SF-83-1 SUPPORTING STATEMENT

for the

2017
National Survey of College Graduates

CONTENTS

2017 NATIONAL SURVEY OF COLLEGE GRADUATES
SUPPORTING STATEMENT

A. JUSTIFICATION

This request is for a three-year renewal of the previously approved Office of Management and Budget (OMB) clearance for the National Survey of College Graduates (NSCG). The NSCG has historically served as a valuable source of information on the education and career paths of the Nation's college-educated population. The most recent NSCG was conducted in 2015 (OMB approval number 3145-0141). The current OMB clearance for the NSCG expires May 31, 2018, which does not cover the complete survey cycle for the 2017 NSCG.

The 2017 survey cycle marks the full implementation of a four-panel rotating panel design that began with the 2010 NSCG. With this stability in the sample design for the survey, the planning for the 2017 NSCG focused on data collection and identified the need for enhancements in the NSCG web survey instrument and in our data collection methodology as described in the sections below. For example, new to the 2017 NSCG, the web instrument will be optimized for use in mobile devices.

1. NECESSITY FOR INFORMATION COLLECTION

In 2010, the America COMPETES Reauthorization Act of 2010[1] established the National Center for Science and Engineering Statistics (NCSES) at the National Science Foundation (NSF) and directed NCSES to "...collect, acquire, analyze, report, and disseminate statistical data related to the science and engineering enterprise in the United States and other nations that is relevant and useful to practitioners, researchers, policymakers, and the public..." Information obtained through the NSCG is critically important to NCSES's ability to measure the education and employment of scientists and engineers. Furthermore, the NSCG data along with the NCSES's Survey of Doctorate Recipients (SDR)[2] data serve as the nation's only source of comprehensive information about the size and characteristics of the science and engineering (S&E) workforce[3]. These data are solicited under the authority of the NSF Act of 1950[4], as amended, and are central to the analysis presented in a pair of congressionally mandated reports[5,6] published by NSF:

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

[1] Section 505, Pub. L. No. 111-358. See Appendix A.
[2] The SDR is a longitudinal biennial survey that provides demographic and career history information about individuals with a research doctoral degree in a science, engineering, or health field from a U.S. academic institution. For more information, see http://www.nsf.gov/statistics/srvydoctoratework.
[3] The S&E workforce includes individuals with degrees or occupations in computer and mathematical sciences, life sciences, physical sciences, social sciences, engineering, health sciences and related fields.
[4] See Appendix B.
[5] 42 U.S. Code § 1863(j)(1)
[6] 42 U.S. Code § 1885(a), 1885(d)

In addition, the Science and Engineering 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." [7] The NSCG and SDR provide much of the information to meet this mandate. The combined data from these two surveys, initially created for the 1993 survey cycle and developed throughout the past two decades, are based on recommendations of the National Research Council's Committee on National Statistics (CNSTAT) report to NSF.[8]

NSCG Background

The NSCG provides data on the nation's college graduates, with particular focus on those in the S&E workforce. The NSCG samples individuals who are living in the United States, have at least a bachelor's degree, and are less than 76 years of age. This survey is a unique source for examining various characteristics of college-educated individuals, including occupation, work activities, salary, the relationship of degree field to occupation, and demographic information.

Collectively, the NSCG and SDR provide comprehensive information on the entire U.S. population of scientists and engineers with at least a bachelor's degree. Historically, these surveys have been conducted every two to three years and, jointly, provide both cross-sectional and longitudinal data on the education and employment of the college-educated U.S. S&E workforce. The NSCG and SDR are the only available sources of detailed information that support a broad range of policy and research topics on the dynamics of the S&E workforce over time.

The NSCG provides information on individuals residing in the U.S. with at least a bachelor's degree including those who received degrees only from foreign institutions. The SDR compliments these data with information on the population of U.S.-degreed doctoral level scientists and engineers. Through 2010, the National Survey of Recent College Graduates (NSRCG) complimented the NSCG and SDR data with the inflow of U.S.-degreed bachelor's and master's level scientists and engineers. Beginning in 2013, the NSCG began capturing the bachelor's and master's level inflow of new graduates and eliminated the need for the NSRCG. As a result, the NSRCG was discontinued after the 2010 survey.

The panel data from the NSCG provide valuable information on careers, training, and educational development of the nation's college graduate population. These data enable government agencies to assess the scientific and engineering resources available in the U.S. to business, industry, and academia, and provide a basis for the formulation of the nation's S&E workforce policies. For example, educational institutions can use the NSCG data in establishing and modifying scientific and technical curricula, while various industries can use the information to develop recruitment and remuneration policies.

[7] 42 U.S. Code § 1885(d)

[8] 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.

2. USES OF INFORMATION

Policymakers, researchers, and other data users use information from the NSCG and SDR 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, policymakers and researchers use the NSCG and SDR data to address questions on topics such as employment of foreign-born or foreign-degreed scientists and engineers, the transition from higher education to the workforce, the role and importance of postdocs as research personnel, 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.

Uses for Policy Discussion
Data from NCSES's 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 NSCG data and the combined NSCG and SDR data are:

- The Executive Office of the President used NSCG data to examine the contributions of immigrants in S&E occupations[9];
- The National Science Board (NSB) used the combined NSCG and SDR data in its investigation to develop national policies for the S&E workforce[10];
- The Commonwealth of Massachusetts Governor's Advisory Council for Refugees and Immigrants used NSCG data to examine the number of foreign-born residents that are trained healthcare professionals[11];
- The U.S. Small Business Administration used NSCG data to investigate differences in STEM entrepreneurship participation between native-born and foreign-born workers[12];
- The importance of information on the S&E workforce to inform public policy can be seen in discussions of the NSB's Task Group on Science, Technology, Engineering, and Math (STEM) Innovators. The task group used the combined NSCG and SDR data to inform its deliberations about the S&E workforce and these data were part of the final report[13];
- The Committee for Equal Opportunity in Science and Engineering (CEOSE), an advisory committee to 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

[9] http://www.whitehouse.gov/sites/default/files/docs/report.pdf and https://www.whitehouse.gov/sites/default/files/docs/economic_effects_of_immigration_ea_february_2015_update_final_v2.pdf
[10] http://www.nsf.gov/nsb/documents/2003/nsb0369/nsb0369.pdf and http://nsf.gov/nsb/publications/2015/nsb201510.pdf
[11] http://www.miracoalition.org/images/stories/gac_task_force_report_final-12.18.14.pdf
[12] https://www.sba.gov/sites/default/files/advocacy/rs432tot-Immigrant-STEM-Entrepreneurs.pdf
[13] http://www.nsf.gov/nsb/publications/2010/nsb1033.pdf

makes extensive use of the combined NSCG and SDR data to highlight key areas of concerns relating to students, educators and technical professionals;

- The Council of Graduate Schools (CGS) used NSCG data to estimate the potential monetary cost and return on investment of pursuing advanced degrees[14], which is a key element of CGS's financial education website – www.gradsense.org; and
- The Educational Testing Service (ETS) and CGS used the combined NSCG and SDR data to examine national benchmarks for career outcomes of master's and doctoral degree recipients by specific field[15].

Uses by NSF

The NSCG data were used extensively in the latest versions of the congressionally mandated biennial reports *Science and Engineering Indicators, 2016* and *Women, Minorities and Persons with Disabilities in Science and Engineering, 2015*. In addition, *Women, Minorities and Persons with Disabilities in Science and Engineering, 2017*, set for release next year, will use NSCG data.

NSF used the NSCG data and the combined NSCG and SDR data in recent reports such as:

- *Immigrants' Growing Presence in the U.S. Science and Engineering Workforce: Education and Employment Characteristics in 2013,* September 2015
- *Characteristics of the College-Education Population and the Science and Engineering Workforce in the United States,* April 2015
- *Employment Decisions of U.S. and Foreign Doctoral Graduates: A Comparative Study,* December 2014
- *Unemployment among Doctoral Scientists and Engineers Remained Below the National Average in 2013,* September 2014
- *Employment and Educational Characteristics of Scientists and Engineers,* January 2013
- *International Mobility and Employment Characteristics among Recent Recipients of U.S. Doctorates,* October 2012
- *International Collaboration of Scientists and Engineers in the United States,* August 2012
- *Diversity in Science and Engineering Employment in Industry,* March 2012
- *Racial and Ethnic Diversity among U.S.-Educated Science, Engineering, and Health Doctorate Recipients: Methods of Reporting Diversity,* January 2012
- *Community Colleges: Playing an Important Role in the Education of Science, Engineering, and Health Graduates,* July 2011

[14] http://www.gradsense.org/gradsense/methodology
[15] http://www.ets.org/c/19574/19089_PathwaysReptqp.pdf

All NSF Publications can be accessed on the NCSES website at http://www.nsf.gov/statistics/reports.cfm.

Uses by Researchers and Analysts

NCSES makes the data from the NSCG available through published reports, our online data tool, downloadable public use files, and restricted-use licenses. The online data tool, available at https://ncsesdata.nsf.gov/sestat/sestat.html, allows users to create customized data tabulations using NSCG data. The NSCG public-use files are available for download through the NCSES data downloads web page at https://ncsesdata.nsf.gov/datadownload/.

Since 2005[16], NCSES has distributed over 800 copies of the more than decade-old 1993 NSCG public-use files, over 1,500 copies of the 2003 NSCG public-use files, and nearly 1,200 copies of the 2010 NSCG public-use files to researchers in government, academia, and professional societies. And, since its release in April 2015, over 900 copies of the 2013 NSCG public-use files have been downloaded from the NCSES data downloads page. The 2015 NSCG data are in the final stages of data review and will be available in the coming months as a standalone public-use file. The NSCG public-use files receive heavy use 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.

The combined NSCG, SDR, and NSRCG public-use files have been downloaded from the NCSES data downloads page over 6,000 times since 2005. In addition to the users of the public-use files, there are currently 19 restricted-use licensees with access to the combined NSCG, SDR, and NSRCG microdata files under a licensing agreement with NCSES.

Some of the research based on the public-use NSCG data, the combined public-use data, and the restricted-use data resulted in papers such as:

- *Occupational and Organizational Effects on Wages among College-educated Workers in 2003 and 2010,* Texas A&M University, 2016
- *The Private and Social Benefits of Double Majors,* St. Lawrence University, 2016
- *Staying in STEM or Changing Course: Do Natives and Immigrants Pursue the Path of Least Resistance?,* Ohio State University, 2016
- *Are College Costs Worth it? How Ability, Major, and Debt Affect the Returns to Schooling,* Temple University, 2016
- *Why Do Women Leave Science and Engineering?,* Rutgers University, 2016
- *Sex, Race, and Job Satisfaction Among Highly Educated Workers,* Vanderbilt University, 2016

[16] The America COMPETES Reauthorization Act of 2010 mandated the name and responsibilities of NCSES. Prior to 2010, the organizational unit that would become NCSES was referred to as the National Science Foundation's (NSF) Division of Science Resource Statistics (SRS). For simplicity, NCSES will be used throughout this report when referring to work completed by SRS or NCSES.

- *Highly Skilled Migrants: Risks and Hedging Mechanisms*, Texas Tech University, 2016
- *The Bachelor's to Ph.D. STEM Pipeline No Longer Leaks More Women than Men: A 30-Year Analysis*, Northwestern University, 2015
- Salary *and Job Satisfaction Among Economics and Business Graduates: The Effect of Match Between Degree Field and Job*, University of South Florida, 2015
- *The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects*, National Bureau of Economic Research, 2015
- *Has the Quality of Accounting Education Declined?*, University of Florida, 2014
- *Trends in Earnings Differentials across College Majors and the Changing Task Composition of Jobs*, Yale University, 2014
- *Are Asian American Women Advantaged? Labor Market Performance of College Educated Female Workers*, Kansas University, 2014
- *Opting Out among Women with Elite Education*, Vanderbilt University, 2013

3. CONSIDERATION OF USING IMPROVED TECHNOLOGY

The data for the 2017 NSCG will be collected by the U.S. Census Bureau under an interagency agreement between NCSES and the Census Bureau. The 2017 NSCG data collection will use a multi-mode approach that begins with a web invitation letter mailed to sample persons asking them to complete the survey on the Internet. Nonrespondents will be followed up using a paper questionnaire mailing and computer assisted telephone interviews (CATI). The data will be collected and managed by the Census Bureau using multiple complementary systems including: Docuprint, Intelligent Mail Barcoding, Enterprise Internet Solutions, Adaptive Design and Intermittent Data Processing, and the Unified Tracking System. These systems are described below.

Docuprint and Intelligent Mail Barcoding
Web invitation letters are produced through an in-house on-demand print process using a Docuprint system which allows personalization and the ability to tailor items to each specific respondent. The letters and questionnaire packets will be tracked using Intelligent Mail Barcoding (IMB). IMB requires separate outgoing and return barcodes to be placed on NSCG envelopes for tracking purposes. Using IMB has the potential to increase the overall efficiency of data collection enabling the collection of detailed tracking information including:

- When an outgoing questionnaire or other mail piece reached a respondent's local post office;
- When an outgoing mail piece left the post office with a postmaster for delivery;
- If the outgoing mail piece was identified as undeliverable-as-addressed (UAA) and is being rerouted for return;

- When a return questionnaire reaches a respondent's local post office; and
- When a return questionnaire reaches its destination.

This information will allow the NSCG to put cases on hold while the returned questionnaire is reviewed to determine whether it is a "good complete." Placing cases on hold will reduce respondent burden by limiting unnecessary contacts. In addition, the IMB tracking will alert the NSCG staff to undeliverable mail pieces while they are still in circulation, allowing the Census Bureau to reduce the NSCG data collection costs by eliminating any future mailings to undeliverable addresses.

Enterprise Internet Solutions and Mobile Optimization
The Enterprise Internet Solutions (EIS) area of the Application Services Division (ASD) at the Census Bureau will host a web-based data collection instrument. Data will be transmitted and processed daily. The web instrument will be hosted on the fully certified and accredited Centurion system (infrastructure, security, and framework). New to the 2017 NSCG, the web instrument will be optimized for use in mobile devices. This enhancement will create a better experience for mobile device users attempting to complete the survey and, as a result, should reduce survey breakoffs and reduce the possibility of measurement errors.

Adaptive Design and Intermittent Data Processing
The 2017 NSCG will continue to expand the scope of adaptive design in an effort to attain high-quality survey estimates in less time and at less cost than traditionally executed survey operations. In 2013, adaptive design implementation focused mainly on developing operational capabilities, while in 2015, the focus was on developing statistical and monitoring capabilities. In 2017, the main focus will be on the increased automation of existing capabilities and the prediction of the expected effects of our data collection interventions. First, the Census Bureau will improve intermittent processing (editing, imputation, weighting) of the response data throughout the data collection period. We refer to this implementation of our complete data processing steps on an intermittent basis as "flow processing." In addition to operational efficiencies, flow processing will allow the NSCG survey team to monitor several quality measures throughout data collection, including R-indicators, benchmarking, stability of estimates, and response propensities by mode.

Second, the 2017 NSCG will include an adaptive design experiment that aims to identify the adaptive design goals most appropriate for the NSCG, and in turn, identify appropriate data collection interventions and the monitoring methods that can be used to drive those interventions. In addition, our focus is on automating much of the decision making process that occurred in the 2015 adaptive design experiment as a way to evaluate whether adaptive design is a framework that could be rolled out in a fully production setting for a large survey with a complex data collection pathway, such as the NSCG. More detail about the 2017 NSCG adaptive design experiment is provided in section B.4. of this report. We will employ roughly the same sample sizes as the 2015 adaptive design experiment in order to provide the statistical power to make definitive statements about statistical differences between the treatment group and the control group on various measures, including response rates, R-indicators, cost, and effect on key estimates. In addition, we will draw upon the previous two adaptive design experiments to set

statistical expectations for the effects of our potential data collection interventions as a way to gauge longer term suitability of potential data collection interventions.

Unified Tracking System

In 2017, the NSCG will be expanding its use of the Census Bureau's Unified Tracking System (UTS) to assist in various aspects of survey management. As in 2013 and 2015, the UTS will provide a full contact history report for the NSCG, giving survey managers a single place to view all contacts integrated from all three modes in the NSCG along with the outcomes of those contacts. This contact history reporting system enables the examination of contact strategies in a number of ways. As an example, if respondents called in to check on the status of their response, NSCG staff are able to quickly and easily access the respondents' contact history and outcomes to provide the current status of their response. In addition, this report provides an easily accessible and interpretable audit trail of all contacts, allowing survey managers to immediately verify if NSCG interviewers are following proper contact protocols, particularly when questions or complaints from respondents arise. For 2017, this contact report will be enhanced by the integration of the previously mentioned IMB data.

The UTS will also be implementing new reports for the 2017 NSCG, focused on monitoring as opposed to aggregation. From a data quality perspective, the UTS will provide daily updates for R-indicators analysis at the cohort-level, so that survey management can understand how data collection operations affect representativeness. In addition, historical R-indicators for the 2013 and 2015 NSCG will also be provided for comparison. Additionally, the UTS will be providing two reports to monitor IMB data. These reports will focus on the difference between the dates provided by the Census Bureau's National Process Center (NPC) and IMB-provided dates for survey monitoring purposes. For outgoing mailings, the report will show the lag between the scheduled mail date of mailings and when NSCG packages actually enter the mail stream. For incoming mailings, the report will provide the dates when UAAs or return questionnaires enter the IMB system versus when they are checked in NPC. Both reports will have these data broken down by mailing geographies. These reports will help us understand the relationship between when sample persons actually receive their mail and when they respond to survey requests. They may also help us develop a better understanding of when we should expect to see significant increases in response relative to mailout operations.

Finally, we plan to incorporate a new UTS report for the 2017 NSCG that documents the interactions of the NSCG sample with the web instrument. This report will provide information like the number of sample persons that have logged in and with what type of device, statistics about the time spent responding, and whether they logged out or submitted the survey. In past NSCG cycles, the analysis of this valuable web paradata occurred at the end of the survey cycle rather than during the data collection effort.

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 1970s. 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 (ACS) 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.

The NSCG and ACS both collect demographic information including gender, race, ethnicity, marital status, citizenship status, and veteran status (with veteran status questionnaire items being added to the NSCG beginning with the 2017 survey cycle). This survey content duplication between the ACS and NSCG is necessary because of the confidentiality restrictions placed on the public release of ACS data. Due to these restrictions, it is not possible for NSF to link the demographic information from the ACS with the detailed education and employment information collected on the NSCG. Since linkage between demographic, education, and employment information is needed for the analysis conducted in the preparation of NSF's congressionally mandated reports, all of this information is collected on the NSCG.

Overlap does exist in the target populations for the NSCG and the SDR. As a result, it is expected there will be approximately 300 individuals selected for sample in both the 2017 NSCG and the 2017 SDR.

In the 2013 NSCG survey cycle, the NSCG and SDR survey contractors identified the individuals selected for both surveys, removed the individuals from the NSCG data collection effort, and, at the completion of the SDR data collection effort, used the SDR responses for these individuals to complete the individual's record on the NSCG data file. This NSCG/SDR deduplication process required the SDR survey contractor to create numerous files containing all SDR sample cases for use by the NSCG survey contractor. Furthermore, given file format and processing differences between contractors, the NSCG survey contractor needed to reformat and manually manipulate many of the SDR files to use them in combination with the NSCG files. The NSCG/SDR deduplication process added over a week of staff time to both the NSCG and SDR processing during the 2015 survey cycle.

Given recent changes to the NSCG questionnaire content, there are noticeable differences in the information collected on the NSCG and SDR. Examples of topics planned for collection on the 2017 NSCG, but not on the 2017 SDR include attainment of certifications and licenses, financial support for education, and community college enrollment. Because of the content differences, the small number of expected duplicates, and the operational challenges of the deduplication process, NCSES will not deduplicate individuals selected for sample in both the NSCG and SDR during the 2017 survey cycle.

5. EFFORTS TO MINIMIZE BURDEN ON SMALL BUSINESS

Not applicable. The NSCG collects information from individuals only.

6. CONSEQUENCES OF LESS FREQUENT DATA COLLECTION

The NSCG data are central to the analysis presented in a pair of congressionally mandated reports published by NSF – *Science and Engineering Indicators* and *Women, Minorities, and Persons with Disabilities in Science and Engineering*. Since these reports are published on a biennial schedule, they rely on the availability of updated data on the S&E workforce every two years. Conducting the NSCG on a less frequent basis would prohibit NSF from meeting its congressional mandate to produce a report that contains an accurate 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 impact of not being able to meet this congressional mandate is that government, business, industry, and universities would have less recent data to use as a basis for formulating the nation's science and engineering policies.

A less frequent data collection would also impact the quality of the NSCG data. Follow-up surveys every two to three years on the same sampled persons are necessary to track changes in the S&E workforce as there are large movements of individuals into and out of S&E occupations over both business and life cycles. To ensure the availability of current national S&E workforce data, the NSCG has been conducted and coordinated with the SDR and the NSRCG on a biennial basis since 1993, and with only the SDR since 2013 after the discontinuation of the NSRCG in 2010. The degradation of any component jeopardizes the integrity and value of these combined surveys to provide comprehensive information on the S&E workforce. [17]

Finally, because the NSCG is a panel survey, conducting the survey less frequently would make it more difficult and costly to locate the sampled persons in follow-up cycles because of the mobility of the U.S. population. The impact would be a higher attrition rate, higher potential for nonresponse bias, and less reliable cross-sectional and longitudinal estimates.

7. SPECIAL CIRCUMSTANCES

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

[17] Through 2010, the NSRCG complimented the NSCG and SDR data with the inflow of U.S.-degreed bachelor's and master's level scientists and engineers. Beginning in 2013, the NSCG began capturing the bachelor's and master's level inflow population and eliminated the need for the NSRCG. As a result, the NSRCG was discontinued after the 2010 survey without any impact on the coverage provided by the NSCG and SDR.

8. FEDERAL REGISTER ANNOUNCEMENT AND CONSULTATION OUTSIDE THE AGENCY

Federal Register Announcement

The Federal Register announcement for the NSCG appeared on July 28, 2016. NSF received one public comment in response to the announcement. See Appendix C for both the announcement and the comment. The comment requested that NCSES consider the use of administrative records, specifically the National Student Clearinghouse, to obtain information on education background in lieu of asking this information from respondents.

NCSES informed the commenter that, at the request of NCSES, the Census Bureau's Center for Administrative Records Research and Application (CARRA) is conducting research to compare administrative records data with the NSCG respondent-provided data. This research will inform survey content discussions for future NSCG survey cycles and will provide insight on the necessity of certain NSCG questionnaire items including the education background items.

Consultation Outside the Agency

NCSES has sought the advice and guidance of survey methodologists, statisticians, demographers, researchers, data analysts, and policymakers to examine numerous issues related to the development of the NSCG.

- Evaluation of the NCSES Effort to Measure the S&E Workforce Population

  The National Academies of Science, Engineering, and Medicine's Committee on National Statistics (CNSTAT), at the request of NCSES has convened an expert panel to review, assess, and provide guidance on NCSES's effort to measure the S&E workforce population in the United States. Given the evolving data needs of NCSES stakeholders and the budget climate uncertainty under which NCSES operates, NCSES would like to develop a framework for measuring the S&E workforce that will enable the flexibility to examine emerging issues related to this unique population while at the same time allowing for stability in the estimation of trend data. This framework would provide direction for numerous issues related to measuring the S&E workforce population including content, data sources, survey design, survey methodology, data collection, data processing, data integration, data dissemination, and data promotion.

  At the end of its review, the panel will issue a report with findings, recommendations, and priorities for improving the relevance, accuracy, timeliness, and cost-effectiveness of S&E workforce data for the next decade and beyond. The hope is that the information included in this report will provide the details, direction, and guidance necessary for NCSES to develop a robust and flexible framework for measuring the S&E workforce over the coming decades.

  CNSTAT Panel Chairs and Panel Members

  CNSTAT Panel Chairs
  Rita Colwell, University of Maryland and Johns Hopkins University
  James House, University of Michigan

  CNSTAT Panel Members
  Jennifer Sue Bond, Council on Competitiveness

Geoff Davis, Verily
Don Dillman, Washington State University
Richard Freeman, Harvard University
Jack Gambino, Statistics Canada
Maresi Nerad, University of Washington
Randy Olsen, Ohio State University
Willie Pearson, Georgia Institute of Technology
Keith Rust, Westat
Nora Cate Schaeffer, University of Wisconsin
James Wagner, University of Michigan
Yu Xie, Princeton University

- Data Processing of New Survey Content on Certifications and Licenses

  The 2015 NSCG collected information on alternative credentials including industry-recognized certifications, occupational licenses, and educational certificates. To aid in the processing of the data collected on these concepts, NCSES used the vast amount of research conducted by the Interagency Working Group on Expanded Measures of Enrollment and Attainment (GEMEnA). GEMEnA is a collaboration among federal statistical agencies established by the OMB Office of Statistical and Science Policy, the Council of Economic Advisors, and the Under Secretary of Education to improve federal data on the attainment of non-degree credentials.

  GEMEnA Member Agency Representatives

  Census Bureau
  Kurt Bauman
  Bob Kominski
  Stephanie Ewert

  Bureau of Labor Statistics
  Dori Allard
  Harley Frazis

  NCSES – National Science Foundation
  Dan Foley
  John Finamore

  Council of Economic Advisors
  Jordan Matsudaira

  OMB Office of Statistical and Science Policy
  Shelly Martinez

  Department of Education
  Jon O'Bergh

  National Center for Education Statistics
  Sharon Boivin
  Lisa Hudson
  Kashka Kubzdela
  Sarah Crissey
  Sarah Carroll
  Andy Zukerberg

- Evaluating Administrative Records as an NSCG Sampling Frame Source

  The staff at the Census Bureau's Center for Administrative Records Research and Application (CARRA) is conducting research examining the potential use of administrative records as an NSCG sampling frame source. In this research, CARRA will identify sampling frame sources, including data available at the Census Bureau, which could be acquired for use in the NSCG processing. CARRA will document the strengths and weaknesses of these files and will produce descriptive statistics on files that are acquired. Simulations using external data and the American Community Survey (ACS) data will assess the quality and viability of using alternative data sources as the NSCG sampling frame. In addition, when

appropriate, studies may be designed to explore uses of administrative records or third party data for interview screening purposes.

- Evaluating Administrative Records to Inform Measurement Error Properties of NSCG Data

  In this task, CARRA will use administrative records and third party data sources to compare with NSCG data. After working with NCSES to identify key NSCG analytical domains, CARRA will use the administrative records to evaluate the quality of the NSCG data as well as the necessity of each NSCG questionnaire item. From the evaluation, CARRA will provide NCSES with insight on the quality of the data being collected in the NSCG and the potential use of administrative records instead of the NSCG for certain questionnaire items.

- Paradata Analysis

  The staff at the Census Bureau's Demographic Statistical Methods Division analyzed the NSCG web survey instrument paradata from the 2013 and 2015 survey cycles. The primary purpose of this research was to understand respondents' interaction with the web survey instrument to identify areas where the instrument needed improvement and then formulate recommendations that target those areas. The findings from this research led to enhancements to the web survey instrument for the 2017 survey cycle.

- Mobile Optimization Research

  The staff at the Census Bureau's Demographic Statistical Methods Division conducted research to develop a mobile optimized version of the NSCG web survey instrument for the 2017 survey cycle. Optimizing a survey instrument for mobile devices means that the survey will detect the device being used and adjust the layout accordingly. The font, buttons, and spacing will appear larger and minimize respondents' need to pinch to zoom. Additionally, horizontal scrolling to view content will be eliminated. New design features that result from mobile optimization can help reduce errors such as a respondent inaccurately recording responses, skipping survey items, satisficing, and exiting the survey prematurely (i.e., breakoff). In addition, incorporating the latest survey design trends that align with usability principles and guidelines for mobile devices can enhance respondents' experience with survey completion.

- Contact Strategies Research

  The Census Bureau's Demographic Statistical Methods Division and the Census Bureau's Center for Survey Measurement conducted research to examine the impact of different contacts on survey response to determine whether there are ways to save money and reduce respondent burden without harming response rates and sample representativeness. The research evaluated both the number of contacts and their content. To assess the number of contacts, the researchers plotted daily response rates, along with contact mailing dates and telephone call dates and times, ran simulations to hypothesize the response rates with fewer contacts, and tracked the outcome codes of telephone calls. The research also used qualitative methods to develop and assess the content of the messages. Specifically, the

research sought to determine what messages motivate people to respond, what modes of communication are preferable, and how people react to multiple contacts. The result of both the qualitative and quantitative components of this research enabled a proposal with several different contact strategies for NSCG to experimentally test as part of production in 2017. This experiment, described in section B.4., will then inform contact strategy development for the 2019 NSCG cycle.

- Adaptive Design

  The 2013 NSCG Terms of Clearance stated that "OMB looks forward to NCSES collaborating actively with the National Center for Education Statistics and the Census Bureau on ways to experiment with and apply "responsive design" methods to the NSCG in order to better measure and reduce bias and improve overall survey efficiency." Over the past two years, NCSES staff continued outreach and collaboration efforts with the Census Bureau, NCES, and other agencies to take stock of the progress made in the field of adaptive design, to identify the obstacles that currently exist, and to explore the adaptive design possibilities for the future. Below are some examples of NCSES's outreach and collaboration efforts related to adaptive design.

  - NCSES, Census Bureau, and NCES staff participated in a topic-contributed session on adaptive design at the 2015 Federal Committee on Statistical Methodology (FCSM) research conference in December 2015.
  - NCSES, Census Bureau, and NCES staff participated in a topic-contributed panel at the 2015 AAPOR annual conference in May 2015. The panel topic was "Innovation in Federal Surveys – Opportunities, Progress, and Challenges."
  - Thanks, in large part, to the collaboration between NCSES and the Census Bureau on adaptive design, NCES requested Census Bureau and NCSES staff present a seminar focused on data quality and adaptive design. This outreach has led to NCES's adoption of data monitoring metrics with an eye towards future adaptive design research and experimentation opportunities.
  - The survey contractors for NCES surveys and NCSES surveys (Research Triangle Institute, Inc. and the Census Bureau, respectively) participated in the Bayesian Adaptive Survey Design Network. This network gathers researchers from academia and national statistical offices to give a strong impetus to theory development and practical implementation of adaptive survey designs.
  - NCSES, Census Bureau, and NCES staff attended meetings of the Adaptive Design Interagency Working Group. This working group, established by the OMB Office of Statistical and Science Policy in 2014, is a collaboration among federal statistical agencies.

- Survey Design and Methodology

  NCSES has sponsored and collaborated on multiple survey design and methodology research projects in an effort to ensure that the NCSES surveys, including the NSCG, are

incorporating best practices for survey design and methodology. NCSES holds ongoing discussions with staff from NCES and the Census Bureau to discuss survey design and methodological issues of interest. In addition, NCSES funds research on survey design and methodological issues. The following provides a listing for some of the ongoing research funded by NCSES related to the NSCG:

- To produce more reliable survey estimates, NCSES funded research to examine and mitigate extreme sample weight variation within the NSCG. Jean Opsomer and Jay Breidt (Colorado State University) were the principal investigators for this research.
- To address the increasing nonresponse trends for governmental surveys, NCSES funded research to examine contact strategies for the NCSES surveys. Jolene Smyth and Kristen Olson (University of Nebraska – Lincoln) were the principal investigators for this research.

9. PAYMENT OR GIFTS TO RESPONDENTS

The 2010 NSCG and 2013 NSCG included incentive experiments to examine the impact of offering incentives on response, data quality, and cost. The results from the incentive experiments[18,19] provided NCSES and the Census Bureau with guidance and direction for using incentives in the 2015 NSCG data collection effort. The incentive usage in the 2017 NSCG will follow the procedures used in the 2015 survey cycle.

As was the case in the 2015 NSCG, we plan to offer a $30 prepaid debit card incentive to a subset of highly influential new sample cases at week 1 of the 2017 NSCG data collection effort. "Highly influential" refers to the cases with a large base weight and a low response/locating propensity. The highly influential cases will be identified by a model-based approach using a weighted response influence, which is the product of a sampled case's base weight and predicted response propensity. We expect to offer $30 debit card incentives to approximately 10,000 of the 48,000 new sample cases included in the 2017 NSCG. The weighted response influence factor is calculated as follows:

$$W_i = \log(\omega_i) * \hat{\phi}_i \text{, where } \hat{\phi}_i = \left(\frac{1}{\hat{\rho}_{Li}}\right)\left(\frac{1}{\hat{\rho}_{Ri}}\right).$$

The weighted response influence for a case, W_i, is the product of the log of the base weight, ω_i, and the response influence, $\hat{\phi}_i$. The response influence is the inverse of the product of the locating propensity, $\hat{\rho}_{Li}$, and the response propensity, $\hat{\rho}_{Ri}$.

[18] Zotti, Allison, "Report for the 2013 National Survey of College Graduates Methodological Research Incentive Timing Experiment," Census Bureau Memorandum from Reist to Finamore and Rivers, April 15, 2014, draft.
[19] Thornton, Thomas, "2013 National Survey of College Graduates (NSCG) Incentive Conditioning Study," Census Bureau Memorandum from Reist to Finamore and Rivers, April 15, 2014, draft.

In addition, using the findings from the 2013 NSCG incentive conditioning study and following our procedures from the 2015 NSCG, we plan to offer a $30 prepaid debit card incentive to past incentive recipients at week 1 of the 2017 NSCG data collection effort. As a result, we expect to offer $30 debit card incentives to approximately 7,000 of the 75,500 returning sample members.

The $30 incentive amount proposed for use in the 2017 NSCG was chosen based on findings from the 2010 NSCG late-stage incentive experiment targeting hard to enumerate cases that had not responded to the survey after multiple contacts. As part of the 2010 experiment, the hard to enumerate cases were allocated to three treatment groups:

- $30 debit card incentive
- $20 debit card incentive
- No incentive

Other than the use and amount of the debit card incentive, the three treatment groups in the 2010 NSCG late-stage incentive experiment received the same data collection contact strategy. At the conclusion of the experimental period (approximately six weeks), the response rate for the three treatment groups differed significantly. The $30 incentive treatment group had a response rate of 29.5%, the $20 incentive treatment group had a response rate of 24.1%, and the no incentive group had a response rate of 6.4%.

In addition to the increase in the response rate for the hard to enumerate cases that were targeted as part of this experiment, the use of the incentive also had a profound effect on the overall representation of the responding sample. The incentive was successful in obtaining responses from individuals who were demographically different than the set of respondents prior to the incentive stage. This ability to increase the demographic diversity of our responding sample helped decrease the potential for nonresponse bias in our estimates.

10. ASSURANCE OF CONFIDENTIALITY

NCSES 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, Title 13, Section 9 of the United States Code, and the Cybersecurity Enhancement Act of 2015. The Census Bureau is conducting the NSCG under the authority of Title 13, Section 8 of the United States Code.

The questionnaire cover will include the following confidentiality statement:

The information collected in this questionnaire is solicited under the authority of the National Science Foundation (NSF) Act of 1950, as amended. The U.S. Census Bureau is conducting this survey under the authority of Title 13, Section 8 of the United States Code. The Census Bureau is required by law to keep your information confidential and can use your responses for statistical purposes only. The Census Bureau is not permitted to publicly release your responses in a way that could identify you. Federal law

protects your privacy and keeps your answers confidential (Title 13, United States Code, Section 9). Per the Federal Cybersecurity Enhancement Act of 2015, your data are protected from cybersecurity risks through screening of the systems that transmit your data. Your response is voluntary and failure to provide some or all of the requested information will not in any way adversely affect you. Actual time to complete the questionnaire may vary depending on your circumstances but on the average, it will take about 30 minutes. If you have any comments on the time required for this survey, please send them to the National Science Foundation, 4201 Wilson Blvd., Suite 295, Arlington, VA 22230, Attn: NSF Reports Clearance Officer.

The cover letters will include additional statements in the Frequently Asked Questions section about the Census Bureau's Title 13 as the data collection authority and assurances of confidentiality. The Census Bureau will include the same appropriate notices of confidentiality and the voluntary basis of the survey to respondents contacted during the web phase and CATI phase of the data collection effort.

NCSES 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.

11. JUSTIFICATION FOR SENSITIVE QUESTIONS

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

12. ESTIMATE OF RESPONDENT BURDEN

NCSES estimates that it will contact approximately 123,500 sample persons by web, mail or computer-assisted telephone interviewing as part of the 2017 NSCG collection. Based on experience administering the NSCG interviews, the questionnaire takes an average of 30 minutes to complete. NSF expects the response rate to be 70 to 80 percent. Based on an estimate of approximately 98,800 completed cases, the total burden hours for the 2017 NSCG data collection are 49,400. The total cost to respondents for the 49,400 burden hours is estimated to be $1,425,000. This estimate is based on an estimated median annual salary of $60,000 per NSCG employed respondent. Assuming a 40-hour workweek and a 52-week salary, this annual salary translates to an hourly salary of $28.85. Salary estimates were obtained using data from the 2015 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 2017 NSCG is approximately $13.8 million, which includes survey cycle costs, and NCSES staff costs to provide oversight of the NSCG and coordination with the SDR. The estimate for survey cycle costs is approximately $13.3 million, which is based on sample size; length of questionnaire; administration; overhead; sample design; mailing; printing; sample person locating; web instrument development; telephone interviewing; incentive payments; 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 NCSES 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

The burden impact decreased between the 2015 and 2017 survey cycles because of a slight reduction in overall sample size. The sample size for the 2017 NSCG is 123,500 cases whereas the 2015 NSCG sample size was 135,000 cases. The main explanation for this sample size reduction is the removal of sample cases that originated in the 2010 NSRCG from the survey.

16. SCHEDULE FOR INFORMATION COLLECTION AND PUBLICATION

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

The time schedule for 2017 data collection and publication is currently estimated as follows:

Data Collection	March 2017 – September 2017
Coding and Data Editing	April 2017 – February 2018
Final Edited/Weighted/Imputed Data File	March 2018
NSCG Info Brief	Summer 2018
NSCG Public Use Data File	Summer 2018

17. DISPLAY OF OMB EXPIRATION DATE

The OMB expiration date will be displayed on the 2017 NSCG questionnaires, postal contacts, and the web instrument introduction page.

18. EXCEPTION TO THE CERTIFICATION STATEMENT

Not Applicable.