SF-83-1 SUPPORTING STATEMENT

For

2010 Survey of Doctorate Recipients

TABLE OF CONTENTS

LIST OF ATTACHMENTS

2010 SURVEY OF DOCTORATE RECIPIENTS

Supporting Statement

A. Justification

This request is for a three-year revision of the previously approved OMB clearance for the Survey of Doctorate Recipients (SDR). The SDR was last conducted in 2008 and the OMB clearance for the 2008 SDR expires July 31, 2011 (OMB No 3145-0020).

The SDR 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 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 (representing persons with a recently earned bachelor's or master's degree in SEH from a U.S. institution) and the National Survey of College Graduates (representing all individuals in the U.S. with a bachelor's or higher degree in a SEH or SEH-related field, or those with a bachelor's or higher degree in some other field, but have an SEH or SEH-related occupation, including individuals who received degrees only from foreign institutions).

The SESTAT integrated database derived from these surveys contains data on 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 SDR is to provide information on doctoral scientists and engineers who were awarded degrees from U.S. institutions and reside in the U.S. It is comprised of two components: 1) a longitudinal panel that tracks doctorate recipients throughout their careers until age 75, and 2) a new cohort component that adds new doctorate recipients after they receive their degree. The panel portion of the SDR provides information on the experienced stock of doctorate recipients, while the new sample in the SDR provides important data on the early career experiences of new doctorate recipients with SEH degrees entering the labor force.

In addition, since 2003 and continuing with the 2006 and 2008 SDR, the NSF tested and reaffirmed the feasibility of developing a complimentary international panel study of U.S. trained doctorate recipients. This sub-sample was comprised primarily of non-U.S. citizens who emigrated after degree award. In 2010, U.S. citizens found living abroad will be surveyed. The 2010 SDR will represent both a National Survey of Doctorate Recipients (NSDR) to be included in the SESTAT, and a smaller International Survey of Doctorate Recipients (ISDR), which will include U.S. citizens as well as non-citizens living outside the U.S. Currently, 33% of U.S. SEH doctorates are awarded to temporary visa holders and nearly 25% of them plan to leave the U.S. upon graduation. The 2010 ISDR will yield new information about the educational and demographic characteristics of U.S. trained SEH doctorate recipients living and working abroad on the reference date, 1 October 2010.

The SDR, as part of the SESTAT data system, is the only available source that provides detailed information at the doctorate level to support a wide variety of policy and research analyses on science, engineering and health (SEH) labor force issues. To provide complete representation of U.S. scientists and engineers 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 1990s, is closely based on the recommendations of the National

Research Council's (NRC) Committee on National Statistics (CNSTAT) report to NSF[1]. That report recommended a data collection design based on three 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 Attachment 1 – National Science Foundation Act of 1950.)

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), the Foundation, 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, 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 the SDR is a major part, provides much of the information to meet this mandate.

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

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.

In addition, the Committee for Equal Opportunity in Science and Engineering (CEOSE), an

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

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. Similarly, ad hoc committees convened by the National Research Council of the National Academies (Advisors to the Nation on Science, Engineering, and Medicine) have used SDR and SESTAT data in Committee reports such as the Committee on Science, Engineering, and Public Policy's 2006 report "*Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future*", and the Committee on Gender Differences in Careers of Science, Engineering, and Mathematics Faculty's 2009 report "*Gender Differences at Critical Transitions in the Careers of Science, Engineering, and Mathematics Faculty*".

2. Uses of Information

The time-series data produced by the SDR on the demographic, employment, and other characteristics of the Nation's SEH doctoral scientists and engineers have been used extensively in the policy and planning activities of the Foundation and the National Institutes of Health. The SDR data are used in assessing the quality and supply of the Nation's S&E personnel resources for educational institutions, private industry, and professional organizations as well as federal, state, and local governments. Other federal agencies, such as the Department of Commerce, USDA, DOE, and NASA, as well as state agencies request and make use of the SDR data for a variety of informational purposes.

SDR data are also an integral part of SESTAT. Researchers, policymakers, and others use information from the SESTAT database to answer questions about the number, employment, education, and characteristics of the S&E workforce. Because SESTAT 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 postdoctoral appointments, 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.

Findings from the 2010 SDR will enable the NSF to continue monitoring employment patterns of recent SEH doctorate recipients, as well as more experienced doctorate recipients in the labor market. The SDR data on the state of SEH doctorates are used for presentations to the National Science Board. Within the Foundation, SDR data are used in the evaluation and development of programs in the Education and Human Resources (EHR) Directorate, and analysis of employment pathways by several research directorates.

The SDR provides data on the educational training, work experience, and career development of persons holding SEH doctorates from U.S. institutions. Without this information, those at the NSF, along with researchers and policymakers, would be less informed when carrying out their responsibilities. The SDR data are made available through published reports, the SESTAT on-line data system, public use files and restricted licenses.

Some recent examples of NSF Publications that used the SDR data (all NSF publications can be accessed on the SRS website at http://www.nsf.gov/statistics) include:

Congressionally mandated reports –

- *Science & Engineering Indicators*

- *Women, Minorities and Persons with Disabilities in Science and Engineering*

Other NSF publications –

- biennial report series: *Characteristics of Doctoral Scientists and Engineers*
- Annual report series: *Science and Engineering State Profiles*
- *Role of HBCUs as Baccalaureate-Origin Institutions of Black S&E Doctorate Recipients (August, 2008)*
- *Thirty-Three Years of Women in S&E Faculty Positions (July, 2008)*
- *Postdoc Participation of Science, Engineering and Health Doctorate Recipients (March 2008)*
- *Why Did They Come To The United States? A Profile of Immigrant Scientists and Engineers* (June 2007)
- *All In a Week's Work: Average Work Weeks Of Doctoral Scientists And Engineers* (December 2005)

Data Dissemination and Access:

The SDR data from the past decade are incorporated in the SESTAT on-line data system for each survey cycle since 1993 and are available as a component of the SESTAT public-use data files, or as separate stand-alone public-use files, or as restricted use files licensed by NSF. The SESTAT on-line system allows Internet users to create customized data tabulations in subject areas of their interest. The SESTAT Home Page can be accessed at http://www.nsf.gov/statistics/sestat.

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

Since 2005, NSF has distributed over 850 copies of SDR public-use files (2001, 2003 and 2006 survey cycles), as well as over 1,600 copies of the SESTAT public-use files (1993-2006 survey cycles). The SDR is also a component of the SESTAT integrated file, which describes the entire science and engineering workforce. There are currently 38 licensees for use of the SDR; there are also 18 licenses for the SESTAT data, which includes the SDR. Several licensing requests for the SDR are pending review and approval by NSF.

Recent examples of use of the SDR data include:

Selected Presentations by non-NSF staff:

- *Internationalization of U.S. Doctorate Education,* National Bureau of Economic Research, March 2009.
- *Why Graduate Students Reject the Fast Track,* University of California-Berkeley Faculty Family Edge Project, Jan 2009.
- *Task Assignments: Generalists vs. Specialists*, Economic Theory Workshop, University of Melbourne, September 2007.
- *Science and the University: Challenges for Future Research,* CESifo Economic Studies on Economics of Higher Education Conference, July 2007.
- *Early Careers for Biomedical Scientists: Doubling (and Troubling) Outcomes*, Science and Engineering Workforce Project, February 2007.
- *Gender Equity in Higher Education: What the President of Harvard Doesn't Know or How Molehills Become Mountains of Inequity,* University of Wisconsin Women's Studies, March 2005.
- *The Success of Female Scientists in the 21st Century,* Faculty Horizons Workshop, July 2005.
- *A Brain is a Terrible Thing to Lose: Locating U.S.-Educated Foreign Nationals Intending to Live*

Abroad, 2005 American Association for Public Opinion Research, May 2005.

Selected Citations of SDR data in other sources:

- *Stay Rates of Foreign Doctorate Recipients from U.S. Universities, 2006,* Science and Engineering Education Program of Oak Ridge Institute for Science and Education, 2009
- *The Impact of Information Technology on Scientists' Productivity, Quality, and Collaboration Patterns,* National Bureau of Economic Research, 2009.
- *UC Berkeley, Leads Nation in Prepping Students for Doctorates,* UC Berkeley News, January 2005.
- *Who's Patenting in the University? Evidence from the Survey of Doctorate Recipients*, Economics of Innovation and New Technology, 2009
- *Ethnic and Technical Clustering: Native-Born Americans versus Foreign S&E Graduates,* International Studies in Entrepreneurship, 2008.
- *Negative Effects of University Patenting: Myths and Grounded Evidence,* Scientometrics, 2008.
- *Problems in the Pipeline: Gender, Marriage, and Fertility in the Ivory Tower,* Journal of Higher Education, 2008.
- *Beyond Bias and Barriers: Fulfilling the Potential of Women in Academic Science and Engineering, Committee on Science Engineering and Public Policy,* 2007.
- *Employment Preferences and Salary Expectations of Students in Science and Engineering,* Bioscience, 2007.
- *Science and the University (several chapters),* University of Wisconsin Press, 2007.
- *Foreign-Born Academic Scientists and Engineers: Producing More and Getting Less Than Their U.S.-Born Peers?* Research in Higher Education, December 2007.
- *Job Satisfaction of The Highly Educated: The Role of Gender, Academic Tenure, and Earnings*, Scottish Journal of Political Economy, May 2006.
- *Educational Mismatch among Ph.D.s: Determinants and Consequences,* National Bureau of Economic Research Working Paper 12693, October 2006.
- *Gender Differences in Major Federal External Grant Programs,* Rand Corporation Technical Report, 2005.
- *Do Babies Matter? Career progress of Women Faculty,* American Association of University Professors (AAUP): Academe, 2004.

3. Consideration of Using Improved Technology

The 2010 SDR will collect data using three different modes of data collection: 1) paper self-administered questionnaires (mail); 2) computer-assisted telephone interviews (CATI); and 3) self-administered online surveys via the World Wide Web (Web). Until the 2003 survey cycle, SDR data were collected by first mailing paper questionnaires to sample persons and then following up the nonrespondents with CATI. The tri-mode data collection effort including mail, CATI and Web was tested in the 2003 SDR and fully implemented in the 2006 SDR. The 2010 survey cycle will be the third round of a fully implemented tri-mode data collection protocol.

During the 2003 SDR, the survey launched a beta Web version and conducted experiments on the efficacy of using the Web and CATI modes as the start mode of data collection. The experiment group sizes were relatively small because the NSF wanted to ensure that using the Web and CATI as the start mode of data collection would result in high quality data, reasonable production costs and respondent satisfaction. Additionally, during the 2003 data collection effort, respondents in all modes were asked to state their mode preference for completing the survey in future rounds.

In its initial rollout in 2003, the Web survey met with very positive reactions. A number of sample members, initially asked to complete the survey in either the CATI or mail mode, completed the survey on Web. Further, of the respondents that answered the mode preference question, 49.0 percent indicated a preference for the Web survey mode.

While the Web mode is the most efficient with regard to cost, NSF wanted to verify that the data quality of this very important survey was not compromised by introducing the Web mode option. Careful data quality analysis was conducted at the unit and item level. Resulting analysis showed that the Web mode showed higher response rates as well as more complete survey and contacting data than the mail mode. Overall, the data obtained from the Web mode was considered to be higher in quality than the data obtained from the mail mode.[2].

In 2006 and 2008 SDR, the Web survey was offered as a start mode option to panel members, who in the preceding round, indicated their preference for the Web mode. With the increase in sample members offered the Web mode, both the 2006 and 2008 SDR showed a large increase in Web completes. Further, many sample members that started in mail and CATI modes also responded by the Web. Approximately 47 percent of the 2006 participants completed a Web survey and 57 percent who provided a response to the future mode preference question indicated a preference for the Web; in 2008, over 57 percent of sample members completed a Web survey and 65 percent of respondents who answered the future mode preference question indicated a preference for the Web. Based on the results of honoring mode preference in the 2006 and 2008 SDR rounds, the 2010 SDR will also honor mode preference. NSF expects that 60 percent or more of the 2010 survey response will be in the Web mode.

The 2010 data collection effort, conducted by NORC, will use a comprehensive computerized case management system that will track data capture across the three modes (mail, CATI, Web) from the 2010 SDR respondents. NORC will use mrInterview, a core product of the Dimensions family of SPSS, in implementing the computer-assisted data entry (CADE), the CATI with a telephone number management system (TNMS) incorporated in the case management system, and the Web instruments. By using one software platform, data from multiple modes of data collection can easily be integrated and delivered as a database. Optical scanning will be used to capture the digital images of the mail questionnaire after keying. The images will be stored in a database for archival purposes.

4. Efforts to Identify Duplication

Duplication does not exist. No other data collection is based on a probability sample of the U.S. trained doctoral population in science, engineering and health fields living in the United States and more recently since the 2003 SDR, living abroad (as part of the ISDR). Data from the Current Population Survey and the American Community Survey provide occupational estimates and only estimates of degree field earned at the bachelor's level. The 2010 survey is necessary to obtain trend data on continuing education and career paths of U.S. trained doctorate holders in science, engineering and health fields as well as data that reflect trends in employment patterns. The 2010 survey also will become the baseline for describing the employment characteristics of the ISDR panel of non-U.S. citizens at birth that have earned a U.S doctorate between 2001 and 2009 and emigrated from the U.S. after receiving their degree. There is no similar information available on this highly trained population that may be used, modified, or made comparable to the SDR.

5. Efforts to Minimize Burden on Small Business

Not applicable. The SDR collects information from individuals only.

6. Consequences of Less Frequent Data Collection

[2] Grigorian, K. and S. Sederstrom, 2005. Qualitative Comparison of Paper and Online Self-Administered Modes. Paper presented at the American Association for Public Opinion Research annual meetings, Miami Beach, FL, May 2005.

Because the SDR is a longitudinal survey, conducting the survey less frequently would make it more difficult to locate the persons in the sample because of the mobility of the U.S. population. This would result in both a higher attrition rate as well as less reliable estimates. Also, NSF's biennial reports and 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 survey cycles 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 Congressionally mandated biennial reports prepared by the NSF.

Follow-up surveys every two years on the same sampled persons are necessary to track changes in the SEH workforce due to large movements in and out of SEH occupations over both business cycles and life cycles. To make sure of the availability of current national data, the SDR is conducted and coordinated with the other two SESTAT surveys, NSCG and the NSRCG. 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 Consultations Outside the Agency

Federal Register Announcement

The Federal Register Notice for the SDR appeared on February 17, 2010 (See Attachment 2). NSF received no public comment in response to the announcement by the closing 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 first decade of 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. In order 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 activities covered several areas: sampling frame, population coverage, sample design, survey content, data system design, data dissemination, and informed redesign of the SESTAT surveys. 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 on Redesign

Both internal and external consultation took 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 commissioned the Committee on National Statistics (CNSTAT) of the National Research Council (NRC) to examine proposed sample design options for the SESTAT surveys. The CNSTAT committee held a two-day workshop on this topic and issued a report with recommendations to NSF on the 2010 and beyond sample design. The recommendations formed the basis for the 2010 NSCG design.[3]

- 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 SESTAT surveys. As a part of this activity, SRS worked with the Census Bureau to test various versions of a FOD question.

- SRS coordinated with OMB on wording for the collection of data on disability items in the SESTAT surveys to increase consistency across the Federal statistical agencies in surveys with such questions. As a result, SRS made two changes in 2010 to all three SESTAT surveys: (1) changed the stem to refer to "functional limitations" and (2) added a category on cognitive limitations based on the ACS item. NSF proposed not making major changes in the disability items that have been historically used in the SESTAT survey (since 1993) because there was not yet consensus on the items that federal surveys should use.

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 2008 and 2010 survey rounds:

- SRS has convened a Human Resources Experts Panel (HREP) to help the Division of Science Resources Statistics (SRS) improve data collection on the S&E labor force through review and renewal of the S&E personnel surveys and to promote use of the data for research and policy analysis purposes. HREP accomplishes 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 is made up of 15 members who represent the sciences, academia, business/industry, government, researchers and policy makers. The panel has met 5 times since it 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 with a selection of current and potential future licensees who presented their research findings and ideas to NSF as well as to the broader research community.

- The SESTAT surveys, particularly the SDR, 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 research, SRS funded a workshop at NSF that brought in experts on

[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.

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

No incentives will be offered to respondents in the initial stages of data collection. Should NSF decide that incentives need to be offered in later stages of the data collection to increase response rates, an incentive plan will be submitted to OMB during the 3rd or 4th month of data collection. See section B.3 for details on the issuance of incentives.

10. Assurance of Confidentiality

The NSF, and its contractor NORC, are fully committed to protecting the confidentiality of all survey respondents. SDR data will be collected under the authority of the National Science Foundation Act of 1950, as amended, and the Confidential Information Protection and Statistical Efficiency Act (CIPSEA) of 2002. Cover letters and survey questionnaires to each selected respondent advise them that the information they provide is confidential (see Attachment 3 – Proposed 2010 SDR Mailing Materials and Attachment 4 – Proposed 2010 SDR Questionnaire). The same notice of confidentiality will be used in the introduction to the CATI interview and will be displayed prior the start of the survey in the Web instrument.

Standard data collection procedures at NORC incorporate numerous safeguards for the data and must conform to a detailed security plan approved by NSF. While collecting SDR data, NORC separates information that could identify a particular sample member from data about that person. Each sample member is assigned a unique identifier, and this identifier is used to store identifying information (such as name, address, etc.) in a separate and secured database apart from the survey response database. The SDR affiliated NORC and NSF staff also receive annual CIPSEA training to reinforce their legal obligations to protect the privacy and confidentiality of the SDR data and staff must sign data use agreements annually to acknowledge this legal obligation.

SDR hard copy questionnaires and other contact materials are housed in a secured storage room at NORC's production facility. Hard copy materials are accessed from the file room only by authorized staff and only when necessary for data collection activities. NORC's electronic systems are on a local area network (LAN). All NORC systems used to store electronic survey data are secure by design and protected by passwords only available to authorized study staff.

NORC takes special steps to ensure that data collected via the Web questionnaire are secure. First, access to the Web instrument is only allowed with a valid Personal Identification Number (PIN) and password correctly entered in combination. Second, data are transmitted by the Secure Sockets Layer (SSL) protocol that uses powerful encryption during transmission through the Internet. If a respondent keeps a Web survey open without any activity, the Web server at NORC closes it after a short period of inactivity, thus preserving the data up to the break-off point and securely closing the connection. The Web system architecture process has been designed in a way that places authentication information and response data on physically separate servers. This strategy provides an extra layer of security to protect response data. Both development and production servers are backed up nightly, as NORC's disaster recovery plan requires.

All data and analysis are reported in aggregate form only and measures are taken so that the identity of individuals or organizations is not disclosed.

11. Justification for Sensitive Questions

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

12. Estimate of Respondent Burden

A statistical sample of approximately 45,700 persons, identified as having a doctorate in a science, engineering or health field from a U.S. university will be selected for the 2010 SDR. The amount of time to complete the questionnaire may vary depending on an individual's circumstances; however, on average it will take approximately 25 minutes to complete the survey. Assuming an 85% response rate (38,845 respondents) NSF estimates that the total burden for the 2010 SDR will be 16,185 hours.

The total cost to respondents for the 16, 185 burden hours is estimated to be $762,657. This is based on an estimated median annual salary of $98,000 per full-time employed SDR respondent from the 2008 SDR data. Assuming a 40-hour workweek over 52-weeks of employment, this annual salary corresponds to an hourly rate of $47.12.

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 the Federal Government

The total estimated cost to the Government for the 2010 SDR is $6.0 million for data collection costs, NSF staff costs to provide oversight and coordination with the other two SESTAT surveys. The cost estimate for the data collection is $5.04 million, which is based on sample size; length of questionnaire; CATI and Web data collection technology; administrative, overhead, design, printing, mail and telephone data collection costs, 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; preparation of documentation and final reports; analysis, and tabulations. The NSF staff costs are estimated at $562,500 ($150,000 annual salary of 1.5 FTE for 2.5 years of the 2010 SDR survey cycle).

15. Reason for Change in Burden

The 2010 SDR will include a slightly larger sample size (from 42,600 in 2008 to 45,700 in 2010) to accommodate the additional subsample added to the ISDR panel. The change in requested burden hours from the 2008 SDR (15,088 burden hours) reflects the increase in the total SDR sample size.

Note: The current approval for the 2008 clearance of the SDR (70,610 respondents) is incorrect. During data entry of the 2008 SDR clearance request, the 2006 clearance was still open and ROCIS added the number of respondents for 2008 (36,210) to the respondents for 2006 (34,400) together (and the same for the burden hours) to get 70,610 respondents, representing two years instead of one.

16. Schedule for Information Collection and Publication

The 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 (Mail, CATI, Web)	October 2010 – June 2011
Coding and Data Editing	December 2010 – September 2011
Final Edited/Weighted/Imputed data file	December 2011
SDR Info Brief	Spring 2012
SDR Detailed Statistical Tables	Summer 2012
SDR Public Use File	Summer/Fall 2012

17. Display of OMB Expiration Date

The OMB Expiration Date will be displayed on the 2010 SDR questionnaire; in the Web survey version, it will be included on the informed consent page of the Web survey, and available in a help screen accessible at any point in the Web survey; in the telephone interview, it will be read to sample members during the introductory informed consent.

18. Exception to the Certification Statement

Not Applicable.