

**SF-83-1 SUPPORTING STATEMENT:
SECTION A**

for the

2019

National Survey of College Graduates

Contents

A.	JUSTIFICATION.....	3
1.	NECESSITY FOR INFORMATION COLLECTION.....	3
2.	USES OF INFORMATION.....	5
3.	CONSIDERATION OF USING IMPROVED TECHNOLOGY.....	8
4.	EFFORTS TO IDENTIFY DUPLICATION.....	10
5.	EFFORTS TO MINIMIZE BURDEN ON SMALL BUSINESS.....	11
6.	CONSEQUENCES OF LESS FREQUENT DATA COLLECTION.....	11
7.	SPECIAL CIRCUMSTANCES.....	12
8.	FEDERAL REGISTER ANNOUNCEMENT AND CONSULTATION OUTSIDE THE AGENCY.....	12
9.	PAYMENT OR GIFTS TO RESPONDENTS.....	17
10.	ASSURANCE OF CONFIDENTIALITY.....	18
11.	JUSTIFICATION FOR SENSITIVE QUESTIONS.....	19
12.	ESTIMATE OF RESPONDENT BURDEN.....	19
13.	COST BURDEN TO RESPONDENTS.....	19
14.	COST BURDEN TO FEDERAL GOVERNMENT.....	19
15.	REASON FOR CHANGE IN BURDEN.....	20
16.	SCHEDULE FOR INFORMATION COLLECTION AND PUBLICATION.....	20
17.	DISPLAY OF OMB EXPIRATION DATE.....	20
18.	EXCEPTION TO THE CERTIFICATION STATEMENT.....	20

2019 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 2017 (OMB approval number 3145-0141). The current OMB clearance for the NSCG expires 29 February 2020, which does not cover the complete survey cycle for the 2019 NSCG.

For the 2019 NSCG, the following three modifications are being implemented:

1. The sample size is being increased to improve the estimation capability for the total population and, specifically, for foreign-earned doctorate recipients (see details in Section B.1);
2. The contact strategy is being modified (e.g., redesigned postal mailings, emails timed to coincide with postal mailings) based on research findings from the 2017 NSCG (see Section A.8. Contact Strategies Research); and
3. Two methodological experiments are being incorporated in response to a CNSTAT panel’s recommendations (specified at Section A.8. Consultation Outside the Agency): one experiment to explore further automating adaptive survey design techniques and the other to examine how including due dates in contact materials might be effective in encouraging survey completions (see Section B.4. and Appendices H.1, H.2, and I).

1. NECESSITY FOR INFORMATION COLLECTION

In 2010, the America COMPETES Reauthorization Act of 2010¹ established the National Center for Science and Engineering Statistics (NCSES) within 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

¹ Section 505, Pub. L. No. 111-358. See Appendix A.

² The SDR is a repeated cross-sectional 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>.

(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.*

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 is a repeated cross-sectional survey conducted to provide data on the nation’s college graduates, particularly those in the science and engineering (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. As of 2017, the NSCG fully implemented a four-panel rotating panel design, in which every new panel receives a baseline questionnaire, followed by three biennial follow-up questionnaires before rotating out of the survey. (See Supporting Statement B, Section 1, for further details about the panel design and sampling methods.) Sample members are invited to complete the NSCG questionnaire online, on paper, or by phone.

The NSCG 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. This survey 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 complements these data with information on the population of U.S.-degreed doctoral level scientists and engineers. Collectively, the NSCG and SDR provide comprehensive information on the education and employment of the entire U.S. population of scientists and engineers with at least a bachelor’s degree. 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.

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)

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.

The NSCG has a history of seeking survey improvements through methodological experiments, and the 2019 NSCG continues that trend with the inclusion of two studies. Following up on results from the 2017 contact strategies experiment, the 2019 NSCG will incorporate new mailing materials and experiment with further revisions pertaining to response deadlines. Similarly, based on the success of past adaptive survey design experiments, the 2019 NSCG will continue to innovate methods to automate adaptive design techniques. (See Supporting Statement A, Section 3, for further information on the adaptive design study and Supporting Statement B, Section 4, for more details about both studies.)

2. USES OF INFORMATION

The 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.

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 the NSCG and SDR provide up-to-date and nationally representative data, policymakers and researchers use these datasets 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 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, Engineering, and Medicine, 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 National Science Board (NSB) used the combined NSCG and SDR data in its investigation to develop national policies for the S&E workforce⁹;
- 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 professionals¹⁰;

⁹ <https://www.nsf.gov/pubs/2018/nsb20187/nsb20187.pdf>,
<http://nsf.gov/nsb/publications/2015/nsb201510.pdf>, and
<http://www.nsf.gov/nsb/documents/2003/nsb0369/nsb0369.pdf>

¹⁰ http://www.miracoalition.org/images/stories/gac_task_force_report_final-12.18.14.pdf

- The U.S. Small Business Administration used NSCG data to investigate differences in STEM entrepreneurship participation between native-born and foreign-born workers¹¹;
- 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 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,¹² which is a key element of CGS's financial education website – www.gradsense.org.

Uses by NSF

The NSCG data were used extensively in the latest versions of the congressionally mandated biennial reports *Science and Engineering Indicators, 2018* and *Women, Minorities and Persons with Disabilities in Science and Engineering, 2017*. In addition, *Women, Minorities and Persons with Disabilities in Science and Engineering, 2019*, 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:

- *Prevalence of Certifications and Licenses among the College-Educated Population in the United States*, January 2017
- *Immigrants' Growing Presence in the U.S. Science and Engineering Workforce: Education and Employment Characteristics in 2013*, September 2015
- *Characteristics of the College-Educated 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

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, restricted-use licenses, and the Federal Statistical Research Data Centers. The online data tool, available at <https://ncesdata.nsf.gov/sestat/sestat.html>, allows

¹¹ <https://www.sba.gov/sites/default/files/advocacy/rs432tot-Immigrant-STEM-Entrepreneurs.pdf>

¹² <http://www.gradsense.org/gradsense/methodology>

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://ncesdata.nsf.gov/datadownload/>.

Since 2005, NCSES has distributed over 1,000 copies of the 1993 NSCG public-use files, over 1,700 copies of the 2003 NSCG public-use files, and over 1,400 copies of the 2010 NSCG public-use files to researchers in government, academia, and professional societies. Since their release in April 2015, over 1,500 copies of the 2013 NSCG public-use files have been downloaded. And, the 2015 NSCG public-use files have been downloaded over 1,600 times since their release in January 2017. The 2017 NSCG public-use files will be available soon. 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.

In addition to the users of the public-use files, there are currently 25 restricted-use licensees with access to the combined NSCG, SDR, and National Survey of Recent College Graduates (NSRCG)¹³ microdata files under a licensing agreement with NCSES.

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

- Santacroce, A. (2018). "Determining Strategies for the Embry-Riddle Aeronautical University College of Engineering Faculty to Use to Increase the Retention Rate of Women in their Undergraduate Engineering Programs." *The Compass* 1(5): Article 6.
- Amuedo-Dorantes, C., D. Furtado and H. Xu (2018). Did OPT Policy Changes Help Steer and Retain Foreign Talent into STEM? *17th IZA/SOLE Transatlantic Meeting of Labor Economists (TAM)*. Inning, Germany. http://conference.iza.org/conference_files/transatlantic_2018/xu_h23885.pdf
- Sassler, S., K. Micheltore and K. Smith (2017). "A Tale of Two Majors: Explaining the Gender Gap in STEM Employment among Computer Science and Engineering Degree Holders." *Social Sciences* 6(3): 69.
- McClough, D. and M. E. Benedict (2017). "Not All Education Is Created Equal." *The American Economist*: 1-22.
- Lin, L., P. Christidis and K. Stamm (2017). Salaries in Psychology: Findings from the National Science Foundation's 2013 National Survey of College Graduates. Washington, DC, American Psychological Association, Center for Workforce Studies. <http://www.apa.org/workforce/publications/2013-salaries/report.pdf>
- Hunt, J. (2017). "Immigrant patents boost growth." *Science* 356(6339): 694-697.
- Islam, A., F. Islam and C. Nguyen (2017). "Skilled Immigration, Innovation, and the Wages of Native-Born Americans." *Industrial Relations* 56(3): 459-488.

¹³ Through 2010, the NSRCG complemented 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.

- Sakamoto, A. and S. X. Wang (2016). "Occupational and Organizational Effects on Wages among College-educated Workers in 2003 and 2010." *Social Currents* 4(2): 175-195.
- Bender, K. A. and K. Roche (2016). "Self-employment and the paradox of the contented female worker." *Small Business Economics* 47(2): 421-435.
- *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
- *Highly Skilled Migrants: Risks and Hedging Mechanisms*, Texas Tech University, 2016

3. CONSIDERATION OF USING IMPROVED TECHNOLOGY

The data for the 2019 NSCG will be collected by the U.S. Census Bureau under an interagency agreement between NCSES and the Census Bureau. The 2019 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). The 2019 NSCG web instrument will be optimized for use on mobile devices, creating a better experience for mobile device users and, thereby, reducing survey breakoffs and the possibility of measurement errors.

Adaptive Design and Intermittent Data Processing

The 2019 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 focus was on increasing the automation of existing capabilities and predicting the effects of data collection interventions. The Census Bureau improved the “flow processing” (i.e., intermittent editing, imputation, and weighting of incoming response data), which allowed the survey team to monitor quality measures throughout data collection.

The 2019 NSCG will build upon the lessons learned in prior rounds in an attempt to fully automate adaptive interventions based on predefined survey goals. More detail about the 2019 NSCG adaptive design experiment is provided in Supporting Statement B, Section 4. We will employ roughly the same sample sizes as the 2015 and 2017 adaptive design experiments 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,¹⁴ cost, and effect on key estimates.

Unified Tracking System

In 2019, the NSCG will continue its use of the Census Bureau’s Unified Tracking System (UTS) to assist in various aspects of survey management. Since 2013 the UTS has provided a full

¹⁴ R-indicators are useful, in addition to response rates and domain estimates, for assessing the potential for nonresponse bias. R-indicators are based on response propensities calculated using a predetermined balancing model (“balancing propensities”) to provide information on how different the respondent population is compared to the full sample population, as well as which variables in the predetermined model are driving the variation in nonresponse.

contact history report for the NSCG, giving survey managers a single place to view all contacts integrated from all three survey modes along with the outcomes of those contacts. This contact history reporting system enables the examination of contact strategies in various ways. For example, if respondents call in to check on the status of their response, NSCG staff are able to access the respondents' contact history quickly and easily. In addition, this report provides an easily 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. In 2017, this contact report was enhanced by the integration of the previously mentioned IMB data.

For the 2019 NSCG, the UTS will continue to provide daily updates for R-indicators analysis at the cohort-level, so that survey management can understand how data collection operations affect representativeness. Additionally, the UTS will provide two reports to monitor IMB data. These reports will focus on the difference between the dates provided by the Census Bureau's National Processing Center (NPC) and IMB-provided dates for survey monitoring purposes. For outgoing mailings, the report will show the lag between the scheduled mail date 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 at NPC. Both reports will have these data broken down by mailing geographies. These reports help us understand the relationship between when sample persons receive their mail and when they respond to survey requests, thus helping us anticipate response relative to mailout operations.

Finally, a UTS report that documents the interactions of the NSCG sample with the web instrument will be continued for the 2019 NSCG. This report provides 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. This report allows such valuable web paradata to be monitored throughout the data collection period.

4. EFFORTS TO IDENTIFY DUPLICATION

Duplication, in the sense of a similar data collection, does not exist. No other data collection captures all components of scientists and engineers in the United States. Data from the Current Population Survey provides occupational estimates but does not collect information on degree field for postsecondary 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.

The NSCG and ACS both collect demographic information including gender, race, ethnicity, marital status, citizenship status, and veteran status. 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. Because linkage between demographic, education, and employment information is needed for the analyses used in NSF's congressionally mandated reports, all 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 360 individuals selected for sample in both the 2019 NSCG and the 2019 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 2013 survey cycle.

Given recent changes to the NSCG questionnaire content, there are noticeable differences between the NSCG and SDR. Information collected on the NSCG but not on the SDR includes attainment of certifications and licenses, financial support for education, community college enrollment, and veteran status. 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 2019 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 on a biennial basis since 1993. The degradation of any component jeopardizes the integrity and value of these combined surveys to provide comprehensive information on the S&E workforce.

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 likely impacts would be a higher attrition rate, higher potential for nonresponse bias, and less reliable estimates.

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 12 June 2018. 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 include measures of sexual orientation and gender identity on the NSCG and on other NCSES surveys (specifically, the SDR and the Survey of Earned Doctorates).

NCSES informed the commenters that it shares their interest in improving federal data collections and providing reliable measures for important segments of the population. Furthermore, NCSES described its process for evaluating possible questionnaire additions, including the extensive experimentation involved and the time and resources required. Finally, NCSES informed the commenters that it is initiating research to evaluate these measures and does not intend to include them in the 2019 NSCG.

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, 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 and methodology, data collection, data processing, data integration, data dissemination, and data promotion.

At the end of its review, the panel issued 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.¹⁵ In the 2019 NSCG, NCSES seeks to address at least three of these recommendations. Recommendation 4.2 suggests evaluating modeling techniques to determine whether they could be used to more efficiently identify foreign-trained doctorate holders. As a first step, NCSES plans to increase the number of foreign doctorates in the 2019 NSCG sample so that sufficient data will be available to conduct such an evaluation. Recommendation 5.1 suggests continuing research into optimal contacts, response modes, and incentives, which we plan to do with an experiment exploring mailout strategies (see Appendix I for details). Lastly, Recommendation 5.5 suggests expanding the use of adaptive design to reduce nonresponse bias and control costs, which we will do with an adaptive design experiment that seeks to automate the identification and selection of cases for interventions (see Appendix H for details).

- 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 continuing research examining the potential use of administrative records as an NSCG sampling frame

¹⁵ <https://www.nap.edu/catalog/24968/measuring-the-21st-century-science-and-engineering-workforce-population-evolving>

source. To date, CARRA has documented the strengths and weaknesses of possible frame sources, conducted research to assess the quality and viability of these sources, and has begun a detailed investigation of two sources that show the most potential promise: The National Student Clearinghouse (NSC) and the Census Bureau's Master Address File Auxiliary Reference File (MAF-ARF). CARRA is currently in the process of examining an extract of the NSC data to analyze its coverage and potential usability compared to the American Community Survey (ACS). For the MAF-ARF, CARRA is examining whether the MAF-ARF contact information could supplement the ACS information for NSCG respondent locating purposes.

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

CARRA is also continuing research using administrative records and third-party data sources to compare with NSCG data to inform measurement error issues for NSCG survey estimates. The two projects underway are (1) An evaluation of earnings data and employment history data from the Census Bureau's Longitudinal Employer-Household Dynamic (LEHD) program and (2) An evaluation of demographic and certification data from a variety of data sources including IRS 1040s, IRS 1099s, NSC, the Census Bureau's Numident File, and the Center for Medicare and Medicaid Services' National Plan and Provider Enumeration System (NPPES). The results from this research will inform measurement error discussions and may provide guidance on the feasibility of using administrative records for NSCG survey replacement or supplementation.

- 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. Staff then evaluated the 2017 paradata to determine whether the enhancements were effective. Overall, the instrument changes between 2015 and 2017 resulted in shorter completion times, fewer breakoffs, and fewer respondents moving backward in the instrument (clicking the Previous button).

- Contact Strategies Research

The Census Bureau's Demographic Statistical Methods Division (DSMD) conducted an experiment in the 2017 NSCG cycle 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, sample

representativeness, and key estimates. The experiment tested three treatments in a fully factorial design: a new mailing strategy, the inclusion of an infographic, and a limit on CATI follow-up calls. The new mailing strategy reduced the number of mailings, used different types of mailings (i.e., perforated letters, tabbed postcards), and included letters with less text that were designed to attract readers to key information and inform them of how their data are used. Additionally, email reminders were sent immediately following a mail contact instead of as stand-alone reminders. The infographic was also intended to inform sample cases to how their data are used and the importance of their response. The final treatment limited CATI calls to 10 per case. Results showed the most successful treatment combination, for both new and returning sample members, was the new mailing materials, no infographic, and no call limit. This combination resulted in nominally higher response rates, the highest R-indicator for the new sample, the second highest R-indicator for the returning sample, and no significant differences for any of the 14 key estimates.

The call limit of 10 CATI contacts was successful in reducing cost by almost \$8 per case, but there was a high number of significant differences in key estimates for both new and returning sample members. Instead of implementing a call limit of 10 in the 2019 NSCG cycle, DSMD will evaluate setting parameters for specific CATI outcomes at the point where additional calls are no longer productive. This will result in new parameters for the outcomes investigated where calls will cease when the parameter is reached, thus providing a method of reducing costs while not impacting estimates.

- 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.” Since that time, NCSES staff have collaborated 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 the outreach and collaboration efforts related to the NSCG’s adaptive design efforts.

- In March 2018, Census Bureau and NCES participated in an invited panel at the Federal Committee on Statistical Methodology. The panel topic was “Nonresponse Bias Studies in the Federal Government.” The panel focused on the state of nonresponse bias studies, suggested necessary improvements, and the ability of various data collection methods, including adaptive design, to assist in reducing nonresponse bias.

- 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 in 2016 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. In 2017, an additional seminar for NCES was jointly presented by Census Bureau and the University of Michigan and focused on the types of regression models commonly used in both static and dynamic adaptive survey designs.
- 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. This network conducted bi-annual meetings in 2013, 2015, and 2017 and presentations related to NCSES- and NCES-sponsored surveys were featured at each of these meetings. 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."
- 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.

Adaptive Design Publications and Presentations Using NSCG Data (2016 – Present)

Publications:

Coffey, S., Reist, B., Miller, P. (2018). Interventions on Call: Dynamic Adaptive Design in the National Survey of College Graduates. *Journal of Survey Statistics and Methodology*. *Under Review*.

Presentations:

Coffey, S., Zotti, A. (2017). Logistic Regression: Practical Examples in Experimental Design. Presentation at the National Center for Education Statistics Monthly Seminar Series. August 2017. Washington, DC.

Coffey, S. (2017). More Information is Better! Where Can We Get It and How Can We Use It? Presentation at the 2017 Joint Statistical Meetings. July 2017. Baltimore, MD.

Coffey, S. (2017). Adaptive Design in the NSCG: Insights from 2015 and Developments for 2017. Presentation at the 2017 AAPOR Conference. May 2017. New Orleans, LA.

Coffey, S. (2016). Concurrent Analysis and Estimation System – 2 Proofs of Concept. Presentation at the 2016 FedCASIC Conference. May 2016.

Coffey, S. (2016). Using System Paradata to Target and Evaluate Data Collection Operations. Presentation at the 2016 WSS Paradata Mini-conference. April 2016.

Coffey, S. (2016). Improving Data Collection Through Targeting and Adaptive Design. Presentation at the University of Michigan Survey Research Center PRS Seminar. March 2016.

- 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 the most efficient manner to create weights for multiple panel estimation within the NSCG. Jean Opsomer and Jay Breidt (Colorado State University) are the principal investigators for this research.
- To address improvements to data quality and increases in nonresponse trends, NCSES funded research to examine these issues. Jolene Smyth and Kristen Olson (University of Nebraska – Lincoln) are 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^{16,17} 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 2019 NSCG will follow the procedures used in the 2015 and 2017 survey cycles.

As was the case in the 2017 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 2019 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

¹⁶ 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.

¹⁷ 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.

sampled case's base weight and predicted response propensity. We expect to offer \$30 debit card incentives to approximately 13,400 of the 67,000 new sample cases included in the 2019 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, $\hat{\rho}_{Ri}$.

In addition, using the findings from the 2013 NSCG incentive conditioning study and following our procedures from the 2015 and 2017 NSCG, we plan to offer a \$30 prepaid debit card incentive to past incentive recipients at week 1 of the 2019 NSCG data collection effort. As a result, we expect to offer \$30 debit card incentives to approximately 11,000 of the 81,000 returning sample members.

The \$30 incentive amount proposed for use in the 2019 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 from 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 Reports Clearance Officer, Office of the General Counsel, National Science Foundation, 2415 Eisenhower Ave., Alexandria, VA 22314.

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 148,000 sample persons by web, mail or computer-assisted telephone interviewing as part of the 2019 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 118,400 completed cases, the total burden hours for the 2019 NSCG data collection are 59,200. The total cost to respondents for the 59,200 burden hours is estimated to be \$1,793,077. This estimate is based on an estimated median annual salary of \$63,000 per NSCG employed respondent. Assuming a 40-hour workweek and a 52-week salary, this annual salary translates to an hourly salary of \$30.29. Salary estimates were obtained using data from the 2017 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 2019 NSCG is approximately \$15.5 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 \$14.9 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 increased between the 2017 and 2019 survey cycles because of an increase in overall sample size. The sample size for the 2019 NSCG is 148,000 cases whereas the 2017 NSCG sample size was 124,000 cases. The main explanation for this sample size increase is an attempt to account for the diminishing response rate in subsequent cycles and to increase the number of foreign-trained doctorates in the sample.

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 2019 data collection and publication is currently estimated as follows:

Data Collection	February 2019 –August 2019
Coding and Data Editing	March 2019 – January 2020
Final Edited/Weighted/Imputed Data File	February 2020
NSCG Info Brief	Summer 2020
NSCG Public Use Data File	Summer 2020

17. DISPLAY OF OMB EXPIRATION DATE

The OMB expiration date will be displayed on the 2019 NSCG questionnaires, postal contacts, and the web instrument introduction page.

18. EXCEPTION TO THE CERTIFICATION STATEMENT

Not Applicable.