

SF-83-1 SUPPORTING STATEMENT

for

2010

National Survey of College Graduates

CONTENTS

Section	Page
A. JUSTIFICATION.....	1
1. NECESSITY FOR INFORMATION COLLECTION.....	2
2. USES OF INFORMATION.....	3
3. CONSIDERATION OF USING IMPROVED TECHNOLOGY.....	5
4. EFFORTS TO IDENTIFY DUPLICATION.....	6
5. EFFORTS TO MINIMIZE BURDEN ON SMALL BUSINESS.....	6
6. CONSEQUENCES OF LESS FREQUENT DATA COLLECTION.....	6
7. SPECIAL CIRCUMSTANCES.....	7
8. FEDERAL REGISTER ANNOUNCEMENT AND CONSULTATION OUTSIDE THE AGENCY.....	7
9. PAYMENT OR GIFTS TO RESPONDENTS.....	9
10. ASSURANCE OF CONFIDENTIALITY.....	9
11. JUSTIFICATION FOR SENSITIVE QUESTIONS.....	10
12. ESTIMATE OF RESPONDENT BURDEN.....	10
13. COST BURDEN TO RESPONDENTS.....	10
14. COST BURDEN TO FEDERAL GOVERNMENT.....	10
15. REASON FOR CHANGE IN BURDEN.....	11
16. SCHEDULE FOR INFORMATION COLLECTION AND PUBLICATION.....	11
17. DISPLAY OF OMB EXPIRATION DATE.....	11
18. EXCEPTION TO THE CERTIFICATION STATEMENT.....	11
B. COLLECTION OF INFORMATION EMPLOYING STATISTICAL METHODS.....	12
1. RESPONDENT UNIVERSE AND SAMPLING METHODS.....	12
2. SURVEY METHODOLOGY.....	13
3. METHODS TO MAXIMIZE RESPONSE.....	14
4. TESTING OF PROCEDURES.....	15
5. CONTACTS FOR STATISTICAL ASPECTS OF DATA COLLECTION.....	18
APPENDIX A: NSF Act of 1950	
APPENDIX B: First Federal Register Announcement	
APPENDIX C: 2010 NSCG Coefficient of Variation (CV) Target	
APPENDIX D: NSCG Questionnaire Changes	
APPENDIX E: Survey Mailing Materials	
APPENDIX F: Draft 2010 NSCG Questionnaire	

**2010 NATIONAL SURVEY OF COLLEGE GRADUATES
SUPPORTING STATEMENT**

A. JUSTIFICATION

This request is for a three-year revision of the previously approved OMB clearance for the National Survey of College Graduates (NSCG). The NSCG was last conducted in 2008. The OMB clearance for the 2008 NSCG expires July 2011.

The NSCG is one of three principal surveys that provide data for the National Science Foundation's (NSF) Scientists and Engineers Statistical Data System (SESTAT). The purpose of the SESTAT database is to provide information on the entire U.S. population of scientists and engineers with at least a bachelor's degree. SESTAT is produced by combining data from the Survey of Doctorate Recipients (SDR; representing persons in the general U.S. population who have earned a doctorate in science, engineering, or health (SEH) from a U.S. institution), the National Survey of Recent College Graduates (NSRCG; representing persons with a recently earned bachelor's or master's degree in SEH from a U.S. institution) and the NSCG (representing all individuals in the U.S. who had a bachelor's degree or higher in a SEH or SEH-related degree before January 1, 2009, or those who had a bachelor's degree or higher in some other field but had an SEH or SEH-related occupation, including individuals who received degrees only from foreign institutions).

The SESTAT integrated database derived from these surveys represents the demographic, educational, and employment characteristics of college-educated scientists and engineers in the United States. All three of these surveys are usually conducted every two years. The primary purpose of the NSCG is to provide information on the U.S. stock of scientists and engineers. The panel portion of the SDR also provides information on the stock, while the new sample in the SDR and the entire NSRCG provide important data on the new graduates with SEH degrees entering the labor force. The NSCG constitutes the bulk of the records in the SESTAT database; accounting for approximately 54% of the records in the SESTAT system and slightly over 88% of the population estimate in 2008.

The SESTAT integrated database is the only available source that provides detailed information to support a wide variety of policy and research analyses on science and engineering (S&E)¹ workforce and personnel. To provide complete representation of U.S. S&E workforce at all degree levels, SESTAT was designed as a unified database that integrates information from all three component surveys. The system of surveys, created for the 1993 survey cycle and developed throughout the past two decades, is closely based on the recommendations of the

¹ S&E workforce includes the individuals with degrees or occupations in computer and mathematical sciences, life sciences, physical sciences, social sciences, engineering, and health sciences.

National Research Council's Committee on National Statistics (CNSTAT) report to NSF.² That report recommended a data collection design based on three surveys, of which one (the NSCG) would be linked to the decennial Census.

Below are summary of changes in the survey methodology in 2010 from previous year:

1) Sampling Frame Change

- The 2010 NSCG new cohort sampling frame has changed from decennial census to the American Community Survey (ACS). In the past, a new cohort sample was drawn from the census long form after each decennial census. The first NSCG survey in the decade has been much more expensive and burdensome than the subsequent year surveys because of the larger sample size required to identify the science, engineering and health (SEH) personnel in the U.S. The sample size required for the 2010 survey (100,000) is significantly less than in the previous decade baseline surveys (numbers for 1993 and 2003) because of the field of bachelor's degree data collected in the 2009 ACS allows for much more efficient sampling that targets the SEH personnel.
- Only half of the 2008 NSCG sample will be replaced with cases drawn from the 2009 ACS. The annual sample size of the ACS is too small to draw the entire NSCG sample from a single year of ACS. The 2012 NSCG sample will be the first NSCG sample to consist entirely of cases drawn from the ACS with half of the 2012 sample drawn from the 2011 ACS. Beginning in 2014, for each NSCG cycle 1/4th of the NSCG sample from the previous NSCG cycle will be rotated out and replaced with new cases drawn from the ACS in the preceding year.
- The 2010 NSCG sample consists of two different sample groups with an old cohort sample based on the respondents from 2008 NSCG and 2008 National Survey of Recent College Graduates (NSRCG), and a new cohort sample based on the respondents from 2009 ACS.

2) New Web Response Mode

For the first time, the 2010 NSCG will employ a web mode in data collection, in addition to a mail questionnaire and Computer Assisted Telephone Interview (CATI). Sample cases will be offered a choice for responding by web or mail. Offering more than one response choice may help obtain an increased number of responses. Younger and recent graduates are highly web-literate, so offering a web response option is apt to be appealing to such respondents.

Usability testing was conducted at the U.S. Census Bureau cognitive lab by the Statistical Research Division staff in the fall of 2009 and in early 2010 to examine the user interface of the NSCG web survey. Eye tracking software was used to test the long list collection, and the placement of navigational buttons. Preliminary results indicated that the users were highly satisfied with the usability of the web instrument. The 2010 NSCG web survey design

² National Research Council. Committee on National Statistics. *Surveying the Nation's Scientists and Engineers: A Data System for the 1990s*. Washington: National Academy Press. 1989.

will be based on the results of these studies and coordinated with the web survey instruments used in other two SESTAT surveys.

1. NECESSITY FOR INFORMATION COLLECTION

The National Science Foundation Act of 1950, as amended by Title 42, United States Code, Section 1862 requires the National Science Foundation to:

“Provide a central clearinghouse for the collection, interpretation, and analysis of data on scientific and engineering resources and to provide a source of information for policy formulation by other agencies of the Federal Government...” (See Appendix A)

In meeting its responsibilities under the NSF Act, the Foundation relied on the National Register of Scientific and Technical Personnel from 1954 through 1970 to provide names, location, and characteristics of U.S. scientists and engineers. Acting in response to a Fiscal Year 1970 request of the House of Representatives Committee on Science and Astronautics (see U.S. Congress, House of Representatives, 91st Congress, 1st Session, *Report No. 91-288*), NSF, in cooperation with the Office of Management and Budget and eight other agencies, undertook a study of alternative methods of acquiring personnel data on individual scientists and engineers.

The President's budget for Fiscal Year 1972, as submitted to the Congress, recommended the "discontinuation of the National Register of Scientific and Technical Personnel in its present form" and that funds be appropriated "to allow for the development of alternative mechanisms for obtaining required information on scientists and engineers." The House of Representatives Committee on Science and Astronautics in its report on Authorizations for Fiscal Year 1972 states that "...it has no objection to this recommendation...." (See U.S. Congress, House of Representatives, 92nd Congress, and 1st Session, *Report No. 92-204*).

Subsequently, the NSF established and continues to maintain the SESTAT system of surveys, the successor to the Scientific and Technical Personnel Data System of the 1980s, which was the successor to the National Register. The Science and Technology Equal Opportunities Act of 1980 directs NSF to provide to Congress and the Executive Branch an “accounting and comparison by sex, race, and ethnic group and by discipline, of the participation of women and men in scientific and engineering positions.” The SESTAT database, of which NSCG is the large majority of records, provides much of the information to meet this mandate.

The longitudinal data from the NSCG provides valuable information on careers, training, and educational development of the Nation’s highly educated science and engineering population. These data enable government agencies to assess the scientific and engineering resources available in the U.S. to business, industry, and academia, and to provide a basis for the formulation of the Nation's science and engineering policies. Educational institutions use the NSCG data in establishing and modifying scientific and technical curricula, while various industries use the information to develop recruitment and remuneration policies.

The NSF uses the information to prepare congressionally mandated biennial reports such as *Women, Minorities and Persons with Disabilities in Science and Engineering* and *Science and Engineering Indicators*. These reports enable NSF to fulfill the legislative requirement to act as a clearinghouse for current information on the S&E workforce.

The Committee for Equal Opportunity in Science and Engineering (CEOSE), an advisory committee to the NSF and other government agencies, established under 42 U.S.C. §1885c, has been charged by the U.S. Congress with advising NSF in assuring that all individuals are empowered and enabled to participate fully in science, mathematics, engineering and technology. Every two years CEOSE prepares a congressionally mandated report that makes extensive use of the SESTAT data to highlight key areas of concerns relating to students, educators and technical professionals.

The importance of information on the scientific and technical workforce to inform public policy can be seen in discussions of the National Science Board's Task Force on National Workforce Policies for Science and Engineering. The taskforce relied heavily on SESTAT data to inform its deliberations about the S&E workforce and SESTAT data were an integral part of the taskforce's final report. (See <http://nsf.gov/nsb/documents/2003/nsb0369>.)

2. USES OF INFORMATION

Researchers, policymakers and other users of the data use information from the SESTAT database to answer questions about the number, employment, education, and characteristics of the S&E workforce. Because it provides up-to-date and nationally representative data, researchers and policymakers use the database to address questions on topics such as the role of foreign-born or foreign-degreed scientists and engineers, the transition from higher education to the workforce, the role and importance of postdocs, diversity in both education and employment, the implications of an aging cohort of scientists and engineers as baby boomers reach retirement age, and information on long-term trends in the S&E workforce.

Data from NSF's SESTAT component surveys are used in policy discussions of the executive and legislative branches of Government, the National Science Board, NSF management, the National Academy of Sciences, professional associations, and other private and public organizations. Some recent specific examples of the use of the SESTAT data are: the American Institutes of Research used the SESTAT data in the evaluation of the NSF's Alliance for Graduate Education and the Professoriate Program; the Commission on Professionals in Science and Technology regularly publishes data from SESTAT in their STEM Trends publications; the General Accounting Office used the SESTAT data to issue a report on education and disability. The Federal Reserve Bank of St. Louis used the SESTAT data to examine the pathway from Community College to a Bachelor's Degree and Beyond; and many Ph.D. students use the SESTAT workforce data in dissertations.

Data Dissemination and Access

The NSF makes the data from the SESTAT system of surveys available through published reports, the SESTAT on-line data system, public use files and restricted licenses. The 1993 and 2003 NSCG data are available as public-use files. The NSCG panel data from all the 1990s and 2000s cycles are also available as a component of the SESTAT database for each survey year (1993, 1995, 1997, 1999, 2003, 2006 and 2008), which are available as SESTAT public-use files.

The SESTAT data were used extensively in the latest versions of the congressionally mandated biennial reports *Science and Engineering Indicators, 2010* and *Women, Minorities and Persons with Disabilities in Science and Engineering, 2009*.

NSF also used the NSCG and SESTAT integrated data in recent reports such as:

- ***2003 College Graduates in the U.S. Workforce: A Profile, December 2005***
- ***What Do People Do After Earning a Science and Engineering Bachelor's Degree? June 2006***
- *Why Did They Come to the United States? A Profile of Immigrant Scientists and Engineers, June 2007*
- *Unemployment Rate of U.S. Scientists and Engineers Drops to Record Low 2.5% in 2006, March 2008*

All NSF Publications can be accessed on the SRS website at <http://www.nsf.gov/statistics>.

To provide better accessibility to information for policy makers and researchers, NSF provides the SESTAT integrated database and the NSCG data on the internet. The SESTAT on-line system allows internet users to create customized data tabulations with a user-specified subject area. Additionally, the NSCG and SESTAT public-use files are available for download through the SESTAT web page at <http://www.nsf.gov/statistics/sestat>.

Results from the SESTAT integrated database and NSCG data are routinely presented at the conferences and professional meetings, such as the annual meetings of the Association for Institutional Research or the American Educational Research Association.

Since 2005, NSF has distributed over 200 files of the more than decade-old 1993 public-use NSCG data set to researchers in government, academia, and professional societies. Close to 700 copies of the 2003 NSCG public-use files have been requested since 2005. In spite of the age of the data, the 1993 and 2003 NSCG data continue to be heavily used because they are the only data sets analysts can use to compare the S&E workforce to the general population of college degree holders in the U.S. Besides capturing people with degrees earned at U.S. institutions, the NSCG between 1993 and 2008 included college degree holders who earned their degrees outside of the United States and who were residing here at the time of the previous census.

There are approximately 20 licensed users for the 2003 and 2006 SESTAT integrated database micro data files under a licensing agreement with SRS. Close to 800 users have downloaded the SESTAT public-use files since 2005. As previously noted, over half of the records in the SESTAT file come from the NSCG.

Some of the research from the public use NSCG data and the SESTAT restricted data licensees resulted in papers such as:

- *Why Do Women Leave Science and Engineering?* National Bureau of Economic Research, 2010
- *Functional Impairment and the Choice of College Major*, University of South Florida, 2010
- *How Much Does Immigration Boost Innovation?* McGill University, 2010
- *Increasing Time to Baccalaureate Degree in the United States*, National Bureau of Economic Research, 2010
- *Higher Education and Disability: Education Needs a Coordinated Approach to Improve Its Assistance to Schools in Supporting Students*, GAO Report, 2009
- *Diversifying Science and Engineering Faculties: Intersections of Race, Ethnicity, and Gender*, Georgia Institute of Technology, 2010
- *Earnings of a Lifetime: Comparing Women and Men with College and Graduate Degrees*, Indiana University Kelley School of Business, 2009
- *Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors*, Harvard University, 2009
- *The Small Firm Effect and the Entrepreneurial Spawning of Scientists and Engineers*, Washington University St Louis, 2009
- *From Community College to a Bachelor's Degree and Beyond: How Smooth Is the Road?* Federal Reserve Bank of St. Louis, 2009
- *Double Your Major, Double Your Return?* St. Lawrence University, 2008
- *Gender Wage Disparities Among the Highly Educated*, University of Chicago, 2008

3. CONSIDERATION OF USING IMPROVED TECHNOLOGY

For the first time, the 2010 NSCG will employ a web mode in the survey collection. The U.S. Census Bureau will collect the NSCG data, under an interagency agreement, using a multi-mode approach, with a paper questionnaire mailed to sample persons along with an option for completing the survey on the Internet, and the nonrespondents will be followed up using computer-assisted telephone interviewing (CATI).

Because the sample contact information will be at least a year old for most sample members by the time the survey is conducted, extra effort will have to be spent to locate respondents. To do this in the most efficient way, the NSCG will employ nonintrusive locating procedures to find valid mailing addresses for cases that are identified as non-mailable after the sample is sent through automated software to check against updates to the National Change of Address (NCOA) database. These nonintrusive procedures include the use of Internet search engines, and name and address locating software such as FastData and InfoUSA. Additionally, the Census Bureau has developed an electronic locating system to improve the efficiency of the locating operation.

The NSCG web survey will be developed using Census Centurion, which is a secure web-based application programmed to meet the stringent Census Bureau data security requirements. The web survey will take advantage of the computer-assisted interviewing system that allows for probes for invalid or inconsistent responses, and missing responses to a few question items critical for a complete interview.

The NSCG will use Census' state-of-the-art keying system, developed in the Visual Basics software, to capture mail questionnaire data, which should decrease the time necessary for this operation and increase accuracy. The telephone-interviewing phase will utilize a variety of improved technologies. Interviews will be conducted using the computer-assisted telephone interviewing (CATI) system. Help screens will be displayed with additional instructions or probes at any given point of data collection. The CATI instrument will be programmed in Blaise. Case management for the telephone interviews will utilize the Census Bureau's advanced WebCATI system. This system allows case workload to be balanced across the Census Bureau's three telephone centers and can assign cases to interviewers based on a variety of skills (e.g. language, refusal conversion expertise). As a result, cases will be handled in a more efficient and effective manner.

Optical scanning will be used to capture the digital images of the mail questionnaire after keying. The images will be stored in a database that is accessible to survey staff at their desktops. This will facilitate easier retrieval of the actual response for use during the data collection of missing critical items, data reconciliation, and editing stages.

4. EFFORTS TO IDENTIFY DUPLICATION

Duplication, in the sense of similar data collections, does not exist. No other data collection captures all components of scientists and engineers in the United States. There is no similar information available other than from this survey, conducted by the U.S. Census Bureau for NSF since the 1960s. Data from the Current Population Survey provides occupational estimates but does not collect information on degree field for higher education degrees. The American Community Survey, starting in 2009, collects the field of bachelor's degrees but does not collect detailed information on education history, work activities, and employment characteristics as the NSCG does, nor is the ACS longitudinal in nature.

5. EFFORTS TO MINIMIZE BURDEN ON SMALL BUSINESS

Not applicable. The NSCG collects information from individuals only.

6. CONSEQUENCES OF LESS FREQUENT DATA COLLECTION

Because NSCG is a panel survey, conducting the survey less frequently would make it more difficult and costly to locate the persons in the sample because of the mobility of the U. S. population. The results would be a higher attrition rate and less reliable estimates. Also, government, business, industry, and universities would have less recent data to use as a basis for formulating the Nation's science and engineering policies.

Expanding the time between interviews would also lessen the accuracy of the recall of information by the respondents. This would affect the reliability of the data collected and reduce the quality of the data for all uses, including the congressionally mandated biennial reports prepared by the NSF.

Follow-up surveys every two years on the same sampled persons are also necessary to track changes in the science and engineering workforce as there are large movements of individuals into and out of science and engineering occupations over both business and life cycles. To make sure of the availability of current national S&E workforce data, the NSCG is conducted and coordinated with the NSRCG and the SDR. The degradation of any single component would jeopardize the integrity and value of the entire SESTAT system of surveys and integrated database.

7. SPECIAL CIRCUMSTANCES

Not applicable. This data collection does not require any one of the reporting requirements listed.

8. FEDERAL REGISTER ANNOUNCEMENT AND CONSULTATION OUTSIDE THE AGENCY

Federal Register Announcement

The Federal Register announcement for the NSCG appeared on February 24, 2010 (See Appendix B.) NSF received no public comment in response to the announcement as of the close date of April 26, 2010.

Consultations Outside the Agency

The Division of Science Resources Statistics (SRS) within the NSF has responsibility for the SESTAT surveys. In the early 1990s, SRS initiated and implemented a major redesign of this system of surveys, and continued to adhere closely to the redesigned approaches in conduct of the surveys throughout the past two decades.

As the SESTAT survey system entered the 21st century, SRS set a goal to further improve the efficiency and relevancy of the SESTAT system in meeting the data needs of policy makers, academic and research communities and industry. To accomplish this goal, SRS carefully planned and engaged in a series of formal and informal evaluations and assessments of each of the three surveys as well as the system as a whole between May 1999 and December 2002. These evaluations covered several areas: sampling frame, population coverage, sample design, survey content, data system design and data dissemination.

After the redesign efforts, SRS began a more systematic set of activities to encourage greater dissemination of the SESTAT surveys, and to encourage greater use of the data by outside researchers.

Meetings and Workshops

Both internal and external consultation has continued to take place through a series of meetings and workshops on various issues related to the SESTAT redesign and survey methodology since 2008.

For the 2010 survey round:

- SRS worked with the U.S. Census Bureau, OMB, and other Federal agencies to add a field of degree (FOD) question to the American Community Survey, to enable more precise sampling for future NSCG surveys. As a part of this activity, SRS worked with the Census Bureau on a methods test to test various versions of a FOD question.
- SRS commissioned the Committee on National Statistics (CNSTAT) of the National Research Council (NRC) to examine proposed sample design options for the NSCG, which will now be based on the American Community Survey (ACS), as opposed to the long form of the Decennial Census. The CNSTAT committee held a two-day workshop on this topic, and issued a report with recommendations to NSF on the 2010 and beyond NSCG sample design. The recommendations formed the basis for the 2010 NSCG design.³
- SRS coordinated with OMB on wording for the collection of data on functional disability question items in the SESTAT surveys to increase consistency across the Federal statistical agencies in surveys with such questions. As a result, a new category on cognitive disability, taken from the ACS, was added to all three SESTAT surveys in 2010, and the introductory sentence was revised to refer to difficulties with specific functional limitations.

³ National Research Council, Committee on National Statistics. 2008. *Using the American Community Survey for the National Science Foundation's Science and Engineering Workforce Statistics Programs*. Washington: The National Academies Press.

Consultations for Outreach and Dissemination

In order to maintain the currency of the SESTAT surveys and to obtain ongoing input from the public and researchers, SRS has engaged in the following activities.

For the 2006, 2008, and 2010 survey rounds:

- SRS has convened a Human Resources Experts Panel (HREP) in order to help improve data collection on the S&E workforce through review and renewal of the S&E personnel surveys and to promote use of the data for research and policy analysis purposes. HREP accomplished its mission by: 1) Suggesting methods to publicize and promote the data; 2) Providing advice on efforts to improve the timeliness and accuracy of S&E labor force data; 3) Providing a mechanism for obtaining ongoing input from both researchers and policy analysts interested in S&E personnel data; 4) Providing perspectives on the data needs of decision makers; 5) Identifying issues and trends that are important for maintaining the relevance of the data; 6) Identifying ways in which S&E personnel data could be more useful and relevant for analyses; and 7) Proposing ways to enhance the content of the SRS human resources surveys. The panel was made up of 15 members who represented the sciences, academia, business/industry, government, researchers and policy makers. Five meetings have been held since the panel was convened in 2007.
- In addition to researchers and the public who use the public-use SESTAT, SDR, NSRCG or NSCG files, there are also individuals who use the restricted-use files under a license. SRS has funded three workshops where current and potential future licensees met at NSF to present their research findings to NSF as well as to the broader research community.
- The SESTAT surveys contain a wealth of information on highly trained individuals in the U.S. labor force. Over the past several years, there has been a great deal of interest in leveraging the survey data that are collected with other information on productivity by some of the same individuals (for example, patenting records or publishing records). In order to pursue the feasibility of this approach, SRS funded a workshop at NSF that brought in experts on database matching. SRS is currently engaged in an activity that will enable the matching of some SESTAT data to various patent and publication databases.
- Through a grant to the Association for Institutional Research (AIR), SRS staff recorded two webinars on the SESTAT website and data tool to encourage broader use of the data.
- ASA/AAPOR invited an SRS analyst to present a webinar on science and technology human resources surveys, data and indicators; the SESTAT data are the source for all of the major indicators and trends on this workforce.

9. PAYMENT OR GIFTS TO RESPONDENTS

NSF is considering a monetary incentive experiment to examine potential nonresponse bias. An incentive plan will be submitted for OMB's review and approval at later time if NSF decides to conduct an incentive experiment.

10. ASSURANCE OF CONFIDENTIALITY

NSF and the Census Bureau are committed to protecting the **confidentiality** of all survey respondents. The NSCG data will be collected in conformance with the Privacy Act of 1974, the NSF Act of 1950, as amended, and the Confidential Information Protection and Statistical Efficiency Act (CIPSEA) of 2002. The Census Bureau is conducting the NSCG under the authority of Title 13 and 15, United States Code, Section 8 and 1525, respectively.

As explained in Section B.1, there are three components of the 2010 NSCG sample design. The first one is 2003 NSCG respondents from the 2000 decennial census, the second is “NSRCG panel” respondents subsampled from the 2001, 2003, 2006 and 2008 NSRCGs, and the third is based on respondents to the 2009 American Community Survey (ACS).

The statement on the questionnaire cover will cite the appropriate data collection authority as the NSF Act and confidentiality assurances under the CIPSEA. The questionnaire cover statement will also inform the respondents that the data will be used for statistical purposes only, and the voluntary nature of their response. Two different cover letters will be used where the content will be the same except that for the sample members from the 2003 NSCG and the 2009 ACS, the cover letters will include additional statements about the Census Bureau’s Title 13 as the data collection authority and assurances of confidentiality (see Appendix E). The Census Bureau will include the same appropriate notices of confidentiality and the voluntary basis of the survey in the introduction to respondents contacted during the CATI phase of the data collection.

NSF and the Census Bureau will operate within the guidelines established by the Privacy Act to protect respondents’ privacy and the confidentiality of the data collected. The Privacy Act states “microdata files prepared for purposes of research and analysis are purged of personal identifiers and are subject to procedural safeguards to assure anonymity.”

The Census Bureau has demonstrated experience in handling sensitive data. Routine procedures will be in place to ensure data confidentiality, including the use of passwords and encrypted identifiers to prevent direct or indirect disclosures of information. Furthermore, the Census Bureau’s management system is in full compliance with the government’s ADP systems requirements.

11. JUSTIFICATION FOR SENSITIVE QUESTIONS

No questions of a sensitive nature are asked in this data collection.

12. ESTIMATE OF RESPONDENT BURDEN

The NSF estimates that it will contact approximately 100,000 sample persons by mail or computer-assisted interviewing. Based on experience administering the NSCG interviews, the

questionnaire takes an average of 25 minutes to complete. An overall response rate of about 80 percent is estimated from 65,000 new cohort sample, and about 90 percent from 35,000 old cohort sample. Based on an estimate of approximately 83,500 completed cases, the total burden hours for the 2010 NSCG data collection are 34,792. The total cost to respondents for the 34,792 burden hours is estimated to be \$1,154,040. This estimate is based on an estimated median annual salary of \$69,000 per NSCG respondent. Assuming a 40-hour workweek and a 52-week salary, this annual salary translates to an hourly salary of \$33.17. Salary estimates were obtained using data from the 2008 NSCG.

13. COST BURDEN TO RESPONDENTS

Not applicable. This survey does not require respondents to purchase equipment, software or contract out services.

14. COST BURDEN TO FEDERAL GOVERNMENT

The total estimated cost to the Government for the 2010 NSCG is approximately \$16.3 million, which includes survey data collection costs, and NSF staff costs to provide oversight and coordination with the other two SESTAT surveys. The cost estimate for data collection is \$15.7 million, which is based on sample size; length of questionnaire; administration; overhead; sample design; mailing; printing; sample person locating, telephone interviewing; incentive payments, critical items data retrieval, data keying and editing; data quality control; imputation for missing item responses; weighting and estimating sampling error; file preparation and delivery; and preparation of documentation and final reports. The NSF staff costs are estimated at \$562,500 (based on \$150,000 annual salary of 1.5 FTE for 2.5 years).

15. REASON FOR CHANGE IN BURDEN

In the past, after each decennial census, a new sample was drawn from the census long form, and that sample followed until the end of the decade. This was done in 1982, 1993 and 2003. The first survey in the decade has been much more expensive and burdensome than the following ones because of the larger sample size required to identify the S&E personnel in the U.S.

The sample size required for the 2010 survey is significantly less than the initial survey (2003) in the previous decade because of the field of bachelor's degree item added to the ACS in 2009 allows for much more efficient sampling that targets S&E personnel. Therefore, the 2010 NSCG burden hours of 34,792 is significantly less than the 74,877 hours estimated for the 2003 survey, which had a sample size of 217,000. The 2010 NSCG burden hours are actually less than the 2008 burden hours of 35,500 with a sample size of 68,000 and a response rate of 90%.

The change in burden from the 2008 survey is because of lower expected overall response rate in 2010 NSCG than 2008. Two-thirds of the 2010 sample will consist of the new cohort sample from the 2009 ACS frame for whom a lower response rate is expected. Sample members

selected from the ACS will be unfamiliar with the survey, and their locating information will not be as complete as the old cohort sample.

16. SCHEDULE FOR INFORMATION COLLECTION AND PUBLICATION

NSF does not plan to use any complex analytical techniques in NSF publications using this data. Normally cross tabulations of the data are presented in NSF reports and other data releases.

The time schedule for 2010 data collection and publication is currently estimated as follows:

Data Collection	October 2010 - June 2011
Coding and Data Editing	December 2010 – September 2011
Final Edited/Weighted/Imputed Data File	December 2011
SESTAT Info Brief	Late Spring 2012
SESTAT Detailed Statistical Tables	Summer 2012
SESTAT Integrated Public Use Data File	Summer/Fall 2012

17. DISPLAY OF OMB EXPIRATION DATE

The OMB Expiration Date will be displayed on the 2010 NSCG questionnaires.

18. EXCEPTION TO THE CERTIFICATION STATEMENT

Not Applicable.