

Supporting Statement B for Request for Clearance:

The Continuous

NATIONAL SURVEY OF FAMILY GROWTH, 2009-2012

OMB No. 0920-0314

(expires April 30, 2009)

Contact Information:

William D. Mosher, Ph.D., Statistician
Project Officer, National Survey of Family Growth
National Center for Health Statistics/CDC
3311 Toledo Road, Room 7318
Hyattsville, MD. 20782
301-458-4385
301-458-4034 (fax)
wmosher@cdc.gov
January 8, 2009

**Supporting Statement for Request for Clearance:
NATIONAL SURVEY OF FAMILY GROWTH,
Continuous Interviewing, 2009-2012**

PART B

B. Statistical Methods

1. Respondent Universe and Sampling Methods

Summary—The Continuous National Survey of Family Growth is based on a national probability sample. The Main Study sample is being drawn from 110 Primary Sampling Units (PSUs), but to control costs, in any one year, the sample is drawn from just 33-35 PSUs. Each year, one-fourth of the total sample will be interviewed, and at the end of 4 years, a full 110-PSU design will have been completed. The data will be collected annually and continuously. Each year, about 14,000 households will be contacted, in order to yield the required 5,000 interviews. Each year of data is an independent national sample, but the full sample of 110 PSU's will be completed in 4 years (mid-2006 to mid-2010).

Target Population of the National Survey of Family Growth. The target population of the National Survey of Family Growth is the household population 15-44 years of age. It excludes residents of military bases and institutions (e.g., long-term hospitals, jails, prisons). College students temporarily away from their homes at college are included by sampling them at their home address; they can be interviewed either at home or at college.

Details of the Sample Design.

- (1) 110 Primary Sampling Units (counties, groups of adjacent counties) are selected at random from the entire set of more than 3,100 counties in the US, including Alaska and Hawaii. PSUs are selected with probability proportionate to population size—that is, counties and cities with large populations have a larger chance of selection, and the cities with the largest populations are always included.
- (2) Within each of those 110 areas, smaller areas called sample segments are selected, again at random. A segment is a geographical area (like a group of blocks in cities or an area bounded by roads in a rural area). It can contain as few as 50 structures in a rural area or several hundred in a densely settled urban area.
- (3) Trained staff are sent out to prepare a list of addresses in the segment. (In urban areas, the listers are verifying a commercially purchased address list; in rural areas, they are listing from scratch.)
- (4) Once the addresses are listed, a sample of the addresses is selected (again, by chance) for the study. The sample size is determined by the total desired number of interviews to be taken.

- (5) After an advance letter is sent to each selected household informing them about the study (**Attachment G1**), a trained survey interviewer visits the household, to conduct the screener—that is, to collect a household roster, in order to see if someone lives there who is eligible for the study. If more than one person is 15-44 and eligible, then one person is selected at random for the interview.
- (6) The data collection activities of the National Survey of Family Growth are continuous, but during each year of the survey, a randomly selected subset of the PSUs is used, so each year is a probability sample of the US household population, albeit smaller than the full sample. This “quarter-sample” consists of about 35 primary sampling areas throughout the country. The largest urban area PSUs fall into every year’s sample.
- (7) At the end of each year, about 25 PSUs are dropped from the sample and 25 more are included. At the end of four years, the cumulative sample contains the full set of about 110 PSUs.

The rotating feature of the PSUs permits a cost efficiency of ongoing sampling and data collection operations by using the field interviewing resources in an optimal manner. It further offers at any single year a full national sample for the study, albeit with standard errors of estimates larger than those of the 4-year cumulative sample.

Group quarters with special living arrangements, such as dormitories, institutions, convents, or institutional group homes (for convicts, the frail elderly, or the developmentally disabled, e.g.) may be listed but will not be selected for interviewing, because they are outside the scope of a sample of the household population. Dormitory residents who otherwise live with their parents will be sampled at their parents’ homes. Members of the active duty military who live in civilian housing (not on military bases) will be eligible for the sample.

The NSFG is a personal visit survey. Telephone contacts are permitted only to

- (a) arrange appointments for interviews after the screener has been conducted; and
- (b) for 3-5 minute verification interviews (**Attachment J**) to ensure that the respondent was interviewed.

2. Procedures for the Collection of Information

The sample size targets are as follows:

Sample Size Targets for NSFG Continuous Interviewing
with 2002 (Cycle 6) sample sizes shown for comparison

	<u>Cycle 6</u>	<u>Continuous</u>	<u>Continuous</u>	<u>2002</u>
				<u>2.5</u>
				<u>years*</u> 4
				<u>years**</u>
TOTAL				12,571
				13,000
				20,000
Hispanic				2,712
				2,600
				4,000
Black				2,460
				2,400
				3,700
White & other				7,399
				8,000
				12,300
Male				4,928
				6,000
				9,000
Female	7,643	7,000	11,000	
15-19 years of age				2,271
				2,600
				4,000
20-44 years of age				10,300

10,400
16,000

*2.5 years = June 2006-December 2008.

**4 years = June 2006 – June 2010.

After 2 ½ years of data collection with these sample sizes, the NSFG data set will be over 13,000 interviews, slightly larger than Cycle 6 (the 2002 NSFG). After 4 years, the sample size should be about 20,000 interviews, more than 50 percent larger than in the 2002 NSFG. This sample size will allow estimates for small but important groups such as low-income minority teenagers, couples who have adopted children, childless infertile women 35 years of age and older, gay and lesbian populations, and those who are at risk of HIV because of their sexual behavior.

The current contractor for the NSFG is the University of Michigan's Institute for Social Research (ISR). Under the supervision and monitoring of NCHS, ISR will recruit and train the interviewers for the NSFG and carry out the fieldwork. Contract No. 200-2000-07001 governs this work. Excerpts from the contract that are relevant to the contractor's work in continuous interviewing are shown in **Attachment K**. The main steps in the fieldwork are described below.

Main steps in field work:

- (1) Before contacting households, the contractor will send an advance letter and pamphlet to all eligible households. These explain who is sponsoring the survey, who is conducting it, why it is being done, and the voluntary and confidential nature of the survey. Spanish versions of the questionnaires, the advance letter, and other introductory materials will be prepared as in past Cycles of the survey. The letters and materials are shown in **Attachment G1-G3**.
- (2) When the housing unit is found to be occupied and the interviewer finds someone (18 or older) at home, the screener interview (**Attachment G4**) is conducted. The purpose of the screener is (a) to list the persons living in the household and their ages, and (b) if one or more are 15-44 years of age, to select one. Age and gender are collected in the screener because teenagers and women are selected at somewhat higher rates than adults and men.
- (3) When a person 15-17 years of age is selected for the sample, signed parental consent will be obtained before the interview is conducted. A parent letter and consent form will be used to explain the survey to the minor's mother, father, or guardian, and ask for their written consent. If either the parent or the minor Respondent refuses to give consent, the case is treated as a refusal. (**Attachment G3**).

Emancipated minors--15-17 year-olds who are married or cohabiting and living away from their parents are rare in a sample of this size. Emancipated minors have been excluded from the continuous NSFG because the number of emancipated minors selected

for the NSFG is so small that excluding this group is unlikely to have any noticeable impact on estimates. Using current IRB rules, however, including them would require special procedures that are too complex and too costly for the NSFG.

- (4) If the Respondent is 18 years of age or older, the interviewer gives the Respondent an Adult Consent Form (**Attachment G3**), which explains the survey and requests signed consent. If the Respondent agrees to do the survey but refuses to sign the form, the interviewer can offer to begin the interview, and ask for a signature at the end of the interview, or sign for the respondent.
- (5) The interviewer gives the respondent \$40 “as a token of appreciation.”
- (6) Then the interview is conducted (**Attachments H and I**), using a laptop computer. This use of the computer makes the interviewer's job easier, reduces interviewer errors, protects confidentiality, and produces higher quality data.
- (7) Finally, at the end of the interviewer-administered interview, the interviewer gives the respondent a pair of headphones and the notebook computer, and shows the respondent how to make simple entries on the computer. The respondent then completes a 10-20 minute Audio Computer-Assisted Self-Interview (Audio CASI). The interviewer cannot see or hear what questions the respondent is being asked over the headphones, and cannot see or hear the respondent's answers. Moreover, no one in the household can hear or see either the questions or the answers. This increased privacy has been found to increase the reporting of sensitive behaviors.
- (8) At the end of the Audio CASI section, the interviewer turns off the computer, thanks the respondent and leaves. The interviewer cannot back up and see the respondent's answers, because the Audio CASI system is locked by the respondent when he or she is done.

Quality control

Computer-assisted interviewing improves data quality in several ways:

- (a) Interviewer errors are reduced because interviewers do not have to follow complex routing instructions; the computer does it for them. Most of the missing data in Cycles 1-4 (which were done with paper and pencil) was due to such interviewer errors.
- (b) Respondent errors are also reduced with CAPI interviewing. The contract requires that selected consistency checks be programmed into the questionnaire so that inconsistent answers can be corrected or explained while the interview is still in progress. Our questionnaire work in continuous interviewing is focused in part on identifying and resolving logical inconsistencies more efficiently than in the 2002 NSFG.

- (c) Coding and coding errors are also reduced using CAPI interviewing, and this makes it possible to prepare the data for analysis faster and more accurately. In Continuous Interviewing, earlier cases (e.g., year 1) are being used to discover and correct errors before they affect later cases (e.g., year 2).
- (d) The "Verification" interview is a quality control procedure in which a 10-percent random sample of both respondents and non-respondents will be contacted (usually by telephone) after the interview to verify that the interview was conducted with the appropriate sample person and that a few key items are correct. **(Attachment J)**
- (e) Editing -- Additional computer editing of the data will be performed by the Contractor in the home office after the interviews are complete. NCHS is also performing checks of the quality of the data files, as it has in past cycles.
- (f) Imputation -- In the 1982, 1988, 1995, and 2002 cycles of the NSFG, a few hundred key variables (called "Recodes") were imputed when missing. On most of these items, missing data was less than 1 percent. The imputation procedure is described further in Vital and Health Statistics, Series 2, No. 142, "National Survey of Family Growth Cycle 6: Sample design, weighting, imputation, and variance Estimation," July 2006; see the NSFG web site).

Two basic types of imputation were used for about 400 variables (out of about 6,000 variables on the data file):

- regression model-based imputation (used for most variables)
- logical imputation (for a few variables with only a handful of missing cases).

The imputation procedures used now are very similar to those used in the 2002 NSFG, and the large majority of imputations are being done by multiple regression imputation using the University of Michigan's software called

"IVEWARE."

As in previous cycles, the public use data files will identify imputed values with imputation "flags," so that data users can assess for themselves whether imputation affects the estimates. (Imputation rarely affects estimates in the NSFG because the levels of missing data are generally very low.)

- (g) Estimation -- Estimation refers to the process of producing weighted numbers and percentages for the population from sample data. For each case, a weight is generated which estimates the number of persons in the population that each sampled person represents. For example, if a woman represents 5,000 women in the population, her weight is 5,000. The weight for each respondent is created in 4 basic steps:
- inflation by the reciprocal of the probability of selection,
 - adjustment for nonresponse within age, sex, and race categories, and
 - post-stratification to independent control totals provided by the Census Bureau.
 - trimming of a few dozen extreme weights.

Probabilities of selection vary because black, Hispanic, and teenage respondents are oversampled, and because non-respondents to the survey are sub-sampled for the “double sample” in the non-response follow-up (the last phase of data collection). Adjustments for non-response are made by multivariate (logistic regression) methods. Post-stratification to control totals is done within cells defined by race and origin, age, and sex.

Variances are being estimated using a Taylor Series linearization approach similar to that used in the 2002 NSFG and described in Series 2, No.142 (available on the NSFG web site at www.cdc.gov/nchs/nsfg.htm, under “Reports.”). Codes were generated that allow data users to compute variances using Taylor Series linearization, Balanced Half-Sample Replication, or Jackknife replication methods. A similar procedure will be used to produce the 2006-8 data file.

3. Methods to Maximize Response Rates and Deal with Non-response

Summary.--As discussed above, we use Advance Letters, highly trained interviewers, a web site, 800 numbers at both the University of Michigan and at NCHS, customized follow-up letters to address particular concerns, and special interviewer training on non-response, to encourage cooperation with the survey, and active survey management using daily paradata to allocate interviewer effort.

Our principal guidance in dealing with non-response is our experience in the 2002 NSFG and to date in continuous interviewing, as well as some recent literature on survey non-response—for example:

Robert Groves and Mick Couper. 1998. Nonresponse in Household Interview Surveys. New York: Wiley.

Robert Groves, Eleanor Singer, Amy Corning. 2000. Leverage-Saliency Theory of Survey Participation. Public Opinion Quarterly 64 (3): 299-308, Fall 2000.

R Groves, G Benson, W Mosher, et al. 2005. Design and Operation of Cycle 6 of the National Survey of Family Growth. Vital and Health Statistics, Series 1, No. 42, August, 2005. National Center for Health Statistics, Hyattsville, MD. Available at: <http://www.cdc.gov/nchs/nsfg.htm>.

R Groves and SG Heeringa. 2006. Responsive Design for Household Surveys: tools for actively controlling survey errors and costs. Journal of the Royal Statistical Society A169, Part 3: 439-457, April, 2006.

R. Groves, W Mosher, et al. 2009. Plan and Operation of the Continuous National Survey of Family Growth, 2006-2008. Vital and Health Statistics, Series 1, National Center for Health Statistics. Forthcoming in 2009.

Procedures are listed separately for non-contacts, and for refusals. For non-contacts, the following procedures are used:

- (a) listers of sample segments document units that have access impediments (e.g., locked apartment buildings) and interviewers will schedule calls on such cases earlier in the field period than others,
- (b) observations are made by the interviewer regarding best times to reach the sample household, and
- (c) multiple calls are made on sample units, at different times of the day and different days of the week.

For refusals, interviewers are trained to use methods to avert refusals by responding to the concerns that potential respondents express. Letters on NCHS letterhead, signed by the NCHS Director, are used for all sample households in order to communicate the scientific goals and practical usefulness of the research and to legitimate the visit of the field interviewer. Letters to local community police are also used in some areas to announce the presence of interviewers in the area.

Guidance to interviewers in continuous interviewing is based on the research cited above, and on extensive paradata—data about the fieldwork—collected and recorded by interviewers and other field staff. These data are summarized using logistic regression equations into a total propensity to respond for an entire segment. These data (and case-specific observations entered into the contractor’s sample management system) can be used to guide further actions on individual cases.

Interviewers are in ongoing contact with their supervisors, permitting them to seek guidance on individual problems they encounter. Throughout this process interviewers are explicitly instructed to treat the sample person’s concerns as legitimate questions that deserve thoughtful answers. NCHS IRB rulings do not allow us to follow up emphatic or “hard” refusals.

Incentives.—

Given that even the good survey practices described above are unlikely to attain an 80% response rate with the budget available to the NSFG, OMB clearance is requested to continue to use a \$40 cash incentive (called a “token of appreciation” in all respondent materials) in 2009-12, as explained in Section A9. Previous research (cited below and in **Attachment C**) suggests that, for long, sensitive in-person surveys, incentives do help raise response rates and help to control fieldwork costs when standard good survey practice is not enough. Incentives appear to be especially effective among minorities, teenagers, and low-income people. That observation is consistent with the NSFG’s experience in the 2002 NSFG and so far in 2006-2008. (Our experience with incentives is discussed in detail in **Attachment C**). Given that interviewer labor costs about \$25 an hour (including indirect costs and supervisor time), this \$40 amount is quickly recovered because it saves interviewers time. Examples of the literature that guides our use of incentives is shown below:

E. Singer, “The Use of Incentives to Reduce Nonresponse in Household Surveys,” pages 163-178 in R Groves et al (editors), Survey Nonresponse. Wiley, 2002.

Kulka R. “The Use of Incentives to Survey ‘Hard to Reach’ Respondents,” pages 256-287, In: Federal Committee on Statistical Methodology, Statistical Policy Working Paper No. 23, Volume 2.

Robert M. Groves, Mick P. Couper, Stanley Presser, Eleanor Singer, Roger Tourangeau, Georgina Piani Acosta, and Lindsay Nelson. 2006. “Experiments in Producing Nonresponse Bias,” Public Opinion Quarterly, 2006; 70: 720 - 736.

Davern, Michael; Todd H. Rockwood, Randy Sherrod, and Stephen Campbell. 2003. Prepaid Monetary Incentives and Data Quality in Face-to-Face Interviews: Data from the 1996 Survey of Income and Program Participation Incentive Experiment. Public Opinion Quarterly, Spring 2003; 67: 139 - 147.

Nonresponse Bias Studies Planned

Nonresponse bias studies in NSFG are built into the daily paradata monitoring of the study. NSFG has the following resources:

- 1) a paradata structure that uses lister and interviewer observations of attributes related to response propensity and some key survey variables;

- 2) daily tracking to monitor the sensitivity of key statistics to calling effort ;
- 3) daily monitoring of 12 key socio-demographic domains key to NSFG estimates (age: 15-19 vs. 20-44; sex: male vs. female; and Race/ethnicity: Hispanic, Non-Hispanic Black, and Non-Hispanic white and other);
- 4) randomized responsive design interventions on key auxiliary variables during data collection in order to improve the balance on those variables among respondents and nonrespondents;
- 5) a two-phase sampling scheme, selecting a probability sample of nonrespondents at the end of week 10 of each quarter and using increased incentives and other procedures to increase response.
- 6) Comparison of alternative postsurvey adjustments for nonresponse.

Nonresponse bias analysis was pre-specified by the responsive design features of NSFG, continuous interviewing. While some have been conducted, they do not represent the final studies, which await fully adjusted estimates that will be produced with the first data release in 2009. A more complete description of these activities appears in **Attachment M**.

4. Tests of Procedures or Methods

The first several weeks of interviewing in 2006 served as the pretest for continuous interviewing. Another pretest is not necessary until a new contractor is hired or questionnaire changes are made (currently scheduled for early in calendar year 2011). The first 600 interviews (about 6 weeks) in 2011 will serve as the pretest. If these initial interviews go well, we will keep the interviews and proceed. We think there is about a 90 percent chance that no major problems will be discovered (no significant problems were discovered in 2006). If significant problems in the questionnaire programs or other materials are discovered, they will be fixed, and interviewing may have to be halted for a few weeks while the fixes are made and sent to the interviewers. We made some questionnaire changes in Year 2 (2007) and 3 (2008). No changes will be made for 2009 or 2010.

5. Statistical Consultants

The statistical consultant (on sample design, variance estimation, and statistical methods) for NCHS is:

Karen E. Davis, M.A., Mathematical Statistician
NCHS Office of Research and Methodology
301-458-4635
e-mail: KEDavis@cdc.gov

The sample selection and data collection are supervised for NCHS by:

William D. Mosher, Ph.D.
Project Officer, NSFG
NCHS, Room 7421
3311 Toledo Road
Hyattsville, MD 20782
301-458-4385
e-mail: wmosher@cdc.gov

Sample selection and data collection are supervised for the contractor by:

Robert M. Groves, Ph.D.,
Project Director, NSFG, and
Director, Survey Research Center,
University of Michigan
426 Thompson St, Ann Arbor, MI 48104
734-647-6078
bgroves@isr.umich.edu

The person responsible for the analysis of the survey is:

William D. Mosher, Ph.D, Project Officer for NCHS:
Phone: 301-458-4385
e-mail: wmosher@cdc.gov
FAX: 301-458-4034

LIST OF ATTACHMENTS

A. Authorizing legislation

- A1. NSFG Authorizing Legislation
- A2. Office of Family Planning Authorizing Legislation
- A3. Adolescent Family Life Authorizing legislation
- A4. NICHD Authorizing legislation
- A5. CDC's Division of HIV/AIDS Prevention (DHAP) Authorizing Legislation
- A6. Children's Bureau (ACF) Authorizing Legislation

B. Justifications for Questions in the Survey

- B1. Justification of the Female questionnaire by topic
- B2. Justification of the Male questionnaire by topic

C. Memorandum to OMB on the Use of Incentives in the NSFG, Nov 19, 2007.

D. Partial list of publications from the Survey

- D1. List of publications from the 2002 NSFG by date published
- D2. List of publications from the 2002 NSFG by topic.

E. Memoranda from other offices and agencies on their use of the NSFG

- E1. NCHS Public Affairs Officer
- E2. Healthy People 2010 Health Objectives on Family Planning, HIV, STDs
- E3. Office of Population Affairs.
- E4. NICHD, NIH
- E5. Children's Bureau, ACF, DHHS
- E6. OASPE (Office of the Assistant Secretary for Planning and Evaluation)
- E7. Division of HIV/AIDS Prevention, CDC
- E8. Division of Reproductive Health, CDC
- E9. Division of Sexually Transmitted Disease Prevention, CDC

F. Consultation outside the agency:

- F1. A Summary of Options for the Design of Cycle 7 of the National Survey of Family Growth, March, 2004.
- F2. Brief Summary of the 2006 Research Conference on the NSFG.
- F3. Brief Summary of Meetings with NSFG Funders in 2007
- F4. Agenda for the 2008 Research Conference on the NSFG.
- F5. Draft Agenda for 2008 Meeting of the NSFG advisory workshop
- F6. 60-Day Notice for the National Survey of Family Growth, 2008

- G. Respondent Materials for the NSFG in 2009-2012**
 - G1—Respondent Letters
 - G2---Brochures and Letter of Authorization.
 - a. Program Brochure
 - b. Confidentiality Brochure
 - c. Introductory letter
 - G3---Consent Forms
 - G4—Screener Questionnaire
- H. FEMALE Questionnaire, 2008-2010 (“Years 3 and 4”)**
- I. MALE Questionnaire, 2008-2010 (“Years 3 and 4”)**
- J. Verification Questionnaire**
- K. Scope of Work for NSFG**
- L. IRB Approval Forms for the NSFG**
- M. Non-Response Bias Analyses for the continuous NSFG**
- N. Testing New Questions**