.
[c]The average total calls is not equal to the sum of average screener and main calls due to rounding, and the fact that not all completed screener interviews resulted in a main interview.

Table 5 shows the mean NSFG interview length overall, and by sex and age group. The mean interview length for females in 2017-2019 was 76 minutes, which is about 3 minutes higher than the mean in 2015-2017 and about 2 minutes higher than the mean in 2013-2015. The mean interview length for males has remained relatively stable: in 2017-2019 it was 52 minutes – about 2 minutes higher than in

2015-2017. These interview lengths are still within the 80 minutes for female interviews and 60 minutes for male interviews approved by OMB.

Table 5. Mean and median length of interview in minutes, for completed female and male interviews by age group: 2006-2010, 2011-2013, 2013-2015, 2015-2017, and 2017-2019 NSFG.

Sex and age	Mean[a] and median length of interview in minutes				
	2006-2010	2011-2013	2013-2015	2015-2017	2017-2019
Total					
Mean	61.6	66.0	64.1	62.6	65.1
Median	57.4	60.9	59.6	57.9	60.0
Female					
Total					
Mean	70.4	77.5	74.5	73.0	76.4
Median	67.6	73.9	71.8	69.3	72.8
15-19					
Mean	52.4	56.9	49.5	51.4	53.9
Median	47.8	51.9	40.3	46.8	49.7
20-44/ 20-49*					
Mean	74.6	82.1	78.8	77.3	80.7
Median	71.4	78.3	75.7	73.0	76.8
Male					
Total					
Mean	51.2	52.8	51.0	49.9	51.9
Median	48.5	50.5	48.4	46.8	48.7
15-19					
Mean	41.7	43.1	42.4	39.7	41.1
Median	39.3	40.9	40.2	37.7	38.5
20-44 / 20-49*					
Mean	54.0	55.6	53.5	52.4	54.6
Median	51.3	53.4	50.7	49.0	51.3

[a] Means exclude interviews with total lengths greater than 3 standard deviations from the mean.

* 20-44 for 2011-2013 and 2013-2015; 20-49 for 2015-2017 and 2017-2019

Table 6 contains the final case counts by sex, age and race/Hispanic origin for the two years (eight quarters) of interview data included in the 2017-2019 NSFG, along with the comparable numbers for 2006-2010 NSFG (four years or 16 quarters), 2011-2013, 2013-2015, and 2015-2017 NSFG (the latter three both being two years or eight quarters). A total of 11,347 completed interviews or sufficient partial interviews were obtained in 2017-2019. Sufficient partials are respondents who at least answered the last applicable question before ACASI; some respondents may stop the interview then or stop somewhere during the ACASI component; 30 of the 11,347 interviews were classified as partials, and they have "not ascertained" codes assigned to all questions and variables that were missed as a result of their breakoffs.

Table 6. Number of completed interviews in the NSFG in 2006-2010, 2011-2013, 2013-2015, 2015-2017, and 2017-2019[a].

Subgroup	2006-2010 (4 years)	2011-2013 (2 years)	2013-2015 (2 years)	2015-2017 (2 years)	2017-2019 (2 years)
Total	22,682	10,416	10,210	10,094	11,347
Sex					
Female	12,279	5,601	5,703	5,554	6,141
Male	10,403	4,815	4,507	4,540	5,206
Age					
Teen 15-19	4,662	2,131	2,027	1,821	2,033
Adult 20-44 / 20-49*	18,020	8,285	8,183	8,273	9,314
Race/Hispanic origin					
Black	4,411	2,192	2,069	2,284	2,316
Hispanic	4,889	2,495	2,259	2,060	2,888
White and all other	13,382	5,729	5,882	5,731	6,143

[a] Counts include partial interviews, deemed sufficiently complete to remain in the data file.

* 20-44 for 2011-2013 and 2013-2015; 20-49 for 2015-2017

Finally, Table 7 shows the response rates from the 2011-2013, 2013-2015, 2015-2017 and 2017-2019 NSFG, both overall and for key subgroups. These response rates correspond to the American Association for Public Opinion Research Response Rate 4 (AAPOR, 2015). Among the unscreened cases, the number that was eligible was estimated. In this case, we used the eligibility rate from the screened cases, and applied it to the unscreened cases. The denominator of the response rate calculation was the number of known eligible cases in the sample plus the estimated number of eligible cases from unscreened households. The numerator was the number of main interviews (including partial main interviews). In the case of the final response rates, a weighting factor was used that is the inverse of the sampling rate used to select each case into the second phase. Additional information on nonresponse bias for data collection 2011 through 2017 can be found here. An update through 2019 is forthcoming.

Table 7. Phase 1 and final response rates[a] by sex, race/Hispanic origin, and age: 2011-2013, 2013-2015, 2015-2017, and 2017-2019 NSFG

	Unweighted Phase 1 response rate (weeks 1-10)				Final weighted response rate			
	2011-2013	2013-2015	2015-2017	2017-2019	2011-2013	2013-2015	2015-2017	2017-2019
Total	56.0%	51.2%	46.2%	45.4%	72.8%	69.3%	65.3%	63.4%
Female	56.8%	52.8%	47.5%	46.6%	73.4%	71.2%	66.7%	65.2%
Black	62.3%	56.9%	55.2%	50.5%	78.5%	74.5%	72.9%	69.6%
Hispanic	58.0%	53.1%	46.0%	47.5%	77.5%	72.0%	66.9%	66.1%
White and all Other races	54.2%	51.2%	45.1%	44.5%	69.9%	69.6%	64.2%	63.1%
Ages 15-19	61.6%	58.6%	51.5%	50.9%	75.2%	70.7%	67.8%	66.0%
Black	69.3%	63.1%	54.0%	56.1%	79.1%	76.2%	67.5%	73.5%
Hispanic	64.3%	61.8%	53.6%	53.1%	78.8%	72.1%	71.2%	71.0%
White and all Other races	57.1%	55.4%	49.5%	47.1%	72.0%	67.9%	66.3%	59.4%
Ages 20-44 / 20-49*	55.8%	51.6%	46.7%	45.9%	73.1%	71.3%	66.5%	65.1%
Black	60.8%	55.7%	55.4%	49.5%	78.3%	74.2%	74.0%	68.9%
Hispanic	56.4%	50.8%	44.2%	46.3%	77.2%	72.0%	65.9%	65.0%
White and all Other races	53.7%	50.4%	44.3%	44.4%	69.5%	70.0%	63.9%	63.7%
Male	55.0%	49.3%	44.9%	44.1%	72.1%	67.1%	63.6%	61.4%
Black	60.8%	57.4%	53.9%	51.6%	75.5%	75.8%	70.0%	67.1%
Hispanic	53.0%	46.3%	41.4%	42.8%	72.1%	68.1%	60.4%	60.7%
White and all Other races	54.1%	48.1%	43.5%	42.4%	71.0%	64.5%	62.9%	60.1%
Ages 15-19	62.1%	59.1%	53.7%	50.8%	74.0%	71.2%	65.6%	65.3%
Black	66.4%	61.5%	60.8%	54.6%	76.2%	80.5%	68.3%	70.0%
Hispanic	63.0%	61.1%	54.5%	54.4%	73.7%	72.8%	65.0%	68.2%
White and all Other races	59.8%	57.3%	50.5%	46.8%	73.1%	67.5%	64.7%	61.9%
Ages 20-44 / 20-49*	53.1%	46.9%	43.0%	42.6%	71.6%	66.1%	63.2%	60.5%
Black	59.0%	56.2%	52.1%	50.8%	75.3%	74.5%	70.5%	66.4%
Hispanic	49.6%	41.7%	37.7%	39.1%	71.6%	66.7%	59.1%	58.4%
White and all Other races	52.8%	46.3%	42.2%	41.7%	70.6%	63.8%	62.7%	59.8%

[a] Phase 1 response rates are unweighted; final response rates are weighted to account for Phase 2 sampling rates.

* 20-44 for 2011-2013 and 2013-2015; 20-49 for 2015-2017 and 2017-2019

7. Data Preparation for Public Use

For a description of the process by which the survey data were cleaned, edited, and recoded in preparation for dissemination in public-use data files, see the 2017-2019 NSFG User's Guide. While the User's Guide also provides some description of the imputation of recoded variables and the disclosure risk review process, this section below provides additional details on these preparations of the NSFG data for public use.

7.1 Imputation of recodes

Most missing recode values were assigned using regression imputation software in which multiple regression was used to predict a value for the case using other variables in the data set as predictors. For each variable with missing data, a regression model was estimated to predict the values for the missing data. The predicted values had a stochastic element added to reflect uncertainty in the coefficients of the model and the predictions. Categorical outcomes were stochastically classified into a single category. This process was repeated sequentially several times for all the variables with missing data. For more details on the method used (sequential regression imputation), see Raghunathan et al. (2001). For specific details on how the approach was implemented with the NSFG, see Lepkowski et al. (2010) and Lepkowski et al. (2013). Regression imputation followed the same logical constraints that are built into the original recode specifications. To the extent possible, imputed values generated by regression modeling were checked to ensure that the imputed values were within acceptable ranges, and were consistent with other recodes and other data reported by the respondent.

In some cases, recodes were imputed using logical imputation, which involved having a subject-matter expert at NCHS examine variables related to the variable in question, and assign a value that was consistent with those other variables. Logical imputation is an educated guess of the true value when there is any ambiguity.

The recodes with the highest rate of imputation involved income. Regression imputation was used for about 10.2% of cases for both poverty level (POVERTY) and total household income (TOTINCR). For no other recodes did the percent of values imputed exceed 2% of all cases.

Regardless of whether any values on a recode were imputed, every NSFG recode has a corresponding imputation flag variable indicating whether the value was based on questionnaire data, logical imputation, or regression model-based imputation. These flags allow users the flexibility to handle imputed cases as they may choose for their own analyses. However, it is the recommendation of NCHS that imputed values be retained in analyses to generate consistent point estimates for the population.

7.2 Procedures to minimize risk of disclosure for individual-level data

Before any NSFG public-use file is released by NCHS, a number of disclosure risk reduction steps are taken to protect the confidentiality of respondents. First, NCHS staff provided specifications to ISR for modifying the data files for public use to prevent disclosure of the identities of the respondents. This included suppression of a significant number of century-month dates and other variables, as well as the creation or collapsing of additional variables that could be used to identify small but visible groups or to match NSFG respondents with external data sets. Next, the proposed NSFG public-use files and related

documentation were reviewed by the NCHS Disclosure Review Board (DRB), chaired by the NCHS Confidentiality Officer. Third, in response to the DRB's review, the NSFG staff and contractor made further changes where necessary to minimize the risk of disclosure. All of these changes are described in the User's Guide, either in the PUF indexes that show the full layout of the files for public use (Appendix 1) or the lists of restricted-use analytic variables to be made available only in the Research Data Center (Appendix 7). In addition, the codebook entries include a special note for each public-use variable where any disclosure risk reduction action has been taken. Lastly, before the release of the public-use files, the values of some variables were altered for some respondents in a process called statistical perturbation. Perturbation is a technique that changes the data before dissemination in such a way that the disclosure risk for the data is decreased but the information content is retained as far as possible. Statistical perturbation was done by a separate team at ISR, and some details of the perturbation process are kept from other ISR or NCHS staff to preserve the integrity of the process. In general, the process involved identifying variables eligible for perturbation, and by determining an appropriate level of perturbation (the proportion of cases whose values may be changed). Once the variables were chosen for perturbation and a perturbation rate determined, a random subset of eligible cases with non-missing values for that variable was selected and the values were deleted for those cases. Then the same sequential regression procedure used to impute for missing data (see above description of imputation) was used to impute the values for those cases set to missing. The resulting "perturbed" values were therefore generally based on the same multivariate models used for the imputation process. These perturbed values were recoded, if necessary, to be consistent with existing recode specifications. The resulting distributions on each of the perturbed variables were then carefully checked to make sure that the recode specifications were satisfied. The NSFG public-use files and the restricted-use files (as made available in the RDC) both contain the same perturbed values for these cases and variables.

7.3 Weighting and variance estimation

The development of weights and sample design characteristics for variance estimation are briefly described here. For more detail, see "2017-2019 National Survey of Family Growth (NSFG): Weighting Design Documentation" and "2017-2019 National Survey of Family Growth (NSFG): Sample Error Estimation Design."

The final case weights for the 2017-2019 NSFG include: 1) a base weight for the housing unit and person selection probabilities, 2) a nonresponse adjustment, and 3) a post-stratification factor. The weights were also trimmed to control the variance of the weights, since highly variable weights may inflate estimates of standard errors. The base probability of selection was calculated from the five separate stages of sampling described earlier.

In order to adjust for any potential bias, nonresponse adjustment factors were developed. Sample-based unit nonresponse adjustments were developed by generating predicted probabilities of response using all available data for respondents and nonrespondents at the screener and main interview levels. Screener and main interview cases had different response processes. Therefore, we have modeled these separately in the adjustment process. In addition, there was slightly different data available at each level. The information on the unscreened cases was somewhat sparser. It included data from the Census Block file as well as information from the paradata, in particular, interviewer observations (see documentation on weighting). Once the probabilities of response were estimated, they were classified into deciles, and the inverse of the response rate within each decile was used as a nonresponse adjustment factor. This was done separately for screener and main interview probabilities.

The last component of the weight is a post-stratification factor. Post-stratification weights the sample to match population totals known from a source such as the Census. This can reduce sampling error and also may help reduce biases due to nonresponse or noncoverage. The selected factors used for post-stratification were age (in seven categories: 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, and 45-49), sex, and race/Hispanic origin (in three categories: black non-Hispanic, non-black non-Hispanic, and Hispanic). This created 42 (7x2x3) separate cells for which we compared population counts to estimated totals. The post-stratification factor for each cell was the population total divided by the sample estimate of that total.

The base probabilities of selection, nonresponse adjustments, and post-stratification factors were then combined to form a single, final weight: WGT2017_2019. Extreme values of this weight were trimmed in order to reduce the variability of the weights. **It is recommended that this weight variable be used for all analyses conducted from the two-year file.**

On the NSFG Combined-File page for 2011-2019, separate, fully-adjusted weights are also available corresponding to the four-year periods 2011-2015, 2013-2017, 2015-2019; the six-year periods 2011-2017 and 2013-2019; and the eight-year period 2011-2019. This page contains technical guidance for using these four-, six-, and eight-year combined data files and the provided four-, six-, and eight-year weights for 2011-2019.

Table 8 shows the mean weights for key subgroups, along with the potential increase in variance due to weighting (as estimated using 1 + L; see Kish, 1992). This measure (1 + L) is a global measure (i.e., not specific to any one variable) that assesses the extent to which the variability of an estimated mean or proportion might be increased because of variability in the weights. A value of 1.0 indicates no contribution to variability due to weighting; a value of 2.0 suggests that there is a potential for the variability of estimate to double due to the weights. After trimming, the minimum weight is 1340.9 and the maximum is 83,332.1.

Table 8. Mean final weights (after post-stratification to Census data and trimming), and potential increases in variance due to the weights (1 + L), by sex, age group, and race/Hispanic origin, 2011-2013, 2013-2015, 2015-2017, and 2017-2019 NSFG.

	Sample size				Mean weight				Increase in variance (1+L)			
	2011-2013	2013-2015	2015-2017	2017-2019	2011-2013	2013-2015	2015-2017	2017-2019	2011-2013	2013-2015	2015-2017	2017-2019
Total	10416	10205	10094	11347	11655.8	12018.5	14249.6	12769.4	2.27	2.03	2.37	2.19
Male	4815	4506	4540	5206	12569.0	13572.4	15774.7	13872.8	2.19	1.93	2.24	2.13
Female	5601	5699	5554	6141	10870.8	10789.9	13002.9	11833.9	2.34	2.10	2.48	2.22
15 to 19	2131	2027	1821	2033	9272.8	9633.6	10746.6	9613.2	2.24	1.89	2.26	2.20
20 to 44 / 20-49*	8285	8178	8273	9314	12268.8	12609.6	15020.7	13458.3	2.25	2.03	2.35	2.15
Hispanic	2495	2258	2060	2888	9370.5	10597.0	13334.4	9991.0	2.36	2.10	2.44	2.32
Black	2192	2068	2284	2316	7803.9	8305.0	8806.2	8641.6	2.41	2.07	2.79	2.35
Other	5729	5879	5731	6143	14125.0	13870.8	16752.9	15631.8	2.08	1.92	2.16	1.98

* 20-44 for 2011-2013 and 2013-2015; 20-49 for 2015-2017 and 2017-2019

In addition to differential weighting, the NSFG design is a stratified cluster sample. This stratification and clustering should be accounted for when estimating variance. In order to reflect the sample design as adequately as possible, without risking disclosure of the identity of respondents, we have created pseudo-strata and pseudo-clusters for variance estimation purposes. The clusters are identified by the variable SECU, and are numbered 1, 2, 3, and 4. These SECUs are nested within pseudo-strata, i.e., unique SECUs are identified by the combination of SEST and SECU. The pseudo-strata are contained in the variable SEST. It is recommended that these variables (SEST and SECU) be used for any estimate of variance (see also Guidelines for Analysis below).

Table 9 shows estimated percentages and standard errors (reflecting the complex design) for four selected statistics, by race/Hispanic origin, age and sex, for 2011-2013, 2013-2015, 2015-2017, and 2017-2019 NSFG. These can be compared to estimates from 2002 NSFG and 2006-2010 NSFG Table X in Lepkowski et al. (2013), but it is important to keep in mind that the latter come from four years of data, while the estimates below are from two years of data in each data release. Also note that the figures presented in this table will not necessarily match those from similarly-described recodes or other variables included in the public-use data files or used in published reports. The variables in this table were created using preliminary data and/or may have been coded differently than the recodes and other variables provided for public use. The primary purpose of this table is to compare Ns and standard errors for similarly-defined variables across all four data releases for the period 2011-2019.

Table 9. Estimated standard errors for four selected statistics, by selected characteristics 2011-2013, 2013-2015, 2015-2017, and 2017-2019 NSFG

Subgroup	2011-2013			2013-2015			2015-2017			2017-2019		
	n	Estimated percent	Standard error	n	Estimated percent	Standard error	n	Estimated percent	Standard error	n	Estimated percent	Standard error
Percent of current contraceptors who were using the oral contraceptive pill												
All	3,308	25.9%	1.34	3,307	25.3%	1.24	3,367	19.4%	1.03	3,815	21.3	1.24
Hispanic	808	19.0%	2.79	740	18.2%	1.82	686	14.4%	2.20	969	13.0	1.37
Non-Hispanic White	1,632	29.3%	1.86	1,715	29.6%	1.83	1,741	22.6%	1.57	1,896	25.6	1.75
Non-Hispanic Black	702	18.2%	2.15	644	20.2%	2.43	769	13.8%	1.94	756	13.5	1.90
Non-Hispanic Other	166	29.0%	5.77	208	17.0%	4.16	171	15.4%	3.50	194	23.2	3.99
Percent of men who are married or cohabiting and intend to have a(nother) birth												
All, age 15-49*	1,215	58.1%	1.90	1,191	58.7%	1.94	1,172	56.1%	1.98	1,351	55.1	1.85
Age 15-19	19	88.8%	8.97	10	79.1%	12.3	7	90.1%	10.92	14	85.3	8.80
Age 20-24	150	84.2%	3.13	133	81.6%	3.47	105	88.3%	4.25	98	85.3	4.84
Age 25-29	323	79.0%	3.71	306	82.0%	2.69	252	85.4%	3.16	276	80.3	2.63
Age 30-34	311	61.5%	3.59	329	64.5%	3.68	274	65.5%	4.10	338	71.2	3.89
Age 35-39	226	46.7%	3.95	235	40.4%	5.00	238	56.7%	5.01	297	42.3	3.36
Age 40-44	186	17.6%	3.51	178	21.0%	3.76	159	22.6%	4.52	202	35.6	4.19
Age 45-49							137	8.2%	2.46	126	9.8	2.25
Percent of females and males 15-19 who have ever had sexual intercourse												
Females age 15-19	1,037	44.8%	2.80	1,010	41.0%	2.38	924	42.0%	3.05	970	41.5	2.28
Males age 15-19	1,088	47.1%	2.22	999	42.1%	1.93	886	37.8%	2.95	1,032	40.2	2.38
Percent of single live births in the last 5 years that were breastfed at all												
All	1,657	75.3%	1.97	2,387	77.4%	1.93	1,267	81.9%	1.79	1,228	85.1	1.50
Hispanic	498	76.4%	3.61	692	86.0%	2.38	320	90.8%	2.22	377	84.3	2.75
Non-Hispanic White	711	78.4%	2.60	1,021	75.5%	2.90	558	82.2%	1.99	485	87.0	2.35
Non-Hispanic Black	366	55.2%	3.33	500	60.4%	4.67	316	68.9%	4.22	294	74.8	2.73
Non-Hispanic Other	82	86.1%	5.19	174	90.4%	2.32	73	71.9%	8.33	72	95.4	3.01

*15-44 years for 2011-2013 and 2013-2015; 15-49 years for 2015-2017 and 2017-2019

8. Accounting for Complex Sample Design in Analysis

The data collected in the NSFG are obtained through a complex, multistage sample design that involves stratification, clustering, and oversampling of specific population subgroups. The final weights provided for analytic purposes have been adjusted in several ways to permit calculation of valid estimates for the noninstitutionalized, household-based population age 15-49 of the United States.

NSFG users are reminded that the use of standard statistical procedures based on the assumption that data are generated via simple random sampling (SRS) will generally produce <u>incorrect</u> estimates of variances and standard errors when used with NSFG data. Applying SRS techniques to NSFG data will generally produce standard error estimates that are, on average, too small, and are likely to generate results that are subject to excessive Type I error. For further details on analysis of complex sample survey data, see Heeringa, West, and Berglund (2010).

Analysts are strongly encouraged to use appropriate statistical software to reflect the complex sample design in their analyses. Several software packages are available for analyzing data collected from complex survey samples. The key design variables for analysis are:
- SEST: Stratum variable
- SECU: Cluster variable
- WGT2017_2019: Final weight variable

Examples of program statements in SAS and Stata that illustrate the correct use of the design variables for variance estimation can be found on the 2017-2019 page under the title, "Variance Estimation Examples."

9. References

The American Association for Public Opinion Research (2015), *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 8th edition*. AAPOR. Available at www.aapor.org.

Couper, M.P., Berglund, P., Kirgis, N., and Buageila, S. (2016), "Using Text-to-Speech (TTS) for Audio Computer-Assisted Self-Interviewing (ACASI)." *Field Methods*, 28(2): 95-111. 10.1177/1525822X14562350.

Groves, R. M. and M. Couper (1998). <u>Nonresponse in Household Interview Surveys</u>. New York, Wiley.

Groves, R.M. et al. (2009), *Planning and Development of the Continuous National Survey of Family Growth*. Vital and Health Statistics, Series 1, No. 48. Hyattsville, MD: National Center for Health Statistics. Available at http://www.cdc.gov/nchs/data/series/sr_01/sr01_048.pdf.

Heeringa, S.G., West, B.T., and Berglund, P.A. (2010), *Applied Survey Data Analysis*. Boca Raton, FL: CRC Press.

Kish, L. (1992), "Weighting for Unequal P_i." *Journal of Official Statistics*, 8(2): 183-200.

Lepkowski, J.M. et al. (2010), *Continuous National Survey of Family Growth: Sample Design, Sampling Weights, Imputation, and Variance Estimation, 2006-2008*. Vital and Health Statistics, Series 2, No. 150. Hyattsville, MD: National Center for Health Statistics. Available at: http://www.cdc.gov/nchs/data/series/sr_02/sr02_150.pdf.

Lepkowski, J.M. et al. (2013), *Responsive Design, Weighting, and Variance Estimation in the 2006-2010 National Survey of Family Growth*. Vital and Health Statistics, Series 2, No. 158. Hyattsville, MD: National Center for Health Statistics. Available at http://www.cdc.gov/nchs/data/series/sr_02/sr02_158.pdf.

Raghunathan, T., Lepkowski, J.M., Van Hoewyk, J., and Solenberger, P. (2001). "A Multivariate Technique for Multiply Imputing Missing Values Using a Sequence of Regression Models." *Survey Methodology*, 27(1): 85-95.

10. Appendix I: Glossary

ACASI – Audio computer-assisted self-interviewing. In the NSFG, a portion of the questionnaire was completed by the respondent using ACASI. The interviewer asked the respondent to use earphones, which delivered an audio recording of the questions. The question text was also displayed on the laptop monitor. The respondent chose a desired response option to each question, using the laptop keyboard. The software directed the respondent to the next appropriate question based on the answers entered. As in all past NSFGs that were computerized, the respondent in the 2017-2019 NSFG performed these steps privately while the interviewer completed other tasks nearby, in an attempt to offer the respondent as much privacy as possible. ACASI has been offered in both English and Spanish since the 1995 NSFG.

Blaise – A software system developed by Statistics Netherlands, which presents the questions in a questionnaire, such as the NSFG. Blaise was programmed to route the respondent to the next appropriate question, store the respondent's answers, and check the consistency of one answer with answers to other related questions. Blaise has been used for the NSFG since 1995.

Call – In-person visit by an interviewer to a housing unit in the NSFG sample. Household calling for screener and main interviews was done only in person in the NSFG. Some calls resulted in a *contact* (speaking with someone in the household), while other calls resulted in no contact (either the address was not occupied or no one was at home). Thus, calls represent any visit, regardless of outcome.

CAPI – Computer-assisted personal interviewing, in which the interviewer used a laptop computer to administer questions in the interview. The laptop displayed question text for the interviewer to read and provided any other necessary instructions to the interviewer. Interviewers recorded the respondent's answers using the keyboard. Software directed the interviewer to the next appropriate question based on the answers entered.

Contact Rate – The percentage of sample households where an interviewer talked with someone at the household at the screener stage (i.e., the screener contact rate); at the main interview stage, the percentage of sample persons who met with the interviewer on one or more visits to the household by the interviewer (i.e., the main interview contact rate).

Cooperation Rate – The percentage of sample households that were contacted and granted a screener interview (i.e., screener cooperation rate); or the percentage of sample persons contacted who granted a main interview (i.e., main interview cooperation rate).

Coverage Error – Deviations between the characteristics (e.g., values of estimated population characteristics) of the sampling frame and the desired target population. Coverage errors arose from the

failure to include some households containing eligible persons in the list of households within segments and failure to list some eligible persons within sample households on the sampling frame.

DSF, or Delivery Sequence File—The Delivery Sequence File from the U.S. Postal Service lists all addresses to which mail is currently delivered by the Postal Service. In most areas, the DSF was the basis for a list of housing units from which listing for the NSFG was done.

Domain – A stratum; a group of sampling units (such as blocks) placed in the same subset from which a sample of units was selected.

Double (or two-phase) sample – A subsample of non-responding sample cases (either at the screener stage or the main interview stage), selected for further follow-up efforts after the completion of the first phase of data collection. NSFG has used such a subsample follow-up approach since the 2002 survey: that is, in Cycle 6 (2002), 2006-2010, 2011-2013, 2013-2015, 2015-2017, and 2017-2019.

Electronic Listing Application (ELA) – A computer application that was used by interviewers for field listing. The application allowed interviewers to update lists of addresses that have been purchased from a vendor. The application applied a set of consistency checks in much the same manner as a CAPI instrument to insure that listings were correct.

Eligible household – A household containing at least one person who is eligible for the NSFG—that is, males or females 15-49 years of age at the date on which the screener was completed, and living in the household population of the United States (all 50 states or the District of Columbia). It is not known whether a selected household has an eligible person until the household screener is conducted. If a household had two or more persons 15-49 years of age, one of these persons was selected randomly for the NSFG main interview.

Eligibility rate – The percentage of sample cases that were members of the target population. In the 2015-2017 and 2017-2019 NSFG the eligibility rate was the percentage of households that contained a person aged 15-49. In the 2011-2013 and 2013-2015 NSFG it was the percentage of households containing a person aged 15-44.

Epsem – Equal probability selection method; a sample design that gives all sample units an equal chance of selection.

Institute for Social Research, University of Michigan – The Institute for Social Research (ISR) at the University of Michigan conducted the fieldwork and data processing for the 2017-2019 NSFG under a contract with NCHS. ISR has several Centers that participated in the NSFG: the Survey Research Center (SRC) provided overall coordination and was responsible for data collection, weighting, and variance estimation; the Inter-university Consortium for Political and Social Research (ICPSR) processed data and developed codebook documentation and web-based systems; and the Population Studies Center provided substantive expertise on demography and family growth.

Institutional Review Board (IRB) – A committee of peer and community reviewers of research procedures involving human subjects that weighs the benefits of the research relative to the risks of harm to human subjects. The NSFG was reviewed and approved by the NCHS IRB, which NCHS refers to as the "Ethics Review Board," or ERB.

Intervention – In the continuous interviewing design (including the 2017-2019 NSFG), changes in interviewing practice based on instructions communicated to field staff by central management staff to resolve imbalances in the sample or to address problems that arose during fieldwork. This included instructions to interviewers to focus on completing screening interviews, and to prioritize cases belonging to categories with lower than average response rates.

Item imputation – The process of assigning answers to cases with missing data ("don't know," "refused," or "not ascertained"). In the NSFG, item imputation was only performed on approximately 600 "recoded variables," or "recodes" (defined below, under "recodes"), rather than all of the thousands of variables in the data set. The purposes of imputation are to make the data more complete, more consistent, easier to use, and, most importantly, to reduce bias caused by differential failure to respond. For example, if a respondent's educational level was missing and a value of "high school graduate" was assigned, education was imputed. As in past NSFG surveys, imputation was done in two ways in the 2017-2019 NSFG, logical and regression imputation. Regression imputation uses a regression equation to estimate a value for a case with missing data. Regression imputation was used to assign most of the imputed values. Occasionally, however, logical imputation was used. Logical imputation uses a subject-matter expert to assign a value based on the value of other variables for the case with missing data. For nearly all of the recoded variables for which imputation was done in the continuous NSFG, less than 2 percent of the cases received an imputed value.

Life history calendar – A visual presentation of a calendar covering the reference period of various questions, used to help the respondent record key personal events used as landmark events to cue memories of the dates of events measured in the survey. In the 2017-2019 NSFG the female interview used a life history calendar as a recall aid for sections of the interview with more challenging recall tasks, such as the pregnancy and contraceptive history sections.

Main interview – An interview sought with the selected household member within sample households containing an eligible target population member. If the screening interview revealed that the household contained one or more persons 15-49 years of age, a main interview was requested from one of those persons. If there were two or more persons 15-49, one such person was selected at random for the main interview, based on pre-programmed selection probabilities.

Measure of Size – A value assigned to every sampling unit in a sample selection. Typically measures of size are a count of units associated with the elements to be selected. This allows different probabilities of selection across the various units of unequal sizes. For a description of the measures of size used by the 2017-2019 NSFG, please see "[2017-2019 National Survey of Family Growth (NSFG): Sample Design Documentation](#)," sections 2.4 and 3.1.

Multi-phase design – A survey design that changes its sample design or recruitment protocol over different sets of sample cases or over time periods of the survey, in order to obtain optimal balance of costs and quality of survey estimates.

National Center for Health Statistics (NCHS) – NCHS is the United States' principal health statistics agency. It designs, develops, and maintains a number of systems that produce data related to demographic and health concerns. These include data on registered births and deaths collected through the National Vital Statistics System, the National Health Interview Survey (NHIS), the National Health and Nutrition Examination Survey (NHANES), the National Health Care Survey, and the National Survey of Family Growth (NSFG), among others. NCHS has conducted the NSFG since 1973. NCHS is one of the

"Centers" for Disease Control and Prevention (CDC), which is part of the U.S. Department of Health and Human Services.

Office of Management and Budget (OMB) Clearance – OMB reviews survey materials and questionnaires proposed for use by government agencies under the provisions of the Paperwork Reduction Act. The review is conducted by the OMB's Office of Information and Regulatory Affairs. No survey of more than 9 persons can be conducted by a U.S. government agency without review and approval by OMB.

Paradata – Information collected via computer software or interviewer observations describing the sample unit, interactions with sample household members, or features of the interview situation. The NSFG used observations of characteristics of sample housing units to reduce the number of callbacks; used statements made by household screener informants in order to diagnose their concerns about the survey; used call record data to model the probability of obtaining an interview on the next visit; and used observations of the respondent during ACASI for measurement error modeling. Some paradata are labeled as "process data."

Phase – A period of data collection during which the same set of sampling frame, mode of data collection, sample design, recruitment protocols, and measurement conditions are used. Starting with the 2006-2010 NSFG, the NSFG has used a two-phase approach combined with 12-week quarters. Phase 1 consisted of weeks 1-10, in which the standard protocol was used, although paradata were used to optimize the efficiency of the interviewers. Phase 2 consisted of weeks 11-12, in which a subsample of non-respondents from phase 1 was offered higher incentives and certain other rules were changed. (See text for detail.)

Public-use file – An electronic data set containing respondent records from a survey with a subset of variables collected in the survey that have been reviewed extensively (within NCHS) to assure that the identities of the respondents are protected. This file is disseminated by NCHS to fulfill its obligations to provide data files for public use as part of the federal statistical system, as well as to fulfill its obligations to the cosponsoring agencies.

PSU – A primary sampling unit. The first stage selection unit in a multistage area probability sample. In the NSFG, PSUs are counties or groups of counties in the United States; there were 215 PSUs selected into the NSFG sample for 2011-19.

Race/Hispanic origin – Race/Hispanic origin was used in this report *as it was used to select the NSFG sample.* Three categories were used for purposes of the NSFG sample design: Hispanic, non-Hispanic black, and all other. Hispanic and non-Hispanic black men and women were selected at higher rates than others in the NSFG, in order to obtain adequate numbers of Hispanic and black persons to make reliable national estimates for these groups. Thus, in this report, tables showing "race/Hispanic origin" show the three categories <u>used to design and select the sample</u>**.** In contrast, in reports that are designed to present <u>substantive results</u>, the "all other" category is often split into "non-Hispanic white" and "non-Hispanic other" categories.

Recodes or recoded variables – It is not possible to edit or impute all of the variables in the continuous NSFG data file. NSFG staff selected about 600 variables from the NSFG data file that were constructed, edited, and imputed. These are called recodes or recoded variables. Recodes are variables that are likely to be used frequently by NCHS and other data users. They were edited for consistency, and missing

values were imputed. Many (but not all) of these recoded variables were constructed from other variables in the NSFG; some were constructed from a large number of other variables. Other variables in the data file were not edited or imputed in this way.

Replicate – A probability subsample of the full sample design. The complete sample consists of several replicate subsamples, each of which is a small national sample of housing units. Replicate samples were released over the data collection in order to control the workflow of the interviewers. In responsive designs, early replicates are used to measure key cost and error features of a survey.

Respondent – A person selected into the sample who provides an interview. In the 2017-2019 NSFG, the "respondents" are the 6,141 women and 5,206 men 15-49 years of age who completed the NSFG interview.

Response rate – Respondents to a survey divided by the number of eligible persons in the sample. In this report, the response rate is the number of respondents (15-49 years of age) divided by the number of eligible persons (15-49 years of age). Given that not all screeners were completed, the number of eligible persons is not known precisely, so this number is estimated.

Responsive design – Survey designs that pre-identify a set of design features potentially affecting costs and errors of survey statistics; identify a set of indicators of the cost and error properties of those features; monitor those indicators in initial phases of data collection; alter the active features of the survey in subsequent phases based on cost/error tradeoff decision rules; and combine data from the separate design phases into a single estimator.

Sample Line – 'Sample line' is a 'hold-over' term from an era in which interviewers were sent to selected area segments (blocks, or linked groups of blocks) to list all housing units. The listing was done on paper, and later keyed to a master list. The sample for any given survey was selected from the master list. The housing units listed were 'lines' on the listing sheet, and the terminology was applied to the electronic records in the master list. This design used primarily U.S. Postal Service Delivery Sequence File (DSF) addresses obtained from a commercial firm in each segment. These lists were updated by field interviewers visiting the segment and comparing the DSF list to the housing units observed by the field interviewer. Housing units not on the list were added. Housing units that were on the list but not present in the segment were deleted. Prior to 2017, in segments where the commercial firm could not provide adequate numbers of addresses (for example, in rural areas where rural delivery routes are used, and no house numbers or street names are available in the DSF), 'scratch listing' was done. Interviewers visited these segments and listed all housing units directly into a laptop. Listed addresses were uploaded to the central office at the end of each day of listing. After 2017, addresses from the DSF were available for all segments. The 'master file' contains addresses from the DSF and from scratch and update listings. We on occasion use the term 'sample lines' to refer to the electronic records in this file. Thus, sample lines are addresses, and not necessarily housing units. They become sample housing units once selected and households when the interviewer visits and finds the housing unit occupied.

Sampling variance – The sampling variance is a measure of the variation of a statistic, such as a proportion or a mean, which is due to having selected a random sample instead of collecting data from every person in the full population. It measures the variation of the estimated proportion or mean over repeated samples. The sampling variance is zero when the full population is observed, as in a census. For the NSFG, the sampling variance estimate is a function of the sampling design and the population parameter being estimated (for example, a proportion or a mean). Many common statistical software

packages compute "population" variances by default; these may under-estimate the sampling variance. Estimating the sampling variance requires special software, such as those discussed in this report.

Sampling weight – For a respondent in the NSFG, the estimated number of persons in the target population that he or she represents. For example, if a man in the sample represents 12,000 men in his age and race/Hispanic origin category, then his "sampling weight" is 12,000. The NSFG sampling weights adjust for different sampling rates (of the age and race/Hispanic origin groups), different response rates, and different coverage rates among persons in the sample, so that accurate national estimates can be made from the sample. Because it adjusts for all these factors, it is sometimes called a 'fully adjusted' sampling weight.

Screening interview – Sometimes called a "household screener," a screening interview is a (usually short) set of questions, asked of a household informant with the chief goal of determining whether the household contains anyone eligible for the survey. In the NSFG, the screening interview consisted of a household roster, collecting age, race, Hispanic origin, and sex. Those households having one or more persons 15-49 years of age were eligible for a main interview. In the NSFG, only persons 18 and older could be screener informants.

Self-representing area – A county or group of counties forming a primary sampling unit with population counts sufficiently large to be equal to or greater than the typical stratum size in the U.S. national sample. Such PSUs were thus represented in all draws of a national sample using the design. The sampling probabilities for persons in such areas were designed to be equal to those applicable in smaller PSUs, called non-self-representing areas.

Segment – A group of housing units located near one another, all of which were selected into the sample.

Simple random sample – A sample in which all members of the population are selected directly and have an equal chance to be selected for the sample. The NSFG sample was not a simple random sample. The NSFG sample was stratified, selected in stages, and employed unequal chances of selection for the respondents, varied by age, race/Hispanic origin, and sex. Such designs are referred to as "complex" and require special software to estimate the variance of statistics computed from a sample with a complex design.

Strata; Stratification – Stratification is the partitioning of a population of sampling units into mutually exclusive categories (strata). Typically, stratification is used to increase the precision of survey estimates for subpopulations important to the survey's objectives. In the 2017-2019 NSFG, those groups included teenagers (15-19 years of age), Hispanic men and women, and Non-Hispanic black men and women. To obtain larger and more reliable samples of these groups, the NSFG sample was stratified: in the first stage of selection, PSUs were stratified using socioeconomic and demographic variables; in the second stage of selection, segments within each PSU were stratified by the concentration of black and Hispanic populations.

SurveyTrak – A software-based sample administration system. The system was used by interviewers on laptop computers to document their sample assignment, to organize the activities of their workday, to prompt them for appointments to be kept, to record results of each call attempt, to record observations of the sample housing unit, and in all other ways to keep track of their job duties.

Target population – The population to be described by estimates from the survey. In the NSFG the target population is the household population of the United States, which refers to the civilian noninstitutionalized population, plus active-duty military who are not living on military bases. "Noninstitutionalized" refers to the omission of prisons, hospitals, dormitories, and other large residences under central control. College students living in dormitories were interviewed but sampled through their parents'/guardians' households.

Trimming – The process of reducing very large weights for individual cases in the data set. Trimming may be done to reduce the effects of very large individual weights on sample statistics, to reduce disclosure risks from such large weights, and to reduce potential bias in statistics resulting from these very large weights. Trimming occurs during the last stage in the process of creating sampling weights.

Weight – See "Sampling Weight."