

2017–2019
Survey of Graduate Students
and
Postdoctorates in Science
and Engineering

OMB Supporting Statement
Section A

September 2017

TABLE OF CONTENTS

Section

n

| | |
|---|----------|
| A. JUSTIFICATION..... | 1 |
| A.1 Necessity for Information Collection..... | 1 |
| A.2 Uses of Information..... | 2 |
| A.2.1 Federal Uses..... | 2 |
| NSF Uses..... | 3 |
| Other Federal Uses..... | 3 |
| A.2.2 Use by Academic Institutions..... | 4 |
| A.2.3 Use by the Carnegie Foundation..... | 4 |
| A.2.4 Use by the Professional Societies..... | 4 |
| A.2.5 Use in Research..... | 4 |
| A.2.6 Media Uses..... | 5 |
| A.3 Consideration of Using Improved Technology..... | 5 |
| A.4 Efforts to Identify Duplication..... | 6 |
| A.5 Efforts to Minimize Burden on Small Business..... | 8 |
| A.6 Consequences of Less Frequent Data Collection..... | 8 |
| A.7 Special Circumstances..... | 9 |
| A.8 Federal Register Announcement and Consultations Outside the Agency..... | 9 |
| A.8.1 GSS Institution Site Visits..... | 10 |
| A.8.2 GSS Coordinator Survey..... | 10 |
| A.8.3 GSS Pilot Data Collection..... | 11 |
| A.8.4 Other Consultations..... | 12 |
| A.9 Payment or Gifts to Respondents..... | 12 |
| A.10 Assurance of Confidentiality..... | 12 |
| A.11 Justification for Sensitive Questions..... | 12 |
| A.12 Estimate of Respondent Burden..... | 12 |
| A.13 Cost Burden to Respondents..... | 16 |
| A.14 Cost to the Federal Government..... | 16 |
| A.15 Program Changes or Adjustments in Burden..... | 17 |
| A.16 Publication Plan and Project Schedule..... | 17 |
| A.17 Display of OMB Expiration Date..... | 17 |
| A.18 Exceptions to the Certification Statement..... | 17 |

EXHIBITS

Number

| | |
|--|----|
| Exhibit 1. Burden Results for 2015 GSS and 2016 GSS Pilot Data Collection..... | 14 |
| Exhibit 2. Expected Composition of the 2017 GSS Frame..... | 14 |
| Exhibit 3. Burden Estimates for the 2017 GSS..... | 15 |
| Exhibit 4. Burden Estimates for the 2018 GSS..... | 15 |
| Exhibit 5. Burden Estimates for the 2019 GSS..... | 15 |
| Exhibit 6. Total Burden Estimates for 2017–19 GSS..... | 16 |
| Exhibit 7. Annual GSS Survey Federal Government Estimated Costs..... | 16 |

LIST OF ATTACHMENTS

- Attachment 1. GSS Instrument Screen Shots
- Attachment 2. Federally Funded Research and Development Centers (FFRDCs) Postdoc Survey Instrument Screen Shots
- Attachment 3. First Federal Register Notice
- Attachment 4. GSS Schedule
- Attachment 5. GSS Code List—Complete List of GSS Eligible Fields and Codes
- Attachment 6. GSS Data Collection Plan
- Attachment 7. GSS Data File Upload Templates
- Attachment 8. CIP and GSS Code Crosswalk
- Attachment 9. GSS Worksheet
- Attachment 10. List of Changes to GSS Taxonomy
- Attachment 11. GSS Taxonomy Tool Prototype
- Attachment 12. GSS Coordinator Survey Report Summary

A. JUSTIFICATION

This submission requests a three-year reinstatement of the previously approved OMB clearance for the National Center for Science and Engineering Statistics (NCSES) (within the National Science Foundation (NSF)) and the National Institutes of Health's (NIH's) Survey of Graduate Students and Postdoctorates in Science and Engineering (GSS). The GSS is an annual survey that was last conducted in fall 2016. The OMB clearance for the GSS will expire on November 30, 2017. With this clearance package, NSF requests approval to collect data for the 2017, 2018, and 2019 survey cycles.

The GSS is the only annual national survey that collects information on the characteristics of graduate enrollment for specific science, engineering, and health (SEH) disciplines at the departmental level. It also collects information on graduate enrollment by race and ethnicity, citizenship, sex, sources of support, and type of support; information on postdoctorates (postdocs) by citizenship, sex, sources of support, and type and origin of doctoral degree; and information on other doctorate-holding nonfaculty researchers (NFRs) (see Attachment 1 for screenshots of the GSS instrument). The GSS has been conducted by NCSES annually since 1972. Additional financial support for the GSS is provided by the NIH.

The GSS is a census of all organizational "units" (departments, programs, research centers, and health care facilities) in SEH fields within eligible academic institutions in the United States that grant research-based master's or doctorate degrees. The survey collects aggregate information on graduate students enrolled in these units, as well as postdocs and NFRs working within these institutions. As a part of the GSS, NCSES also periodically surveys Federally Funded Research and Development Centers (FFRDCs) to collect information on the postdocs such as race/ethnicity, sex, citizenship, source of support, area of research (see Attachment 2 for screenshots of the FFRDC Postdoc Survey instrument).

A.1 Necessity for Information Collection

In 2010, the America COMPETES Reauthorization Act of 2010¹ established, the previously named Science Resources Statistics division as NCSES within the 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

¹ Section 505, Pub. L. No. 111-358.

useful to practitioners, researchers, policymakers, and the public...” Information obtained through the GSS is critically important to NCSES’s ability to measure science and engineering resources in the United States. Furthermore, the GSS data serve as the nation’s only source of comprehensive graduate enrollment information for specific SEH disciplines at the departmental level. These data are solicited under the authority of the NSF Act of 1950, as amended, and are central to the analysis presented in a pair of congressionally mandated reports^{2,3} published by NSF, the *Science and Engineering Indicators* and the *Women, Minorities, and Persons with Disabilities in Science and Engineering*.

A.2 Uses of Information

A.2.1 Federal Uses

NSF and NIH extensively use the information on the number and characteristics of students currently enrolled in graduate SEH programs and of persons engaged in postdoctoral programs to assess future stock of trained SEH personnel. A variety of more general information needs are met through the annual release of data in electronic format. NSF publishes a short InfoBrief and a set of statistical tables, *Survey of Graduate Students and Postdoctorates in Science and Engineering Data Tables*, available on the NCSES website.

Data from the GSS are also available as public use files, and on the Web through the WebCASPAR (Computer Aided Science Policy Analysis and Research) system (<https://ncesdata.nsf.gov/webcaspar/>). WebCASPAR is an institution-based data system, and it contains institutional and summary data from all NCSES’ academic sector surveys for all institutions offering graduate-level instruction and/or maintaining research and development (R&D) activity in SEH fields.

Primary uses of the GSS data include: reviewing changing enrollment levels to assess the effects of NSF initiatives; tracking student support patterns; and analyzing participation in SEH fields by targeted groups for all disciplines or for selected disciplines and for selected groups of institutions. Program officers check departmental and institutional records, including data from the GSS and the IPEDS surveys, to determine department eligibility for NSF programs targeted to special populations or instructional programs.

² 42 U.S. Code § 1863(j)(1)

³ 42 U.S. Code § 1885(a), 1885(d)

A recent study examined the quality of the non-faculty researcher data in the GSS. The working paper, “Examining the Reporting of Nonfaculty Doctorate Researchers in the Survey of Graduate Students and Postdoctorates in Science and Engineering” (<https://nsf.gov/statistics/2015/ncses15201/>), was released in 2015. This working paper examined the consistency of NFR reporting in GSS since 2010, and analyzed key reporting patterns and attempts to validate the GSS data. The study found that the 2010–12 NFR counts are much more reliable and accurate in gauging the size and distribution of this population across the GSS academic institutions than prior estimates.

NSF Uses

Special tabulations from the GSS data constitute a key resource in meeting policy and program information needs of the Foundation. Major examples of GSS data uses are in the Foundation’s two congressionally mandated biennial reports, *Science and Engineering Indicators* and *Women, Minorities, and Persons with Disabilities in Science and Engineering*.

The GSS is one of four NCSES surveys whose microdata are combined into an integrated database to produce the publication *Academic Institutional Profiles*. The other three surveys are (1) the SED; (2) the Higher Education Research and Development (HERD) Survey; and (3) the Survey of Federal Science and Engineering Support to Universities, Colleges, and Nonprofit Institutions. As explained in the next section, these data are further integrated with institutional data from other NCSES surveys and with surveys conducted by NCES. Together these data provide policy makers with information on the role of higher education in the context of the national R&D effort.

Other Federal Uses

Data derived from the GSS are routinely provided to Congress and to various agencies of the Executive Branch. Recent examples of provided data include::

- Data on graduate SEH enrollment provided annually to NCES for comparison purposes and are published in the *Digest of Education Statistics*.
- Data in specially prepared GSS tabulations used by the NIH to answer specific questions to help their agencies prepare budgets and conduct program evaluation studies.

A.2.2 Use by Academic Institutions

The surveyed institutions themselves are major users of the GSS data. Institutions use the NCSES' GSS data reports or the WebCASPAR system to study selected groups of peer institutions for planning and comparative purposes. They combine the NCSES data with information from state and local governments on institutions in their geographic areas. Institutions also use the comparative data to review the strength of their own programs on the basis of factors such as support of students by various federal agencies and progress in reaching special target populations.

A.2.3 Use by the Carnegie Foundation

Data from the GSS are used by The Carnegie Foundation for the Advancement of Teaching in developing the Carnegie Classification of Institutions of Higher Education. The foundation uses the GSS data on nonfaculty research staff with doctorates as one component of the "research activity" measure constructed for doctorate-granting universities (for more detail see: <http://carnegieclassifications.iu.edu/methodology/basic.php>).

A.2.4 Use by the Professional Societies

Data users include American Association of Colleges of Nursing, American Association of Universities, American Chemical Society, American Council of Education, American Geological Society, American Institute of Physics, American Physical Society, American Society for Engineering Education, Association of American Medical Colleges, Association of International Educators, Commission on Professionals in Science and Technology, Computing Research Association, Council of Graduate Schools, Federation of American Societies for Experimental Biology, and the National Postdoctoral Association. Generally, associations use GSS data to monitor trends in enrollment by field of study, and many are also interested in tracking the numbers of postdoc and NFRs.

A.2.5 Use in Research

Researchers studying policy issues relating to the SEH labor pipeline, the gender gap in SEH fields, and financial support for SEH training and research activity have used GSS data in their investigations. Recent research studies using GSS data have shown that enrollments in biomedical sciences graduate and postdoc programs are more responsive to changes in the

availability of NIH fellowships, traineeships, and research funding than they are to fluctuations in biomedical scientists' wages at the time of enrollment.⁴ The data also show that increases in federally-funded R&D are associated with increases in the number of postdoc positions.⁵

After the NIH Advisory Committee to the Director recommended a greater emphasis on traineeships and fellowships over research assistantships, a study used GSS data to demonstrate that biological and biomedical sciences programs that increased their NIH-funded traineeships and fellowships saw increases in their graduate enrollments, primarily among U.S. citizens and permanent residents.⁶ GSS data also were used to forecast trends in U.S. graduate student enrollment rates, including those for underrepresented minorities, relative to foreign student enrollment rates.⁷

A.2.6 Media Uses

Enrollment of graduate students in S&E fields are well reported by the press, including *Forbes*, *The Chronicle of Higher Education*, and *Science*. A recent example of the use of GSS postdoc data is a *Science* article on December 9, 2015, entitled "[The case of the disappearing postdocs](#)". Another recent article in *Forbes* used GSS data to address high skill immigration policy in a September 29, 2015 article entitled, "[Should U.S. Companies Ignore Over 70% of the Potential Tech Labor Force?](#)"

A.3 Consideration of Using Improved Technology

NCSES has engaged in a process of making improvements to GSS, involving technical innovations to increase the utility of the data collected and reduce the burden for respondents. During the 2016 GSS data collection, a stratified random sample of 80 institution coordinators was selected for a Pilot survey to test the feasibility of implementing the redesign changes. Pilot coordinators were instructed to report master's and doctoral student data separately, and upload their data file on the GSS web system using CIP codes instead of GSS codes. Two data upload

⁴ Blume-Kohout, Margaret E., and John W. Clack. "Are graduate students rational? evidence from the market for biomedical scientists." *PloS ONE* 8.12 (2013): e82759.

⁵ Howard H. Garrison, Louis B. Justement, and Susan A. Gerbi. "Biomedical Science Postdocs: An End to the Era of Expansion", *FASEB Journal* (2015).

⁶ Blume-Kohout, Margaret E., and Dadhi Adhikari. "Training the scientific workforce: Does funding mechanism matter?" *Research Policy* 45.6 (2016): 1291-1303.

⁷ Sanfilippo, Antonio, Chase Dowling, and Sofiane Abbar. "Do International Students Displace US Students in the Pursuit of Higher Degrees in Science and Engineering? A Forecasting Analysis." *Higher Education Policy* 29.3 (2016): 335-354.

options were offered to the Pilot coordinators and the uploaded data were automatically loaded into the GSS web instrument. Pilot coordinators were allowed to review and edit the data before submitting to NSF. Of 80 Pilot coordinators, 76 uploaded their data as part of the 2016 GSS pilot survey (see A.8.3 for more information).

NCSES expects to leverage these technical innovations in the proposed 2017 GSS data collection methods, which are based on the results of the research and 2016 GSS pilot survey. Therefore, in 2017, all schools will be asked to use one of the two following data uploading options:

- Upload a file containing de-identified individual records that the Web system automatically aggregates to the unit-level format and then populates the appropriate cells in the GSS survey.
- Upload a file that contains an Excel macro program that aggregates individual-level data into unit-level data locally. This option is available for SCs who do not wish to transmit individual-level data over the Internet.

NCSES expects the expansion of available upload options to increase the number of SCs that supply GSS data through file uploads and will reduce the overall burden of completing the survey through the web instrument.

A.4 Efforts to Identify Duplication

NCSES staff consults regularly with other federal agencies and private organizations to prevent duplication of data collection activities and to stay abreast of changes in other surveys. Such consultations take place with the NCES, the Council of Graduate Schools (CGS), and others. Specific surveys conducted by these groups will be discussed below. In addition, NCSES staff participate in a variety of NCES-related activities, including serving on the 2010 CIP Working Group and Technical Review Panels. The routine data uses of the federal agencies described in Section A.2.1 have largely determined the content of the GSS questionnaire.

Only the GSS collects the following information at the level of detailed SEH fields of study:

- For full-time graduate students, aggregate counts by
 - Sources of major financial support (federal agencies, institutions, self-support, etc.)
 - Mechanisms of major financial support (fellowships, teaching assistantships, etc.)
 - gender

- citizenship
- enrollment status (full-time or part-time; first time)
- race/ethnicity background of U.S. citizens
- For part-time graduate students, aggregate counts by
 - gender
 - citizenship
 - race/ethnicity background of U.S. citizens
- For postdocs, aggregate counts by
 - sources of major financial support
 - Mechanism of major financial support
 - gender
 - citizenship
 - type of doctoral degree
 - doctoral degree origin
- For NFRs, aggregate counts by
 - gender
 - type of doctoral degree

Because the data are collected from all eligible institutions with graduate SEH departments, data are available at the detailed field of study by institutional characteristics, such as highest degree granted, geographical location, type of control (public or private), or any other special grouping (medical schools, historically black colleges and universities, land-grant institutions, etc.) as well as by rankings on various characteristics (foreign enrollment, minority enrollment, field-specific enrollment, etc.)

Some graduate enrollment data are collected by other organizations, either federal or private, but none of the other data collection efforts contain the detailed field distribution that is required for analyses and provides the necessary data for NSF and NIH. IPEDS, for example, collects race and ethnicity data every 2 years for only nine select fields (of which four are within the NSF definition of science and engineering, but are at more general level than is collected for GSS). The IPEDS annual fall enrollment data collected by race and ethnicity category are not reported by the field, and hence, they do not provide a viable substitute for the race and ethnicity data collected in the GSS. No data are collected on source of support or postdocs and NFRs.

The CGS conducts an annual survey of graduate enrollment in cooperation with the Graduate Records Examinations (GRE) Board, surveying 776 institutions in 2015 that were

members of the CGS or one of the four regional graduate school associations—the Conference of Southern Graduate Schools, the Midwestern Association of Graduate Schools, the Northeastern Association of Graduate Schools, and the Western Association of Graduate Schools. The survey had a response rate of 80%, with 617 schools responding. The survey collects data by 51 fine fields of study using the GRE discipline codes as its taxonomy, type of institutional control, and highest level of degree offered, but has no data on source of financial support. It also collects information on postbaccalaureate and post-master’s certificates and applications to graduate schools. Only the GSS maintains detailed data grouped into ninety-seven fine fields of study on all SEH degree fields at all eligible institutions and institution-provided data on source of financial support.

A number of surveys are conducted by other professional societies or by groups of institutions, and are limited to a single field or group of related fields or to institutions that are members of the organization. These surveys may collect far more detailed data on the fields of interest to the organization conducting the survey, and may even collect data on topics not covered by the GSS (e.g., on undergraduate enrollment), but they do not provide compatible data on all SEH fields, nor do they often address the issue of types and sources of financial support for graduate students.

A.5 Efforts to Minimize Burden on Small Business

Not applicable. The GSS does not collect information from small businesses.

A.6 Consequences of Less Frequent Data Collection

A less frequent survey cycle would have several serious consequences. First, there would be the loss of information. Because of the data uses described previously, biennial or less frequent collection means that data users would be unable to access current information. Collecting the GSS annually also increases the value of the data for monitoring trends, particularly the effects of dramatic changes in the larger context. Minor shifts in enrollment trends are monitored as early indicators of likely future changes in the supply of SEH professionals.

Other examples of trend monitoring are changes in the foreign graduate student enrollment and postdoc employment counts that correspond to the events such as September 11,

2001, the 2007-2009 Great Recession, and immigration policy changes. Less than annual data collection may not capture such changes or reveal the inflection point of a changing trend. Following the September 11, 2001, the release of the GSS fall enrollment data was eagerly anticipated to examine trends in SEH graduate enrollment by foreign visa holders. Foreign student enrollment did not drop immediately (i.e., in 2001), and the trends varied by several years for first-time enrollment and total enrollment. Those nuances would have been lost if the data had not been collected every year.

Annual collection also helps reduce respondent burden. Most colleges and universities have automated record keeping systems, facilitating their ability to respond to the GSS on an annual cycle. These automated record systems considerably reduce the time required to assemble and report information needed for the GSS related to graduate enrollment by field, demographics, postdoctoral appointments, and sources and mechanisms of support, etc. Thus, because the database and software are retained, kept current, and easily accessed, collecting consistent data annually considerably reduces respondent burden for academic institutions with automated data systems.

Annual collection also helps to maintain contacts with the SCs within institutions. Having this continuity helps the SCs maintain their databases and, therefore, maintain the quality of the data.

A.7 Special Circumstances

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

A.8 Federal Register Announcement and Consultations Outside the Agency

The Federal Register notice was published on April 26, 2017 (see Attachment 3). No public comments were received.

As described in the next sections, in the past three years, several consultations with the respondents have taken place to examine different aspects of the GSS data collection and to inform the changes introduced in 2017.

A.8.1 GSS Institution Site Visits

Between July and November 2015, NCSSES conducted visits to seven areas of the country to follow up on the data reporting experiences and capabilities of institutions participating in the GSS. The visits included 23 universities, one FFRDC, and two university system offices. The discussions were designed to identify ways to minimize institutional reporting burdens and obtain input from institutions about the specific changes considered in the GSS, including separate reporting of master's and doctoral student data, reporting data by CIP code, and use of data uploads for GSS reporting. The results from the site visits were encouraging as all visited SCs indicated that they would be able to distinguish master's students from doctoral students, and report student data using CIP. Most of the SCs expressed interest in the data upload feature and thought that most of the GSS data could be extracted from centralized institutional databases and subsequently formatted for data upload.

A.8.2 GSS Coordinator Survey

To explore the feasibility of implementing the proposed changes to the GSS, a survey was administered to all GSS SCs that participated in the 2015 GSS data collection (see Attachment 12). Conducted in July 2016, a total of 840 SCs were invited to take the survey, and 676 participated, for a total response rate of 80.5%. The survey included questions on the feasibility of separate reporting of master's and doctoral student data and of using CIP codes to report data by academic discipline. The survey data showed:

- CIP codes are available for the majority of academic units in which graduate students are enrolled. Among SCs responsible for reporting graduate student data, 88.6% stated CIP availability for some or all of the academic units of their institution. Only 1.8% indicated that CIP codes were not available, while 8.7% did not know.
- The majority of those with responsibility for reporting student data indicated that using CIP codes to report student data would be a neutral or beneficial change for them. About one-third (34.5%) of respondents indicated the estimated burden for using CIP codes would be about the same as using GSS codes, while 32.4% indicated CIP would require less effort compared with the current taxonomy.
- The majority of respondents at institutions that offered both master's and doctoral degrees state that their intuitional records allow them to distinguish between the two degree types: 85.0% indicated this was possible for most programs, while an additional 12.3% said it was possible for some programs.
- Of the responding SCs who used the data upload feature for 2015 GSS, a large majority (86.0%) found the feature somewhat or very easy to use.

A.8.3 GSS Pilot Data Collection

As part of the 2016 GSS data collection, a parallel, Pilot data collection was conducted with a stratified random sample of 80 SCs. The Pilot data collection differed from the regular data collection in the following ways:

- Separate reporting of enrollment and financial support data for master's and doctoral students
- Collecting data based on the CIP codes as a disciplinary field, instead of GSS codes
- Expanding the institutional use of data upload options for data submission instead of the manual entry of data in the GSS web instrument

The sampling strata used to select the sample of 80 consisted of four, non-mutually exclusive groups:

- 15 SCs that uploaded in 2015 (Uploaders)
- 15 SCs that reported only master's degree students in 2015 (Master's Only)
- 25 SCs that reported both master's and doctoral students in 15 or fewer organizational units (Small Reporters)
- 25 SCs that reported both master's and doctoral students in over 15 organizational units (Large Reporters)

These groups were chosen to provide a breadth of characteristics expected to be relevant to the changes being introduced, as well as a breadth of characteristics expected to be relevant in estimating response burden. It should be noted that all SCs in the Uploader stratum also represented schools reporting both master's and doctoral students with over 15 units (i.e., Large Reporters), and were treated as such for burden estimation purposes (section A.12).

Pilot SCs were provided with variable and file specifications to create a data upload file for reporting their institution's data at the unit-level on the GSS web instrument via a secure web-connection. Of the 80 Pilot SCs selected to participate, one declined participation at the outset (and chose to complete the regular GSS), two additional SCs did not respond to the data request, and two other SCs that happened to be working at the same institution (one from the graduate school, one from the medical school) chose to merge their data collection efforts into a single response. From the remaining 76 SCs, all were able (where applicable) to separately report master's and doctoral students, nearly all were able to upload their data by CIP codes. In a short follow-up debriefing survey conducted upon completion of the Pilot, over 90 percent of respondents indicated they were likely to continue to upload data in future GSS data collections.

Reported burden estimates from Pilot participants were compared to estimates provided in the previous year for all SCs that reported their burden in both 2015 and 2016 (n=47). These data have been used to estimate the burden for the 2017-19 cycles (see Section A.12).

A.8.4 Other Consultations

The NCSES conducted a GSS Data User Workshop in April 2016 and the Human Resource Expert Panel meeting in June 2016 to obtain feedback on making this change to the survey. There was overwhelming support for the collection of separate data on master's and doctoral level student enrollment and financial support data, which significantly increases the usefulness and value of the GSS data.

NCSES regularly consults with the Department of Education's NCES, and other federal agencies, such as NIH, professional societies, and institutions. NCSES staff members maintain frequent contact with members of the data-using community as well as with major academic data providers through attendance at professional society meetings and consultation with institutional and agency officials. GSS sessions are typically held at the Association for Institutional Researchers (AIR) Annual Forum and the CGS Annual Meeting each year to obtain respondent input.

A.9 Payment or Gifts to Respondents

Not applicable. There are no payments to GSS respondents.

A.10 Assurance of Confidentiality

No pledge of confidentiality is given to institutions providing data to the GSS because all data collected in the GSS are aggregate counts of students, postdocs, and NFRs. Data are published only at the departmental summary level.

A.11 Justification for Sensitive Questions

The survey does not contain any questions of a sensitive nature.

A.12 Estimate of Respondent Burden

Each survey cycle, when respondents reach the end of the GSS web instrument, they are asked to report how long it took them to complete the GSS. In the past three cycles (2013-2015

data collections), the average burden per organizational unit reported each cycle was 2.5 hours. However, burden varies considerably across respondents. Factors impacting burden include the number of organizational units at the institution, the degree to which requested data can be queried from centralized institutional databases, whether the GSS SC relies on the Unit Respondents (URs) in various units, for some of the requested data, and whether the SC uploads their data or manually enters it into the GSS web instrument.

In previous data collection cycles response burden was estimated on a per organizational unit basis. That is, the number of coordinating and reporting hours divided by the number of organizational (academic or research) units. The total number of hours requested was based on the expected number of organizational units in the data collection multiplied by the hours per unit burden. However, beginning with the 2017 GSS survey cycle, the use of burden per unit is no longer tenable. One key takeaway from the 2016 Pilot data collection is that when SCs directly query their institutional databases to extract GSS data, the number of unique organizational units can increase dramatically. Often, these increases are the result of the way in which data are stored in the institutional database rather than a reflection of increased organizational complexity. For example, within a single academic department, each professor's lab that employs postdocs might be stored separately in the institutional database.

In the 2016 GSS Pilot data collection, the number of organizational units increased by nearly 25 percent over what was reported in 2015. Consequently, the burden estimates for the 2017–19 data collection will focus solely on the overall number of hours required to report GSS data on a per school basis. This should present a more straightforward and readily interpretable approach to burden estimation.

The Pilot participants were asked to report their burden hours for completing the data collection, which allowed NCSES to directly measure the impact of the burden changes for the 2017 GSS survey cycle. Of the 80 SCs in the Pilot survey, 47 reported burden estimates for both the 2015 GSS and 2016 Pilot data collection. As seen in Exhibit 1, burden varies considerably with master's only and small scale (15 or fewer units) reporters reporting relatively small burden estimates compared to large-scale (over 15 units) reporters. It is worth noting that burden is substantially reduced for large-scale reporters with the data collection methods employed by the 2016 Pilot and incorporated into the 2017 GSS survey cycle. Burden levels increased very slightly for master's only and small-scale reporters. These results support the proposition that the

changes introduced into the 2017 GSS survey cycle will result in a considerable reduction in burden for many SCs.

Exhibit 1. Burden Results for 2015 GSS and 2016 GSS Pilot Data Collection

| School Type | Respondents (N) | 2015 GSS Burden (hours) | 2016 GSS Pilot Burden (hours) |
|---|------------------------|--------------------------------|--------------------------------------|
| Master’s Only | 11 | 5.7 | 5.9 |
| Master's/Doctorate: 15 or fewer units | 13 | 16.1 | 17.1 |
| Master's/Doctorate: More than 15 units (includes previous data uploaders) | 23 | 142.3 | 86.6 |
| Weighted Average | | 77.3 | 49.7 |

The expanded use of data upload is expected to mitigate the additional burden of separate reporting of master’s and doctoral degree data in 2017 GSS. In addition, respondents will be able to upload data using eligible CIP codes rather than GSS codes, which should further reduce burden.

To estimate burden for the next three data collection cycles, the GSS frame is split by institution reporting type: master’s only, small-scale reporters, and large-scale reporters. Based on the 2016 GSS, 41.0 percent of schools were master’s only, 24.8 percent were small-scale reporters, and 34.1 percent were large-scale reporters. The expected frame for the 2017 GSS includes 712 institutions comprising 826 responding schools. Applying the reporting type percentages, 339 master’s only schools, 205 small-scale reporters, and 282 large-scale reporters are expected in the 2017 GSS (see Exhibit 2). Given the historically high levels of participation, a 100 percent school response rate is used in these estimates.

Exhibit 2. Expected Composition of the 2017 GSS Frame

| Institution Type | # of Schools | Percent |
|--|---------------------|----------------|
| Master’s Only | 339 | 41.0 |
| Master's/Doctorate: 15 or fewer units | 205 | 24.8 |
| Master's/Doctorate: More than 15 units | 282 | 34.1 |
| Totals | 826 | 99.9 |

Note: Percents do not add to total due to rounding.

Burden estimates for the 2017 GSS project the burden reported by SCs in the Pilot data collection by reporting type. Additionally, the 2017 GSS data collection will include a biennial

Survey of Postdocs at the FFRDCs. Response burden for FFRDCs is estimated based on the 2015 data collection. In the 2015 data collection, FFRDCs required an average of 3.7 hours per center to complete the information request. Estimates are provided in Exhibit 3.

Exhibit 3. Burden Estimates for the 2017 GSS

| Institution Type | Respondents (# of schools) | Average Burden (hours) | Total Burden (hours) |
|--|---------------------------------------|-----------------------------------|---------------------------------|
| Master's Only | 339 | 5.9 | 2,000 |
| Master's/Doctorate: 15 or fewer units | 205 | 17.1 | 3,506 |
| Master's/Doctorate: More than 15 units | 282 | 86.6 | 24,421 |
| FFRDCs | 43 | 3.7 | 159 |
| Estimated total | 869 | | 30,086 |

Estimates for the 2018 and 2019 GSS data collections assume a one percent growth in the number of eligible schools for each respondent category. Exhibits 4 and 5 present burden estimates for the 2018 and 2019 GSS data collections, respectively. Exhibit 6 shows the total burden estimates for the 2017-19 GSS cycles and the FFRDC postdoc survey planned for 2019. In addition, the burden estimate includes 800 hours for methodological testing for continued evaluation and refinement of the GSS data collection procedures as technology evolves.

Exhibit 4. Burden Estimates for the 2018 GSS

| Institution Type | Respondents (# of schools) | Burden (hours) | Total Burden (hours) |
|--|---------------------------------------|---------------------------|---------------------------------|
| Master's Only | 343 | 5.9 | 2,024 |
| Master's/Doctorate: 15 or fewer units | 208 | 17.1 | 3,557 |
| Master's/Doctorate: More than 15 units | 285 | 86.6 | 24,681 |
| Estimated total | 836 | | 30,262 |

Exhibit 5. Burden Estimates for the 2019 GSS

| Institution Type | Respondents (# of schools) | Burden (hours) | Total Burden (hours) |
|--|---------------------------------------|---------------------------|---------------------------------|
| Master's Only | 347 | 5.9 | 2,047 |
| Master's/Doctorate: 15 or fewer units | 210 | 17.1 | 3,591 |
| Master's/Doctorate: More than 15 units | 288 | 86.6 | 24,941 |
| FFRDCs | 43 | 3.7 | 159 |
| Estimated total | 888 | | 30,738 |

Exhibit 6. Total Burden Estimates for 2017–19 GSS

| Survey Cycle | Respondents (# of schools) | Total Burden (hours) |
|--|---------------------------------------|---------------------------------|
| 2017 GSS | 869 | 30,086 |
| GSS Institutions | 826 | 29,927 |
| FFRDCs | 43 | 159 |
| 2018 GSS | 836 | 30,262 |
| 2019 GSS | 888 | 30,738 |
| GSS Institutions | 845 | 30,580 |
| FFRDCs | 43 | 159 |
| Future methodological testing (across all 3 years) | | 800 |
| Total estimated burden | 2,593 | 91,886 |
| Estimated average annual burden | 864 | 30,629 |

A.13 Cost Burden to Respondents

This survey does not require the purchase of equipment, software, or services beyond those normally used in universities as part of customary and usual business.

A.14 Cost to the Federal Government

The average cost per cycle of conducting the GSS is \$2.5M based on the total estimated value of the current contract (\$10M) to conduct four cycles, 2014–17 GSS. The total cost of the GSS to the federal government is \$2.86M per cycle. Exhibit 7 presents more detailed information on this estimate.

Exhibit 7. Annual GSS Survey Federal Government Estimated Costs

| GSS Resources and Activities | Total (\$) |
|--|-------------------|
| Data collection and processing contract | 2,500,000 |
| GSS survey manager (1.0 person year) | 150,000 |
| Other NCSSES staff (program manager, statistician, editor, etc.) | 210,000 |
| Publication Web posting, printing and mailing costs | 1,000 |
| Estimated total | 2,861,000 |

For the 2016 GSS, NIH contributed \$425,000 (15%) of the annual contract costs. It is assumed that NIH will continue that level of support. The NSF funds the remainder of the annual costs to the federal government.

A.15 Program Changes or Adjustments in Burden

Based on the 2016 pilot, which incorporated several proposed changes to data collection, NCSES expects considerable reduction in burden for the larger institutions, while burden for smaller institutions may increase very slightly.

A.16 Publication Plan and Project Schedule

The GSS project schedule (Attachment 4) for the entire project from design to final publication is similar each year. Institutions are contacted to confirm the school SCs in September, and the survey is launched in October, with a final closeout date in May of the following year. The most recent InfoBrief was published in February 2017 along with the detailed data tables, and a description of the survey methodology (<https://ncesdata.nsf.gov/datatables/gradpostdoc/2015>). There are no complex analytical issues, except imputations for nonresponse (see Section B.2.3).

A.17 Display of OMB Expiration Date

The OMB expiration date appears on the GSS Web survey login page and on GSS worksheets provided to respondents for reference purposes (worksheets are no longer used for actual data submission).

A.18 Exceptions to the Certification Statement

Not applicable. There are no exceptions.