SF-83 SUPPORTING STATEMENT

for

Survey of Earned Doctorates

2018 and 2019 survey cycles

Section A

TABLE OF CONTENTS

LIST OF ATTACHMENTS

Attachment 1:   Current 2017 SED Questionnaire & Proposed 2018 Questionnaire
Attachment 2:   Changes to the SED 2018-2019 Questionnaire Compared with SED 2017
Attachment 3:   Authorizing Legislation of Sponsoring Agencies
Attachment 4:   Current Representatives from Sponsoring Agencies
Attachment 5:   List of Persons who reviewed the SED Questionnaire
Attachment 6:   Federal Register Announcement
Attachment 7:   NSF Staff and Contractor Data Use Agreement for Individuals
Attachment 8:   Materials Sent to Institutions
    8.1:   Example of an Institution Profile
    8.2:   Job Aid for Institution Contacts
    8.3:   Institutional Contact Transmittal Form
    8.4:   Example of the Interim Results Form
    8.5:   Address Roster Form
    8.6:   Missing Information Roster
Attachment 9:   Example of Nonrespondent Follow-up Email
Attachment 10: SED Methodological Research Listing

SECTION A: JUSTIFICATION

This request is for an OMB clearance for two years covering the 2018 and 2019 cycles of the Survey of Earned Doctorates (SED). The request represents an extension of a currently approved data collection (OMB No. 3145-0019). The SED underwent cognitive testing in 2015 to identify revisions and additions that would reduce respondent burden, improve data quality, and reduce survey processing costs. The recommendations that resulted from the testing fell into two categories: (1) improvements to existing items and (2) the identification of new items. NSF received OMB approval of suggested improvements to existing items in response to a non-substantive change request submitted in April 2016. The remaining recommendations entail adding new items to the SED questionnaire, and these are described in section B.4 of this report. Both the current (SED 2017) and future (SED 2018) recommended paper questionnaire instruments are presented in Attachment 1. A list of recent methodological reports that have informed the design of the SED appears in Attachment 10.

A.1. Necessity for Information Collection

The SED began in academic year (AY) 1958 to collect data annually on the number and characteristics of individuals receiving research doctoral degrees from accredited U.S. institutions. Since then, all individuals receiving such doctorates are asked to complete the survey. A research doctorate is a doctoral degree that (1) requires the completion of an original intellectual contribution in the form of a dissertation or an equivalent culminating project (e.g., musical composition) and (2) is not primarily intended as a degree for the practice of a profession. The most common research doctorate degree is the PhD. Doctorate recipients of professional doctorate degrees such as MD, DDS, JD, DPharm, and PsyD are not included in the survey, unless they also received a research doctorate. The results of this annual survey are used to assess characteristics and trends in doctorate education and degrees. This information is vital for education and labor force planners within the federal government and in academia.

The SED is also used to identify sample members for NSF/NCSES's Survey of Doctorate Recipients (SDR). The SDR is designed to provide demographic and career history information about a sample of individuals with doctoral degrees in science, engineering and health (SEH) fields. Contact information obtained by the SED is used for locating the recently awarded doctorate recipients, who are added to the SDR sample every two years. The SDR results are used by all sectors (education, industry, and government) to understand trends in employment and salaries for doctorate holders in SEH fields. Results are also used to evaluate the effectiveness of equal opportunity efforts. Additionally, the results are important for internal planning because most NSF grants and fellowships are awarded to individuals who have earned, or are in the process of earning, doctoral degrees.

The SED is sponsored by the National Science Foundation (NSF) in collaboration with the National Institutes of Health (NIH), the U.S. Department of Education (ED), and the National Endowment for the Humanities (NEH). Sponsoring agencies typically provide funding for the SED, obtain special tabulations from the survey customized to their unique needs, and receive data tabulations/reports. The representatives of each sponsoring agency are listed in Attachment

4, and the list of persons who have been consulted and/or have reviewed the 2018-2019 SED questionnaire are listed in Attachment 5. The participating federal agencies are subject to change, pending funding availability. The National Center for Science and Engineering Statistics (NCSES) within the National Science Foundation has lead responsibility for the SED, which is currently conducted under contract by RTI International. RTI was competitively awarded a new procurement in 2016 that covers the SED operations from 2016 through 2018, with an option period through 2020.

The authority to collect information for the SED is established under the National Science Foundation Act of 1950, as amended, Public Law 507 (42 U.S.C. 1862), Section 3(a) (6), which directs the NSF "…to provide a central clearinghouse for the collection, interpretation, and analysis of data on scientific and engineering resources and to provide a source of information for policy formation by other agencies of the federal government…" Furthermore, Executive Order 10521 (March 17, 1954) states: "The Foundation shall continue to make comprehensive studies and recommendations regarding the Nation's scientific research effort and its resources for scientific activities, including facilities and scientific personnel, and its foreseeable scientific needs, with particular attention to the extent of the federal government's activities and the resulting effects upon trained scientific personnel." More recently, NCSES was established within the National Science Foundation by Section 505 of the America COMPETES Reauthorization Act of 2010 and given a broader mandate to collect data related to STEM education, the science and engineering workforce, and U.S. competitiveness in science, engineering, technology, and R&D.

The other Federal collaborating agencies also have statutory authority for the collection of information relevant to their mission. The following is a list of the applicable legislation:

1. NIH: Title I of the National Research Act of 1974 (PL 93 348);

2. ED: Education Sciences Reform Act of 2002;

3. NEH: Section 956(k) of the Arts, Humanities, and Museums Amendments of 1990, as enacted in Public Law 10 1 -512.

Attachment 3 provides the cited legislation for each collaborating agency.

A.2. Uses of Information

The SED is an accurate, timely source of information on one of our nation's most precious resources –individuals with research doctorates. The SED uniquely provides comprehensive information on the educational history and early career commitments of recent U.S.-educated doctorate recipients. The resulting information is a valuable resource for government agencies, universities, professional societies, academic researchers, policymakers, program evaluators, and individuals doing research in science policy, graduate education, economics, and human resource planning.

Each academic year the results of the SED become part of the Doctorate Records File (DRF), a complete database of more than 2 million U.S.-educated doctorate recipients from 1920 to the present (currently 2015).

The four collaborating agencies have made extensive use of the SED. Detailed tables, tabulations, and data are used by these agencies in program planning/evaluation, policy development, and dissemination. Similarly, detailed tables and data files are available to the doctorate-granting institutions that participate in the SED.

There is no public-use SED data file available; however, selected SED data items are available to the public on the internet through the WebCASPAR database (http://webcaspar.nsf.gov) and the SED Tabulation Engine (https://ncses.norc.org/NSFTabEngine/#WELCOME). Moreover, NCSES publishes detailed statistical tables each December (12 months after the close of data collection for the previous academic year), followed by additional statistical reports. In addition, organizations and individuals can request special tabulations from NCSES or the survey contractor. Lastly, researchers at U.S. institutions may gain access to the DRF by completing an NCSES Restricted-Use Data Licensing Agreement (http://www.nsf.gov/statistics/license/).

Current Uses of the SED at the Federal Level

The use of SED data and reports is widespread among collaborating federal agencies and other federal organizations. The data are used for policy development, program administration, and program evaluation. Some of the more important recent uses, organized by agency, are listed below.

a. The National Science Foundation

NSF/NCSES has conducted the SED since AY 1958. Special survey data tabulations constitute a key resource in meeting NSF policy and program needs. Examples of SED uses include the following:

- Inclusion in two of NSF's Congressionally-mandated biennial reports, *Science and Engineering Indicators* (http://www.nsf.gov/statistics/2016/nsb20161/#/), and *Women, Minorities and Persons with Disabilities in Science and Engineering (http://www.nsf.gov/statistics/wmpd/2013/start.cfm).*

- Programs within NSF, especially those dealing with women, minorities, and persons with disabilities, use data from the SED for program planning. For instance, NSF's Graduate Research Fellowship Program (GRFP) has used SED information on those who complete a PhD to evaluate the effectiveness of the GRFP and its design requirements. While most NSF programs focus on U.S. citizens, SED data on foreign citizens studying in the U.S. for their PhD are also useful for international comparisons and for quantifying the attraction of U.S. graduate education around the world.

- The universe frame for sample selection of doctoral scientists and engineers.

- Presentations of data on doctorates awarded to minorities and women to the National Science Board and the Committee on Equal Opportunities in Science and Engineering (CEOSE) for examining the participation of these groups in graduate education.

- Improvements in support of the Careers of Doctorate Holders project, an international effort led by the Organization for Economic Co-operation and Development (OECD) of the quality and comparability of international data on doctorate holders.

- Information to outside users at the national level, including the White House Office of Science and Technology Policy and the National Academy of Sciences.

- Publication of detailed statistical tables and reports on science and engineering doctorates. The first report to be released each year is available publicly in December (http://www.nsf.gov/statistics/2016/nsf16300/), 12 months after the close of data collection for the previous academic year. Additional reports follow which provide more detailed data or more analysis of the results from the SED.

- Aggregation of selected variables are available for each academic institution through the WebCASPAR database, available to the public (http://webcaspar.nsf.gov). A complementary tool, the SED Tabulation Engine, provides additional variables to the public at (https://ncses.norc.org/NSFTabEngine/#WELCOME).

b. The National Institutes of Health

For more than 30 years, NIH has used the results of the SED to meet a variety of planning, evaluating, and reporting needs:

- **Planning for the medical research workforce.** NIH relies on the results of the SED to monitor PhD production in the biomedical, behavioral, and clinical sciences. This information helps NIH determine the need for investigators in these fields and, in turn, the size and distribution of its research training programs.

- **Evaluating NIH research training programs.** Because the SED has proven to be such a reliable and comprehensive source of information on new PhDs, NIH routinely uses SED results to monitor the educational outcomes of NIH predoctoral trainees and fellows and to assess its research training programs. By comparing its internal records with the results of the SED, NIH regularly monitors PhD completion rates for students participating in NIH-sponsored training programs, their time to degree, and their plans for postdoctoral study or employment. In evaluating its research training programs, NIH also uses the SED to identify comparison groups of non-NIH-sponsored students in the same fields of study.

- **Fulfilling reporting requirements.** The SED is a critical tool in Government Performance and Results Act (GPRA) reporting on the effectiveness of NIH predoctoral training grants. In addition, since 2008, NIH has used the SED results to report on the number and type of graduate degrees awarded with NIH support, in its biennial report to Congress.

c. The Department of Education

ED has been a sponsor of the SED since inception in AY 1958. The Department's National Center for Education Statistics (NCES), Sample Surveys Division, provides funding for the

survey and makes extensive use of a range of SED data. Reports have been published on a time-series analysis of doctorates in the field of education, as well as in other fields.

NCES has also used data on the postgraduate plans of new doctorates. Tables with trend data are annually presented in the Center's publication *Digest of Education Statistics*. NCES has also published tables using the DRF that compare education doctorates to doctorates in other fields, by selected characteristics.

In addition to NCES, ED's programs, such as the Office of Student Financial Aid, the individual program offices, and the Office of Planning, Evaluation, and Policy Development also use SED data for evaluation purposes.

d. The National Endowment for the Humanities

The authorizing legislation for NEH tasks the Endowment to "develop a practical system of national information and data collection on the humanities, scholars, educational and cultural groups, and their audiences." The SED meets this mandate and gives university administrators, federal funding agencies, and private foundations an annual reading of a vital index of teaching and scholarship, the national output of humanities doctorates. NEH is currently participating in an effort led by the American Academy of Arts and Sciences to develop and regularly release via the web, a set of *Humanities Indicators*. SED data on doctorate production provides a key "indicator" of the health of the humanities workforce.

e. Other Federal Agencies and Congress

Other Federal agencies have used the SED in several ways – through requests for special tabulations and tables, data files, and licensing agreements. Congressional staffers have contacted NSF for information on several topics relevant to legislation development (such as the percent of degrees awarded to temporary visa holders and debt levels of science doctorates at graduation) and national security interests (such as nuclear engineering doctorates awarded to foreign citizens).

Academic Uses of the SED

The nation's doctorate-granting institutions not only provide SED data but also use the data. Each year since 1997, NSF/NCSES has provided to the dean of each graduate school profiles of their doctorate recipients' demographic characteristics, debt status, postgraduation plans, employment and other activities, compared with national and peer-institution data (see Attachment 8.1 for an example of an Institutional Profile).

Graduate and baccalaureate institutions use SED data for program planning, comparison with other institutions or with national figures, and in the development of affirmative action plans. The number of SED research doctorates awarded to racial/ethnic minorities by field of study is used extensively by institutions as the only reliable source on the diversity of the potential pool of applicants for academic employment positions.

Doctorate-granting institutions participating in the SED may request cumulative (going back to 1920) and preliminary (current AY only) microdata files of their own institution's doctorate recipients. During the past two years, NSF/NCSES has responded to over 170 requests for data by graduate deans, other academic administrative offices, and individual researchers, 66 for institutional datasets and 50 for preliminary institution files. The universities help administer the SED, and in return they have access to their data; it is a symbiotic data collection effort.

Researchers can apply to use selected SED microdata under the NCSES Restricted-Use Data Licensing Agreement, if publicly available data do not address their specific needs. The NCSES Licensing Agreement, executed between an institution and NCSES, requires implementing stringent security procedures to ensure the protection of confidential data against unlawful disclosure.

A.3. Consideration of Using Improved Technology

Planning for a web-based SED started in late 1999 and has been implemented, refined, and expanded since that time. The web-based survey eases the burden on students by offering an alternative completion mode, helps assure continued high response rates, improves data quality through built-in edit processes, and provides an electronic option for institutions using electronic graduation packages.

The survey population, graduating doctoral students, is quite familiar with the internet, and offering a paperless survey version which can be submitted electronically is appealing to many students. It is also very practical for respondents who are often relocating at the close of their studies and are not near the graduate offices for submission of completed paper questionnaires.

The SED's electronic procedures are focused on three components: a web survey (with a parallel paper questionnaire); nonrespondent follow-up via email; and a web interface for SED institutions. Students access the web survey directly after registering online. They are provided with a PIN and password via email should they need to complete the survey at a later time, including after graduation.

To inform web survey revisions, cognitive interviews were conducted with doctoral graduates. The recommended non-substantive changes were incorporated in the 2017 SED. The 2017 SED was also designed for survey access across different web platforms (e.g., on tablets). The recommended substantive changes for the 2018 SED are covered in this package.

The Institutional Coordinators (ICs) at all participating institutions are provided with password-protected access to a website for survey administration. The ICs Administrative Tool (ICAT) allows for two levels of interaction. Level 1 includes general information and communications, including a calendar with important survey dates and a forms library with generic and publicly available forms. Level 2, which is institution-specific and only available to each unique IC, includes current student roster data. ICs can monitor the current survey completion status of their graduates, upload revised files with student-level updates, and compare their records with the SED contractor-maintained database to track the response rate for their institution.

The web-based system has been widely accepted by both graduates and institutions. In the 2015 SED, 93 percent of questionnaires were completed via web. This demonstrates a continued increase from 2013 and 2014, when web completes were 84 percent and 91 percent, respectively.

The IC, usually located in the graduate dean's office, is the main SED interface with the doctorate recipient. They help administer the survey (either via the registration URL or the paper questionnaire), track survey completion status, and submit the completed questionnaires and final graduation lists to the SED survey contractor. The SED survey contractor works with the IC to adjust administrative tasks, as needed, to fit with each institution's procedures for processing and awarding doctoral degrees.

Significant efforts have been made to increase SED schools' web participation. As part of this outreach, schools were offered direct assistance from SED staff, who explained the web process, answered questions, and provided technical support. An additional benefit of the web option is that institutions can link the SED survey to their institution-specific exit survey, a seamless transition for students. This feature not only reduces the students' burden, but it reduces the need of both the survey contractor and the institutions to prompt students to complete the survey.

A.4. Efforts to Identify Duplication

During collaborations with other agencies and organizations, NSF/NCSES has confirmed that no other government survey gathers identical information. NSF/NCSES also learns about other survey efforts and potential duplication through contacts with professional societies and groups (such as the Council of Graduate Schools) within both the higher education and data collection communities.

SED survey content is coordinated with NSF/NCSES's SDR and National Survey of College Graduates to assure relevant uniform approaches on similar items such as race/ethnicity and specific functional limitations.

The Integrated Postsecondary Education Data System (IPEDS) also provides information on doctoral degrees. Differences between the SED and IPEDS are outlined below. While SED collects data from individuals, the IPEDS Completions survey, conducted by the National Center for Education Statistics (NCES) of the Department of Education, collects aggregate data from institutions on numbers of degrees at each level by discipline and on recipients (including race/ethnicity and sex), while the SED collects from individual doctorate recipients information on more than two dozen variables, not collected by the IPEDS survey and used by NCES.

There are four duplicative data items collected on both the SED and IPEDS: field of degree and the demographic variables of citizenship, sex, and race/ethnicity. However, important purposes are served by including these variables in both surveys:

- In the SED, field of degree, citizenship, sex, and race/ethnicity are frequently used in analyses with other key SED variables, such as the length of time spent pursuing the degree and the amount of debt accumulated during graduate education. The other key

variables cannot be collected from the IPEDS institutions. The field of degree and demographic variables are also used to identify individuals in "rare subgroups" for oversampling in the SDR.

- IPEDS' inclusion of field of degree, citizenship, sex, and race/ethnicity permits comparative analyses of trends in degree production at different degree levels. SED data cannot be substituted for IPEDS in such comparisons because of the inevitable differences in responses from institutional and demographic surveys. For example, individuals' racial/ethnic self-identification on these variables may differ from those maintained by the institutions. Also, IPEDS collects data on types of doctoral degrees that are explicitly excluded from the SED (i.e., doctorate degrees intended for the practice of a profession, such as an MD). Hence, eliminating doctoral degrees from the IPEDS data collection would result in a loss of information about these other important types of doctoral degrees.

- Including field of degree, citizenship, sex, and race/ethnicity questions on both surveys provides important validity checks for both surveys at the aggregate level.

A.5. Efforts to Minimize Burden on Small Business

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

A.6. Consequences of Less Frequent Data Collection

The SED is an important source for monitoring changes in participation in the various fields of study by demographic groups of interest (including U.S. and non-U.S. citizens on both permanent and temporary visas). The SED data pertaining to respondents' postgraduation plans provide an annual barometer of the market conditions encountered by new doctoral degree recipients and are therefore an integral component in policy implementation and program design.

Less frequent data collection would result in a more complicated survey administration by the institutions. Currently the IC staff requests SED completion from each person receiving a research doctorate at the time they complete their degree requirements. Staff (ICs) at the graduate deans' offices include the link to the web survey with their other electronic documents, post the PDF version of the survey on their website, or insert the SED questionnaire into the package of materials for doctorate recipients. Any less frequent collection of the SED would yield far lower response rates because the graduate deans' offices would be uncertain about the timing and distribution of questionnaires to prospective doctoral graduates, a process which now occurs continuously throughout the survey year. Discussions with the Council of Graduate Schools and several universities confirm that graduate schools would face extreme difficulty if the survey were operated on a non-annual basis. Stability of both the survey questionnaire and of the survey collection process is imperative for data usefulness and ease of administration.

If the SED were conducted less frequently, there would also be significant repercussions on the SDR sample selection. Locating information obtained from the SED is necessary for contacting the new PhDs who are added to the SDR sample. The coordination of timing, content, and

procedures of these two studies is, therefore, critical to the success of both the SED and SDR surveys.

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 announcement for the SED appeared on September 16, 2016 (see Attachment 6). Two public comments were received during the 60-day comment period. One requested a copy of the draft SED materials for the OMB information collection request, and the other was a note written in support of renewing the SED.

Consultations outside the Agency

NSF/NCSES often invites others to comment on the SED. Comments have come from the SED collaborating federal agencies, expert panels convened by NCSES, the Council of Graduate Schools, and other agencies and academic institutions. NCSES has also received comments from respondents, university faculty advisors, graduate deans' offices, and professional researchers by telephone, email, mail, and in-person contacts. NCSES seeks input from university representatives at venues such as professional conferences, meetings and personal site visits to institutions. These consultations have identified problems with survey administration or in the interpretation of certain data items. NCSES often discusses identified issues with ICs for their conceptual validity and applicability to all fields of study, and assesses next steps based on needs and respondent and institutional burden.

The collaborating agencies meet periodically to discuss the SED design, operation and dissemination activities, and to plan future activities. They review recent trends in the number of doctorate recipients receiving degrees in emerging fields of study – that is, fields of study not currently coded within the SED taxonomy – and in fields of study for which there are few graduates. This review is the basis for the decisions made every two years on SED taxonomy changes.

NCSES has convened multiple meetings of a Human Resources Expert Panel (HREP) to improve data collected on the education and employment of the science and engineering (S&E) labor force through review and renewal of the program's surveys, and to promote use of the data for research and policy analysis purposes. HREP accomplishes its mission by: 1) suggesting methods to publicize and promote the data; 2) providing advice on efforts to improve the timeliness and accuracy of S&E education and labor force data; 3) providing a mechanism for obtaining ongoing input from the community interested in S&E personnel data; 4) providing perspectives on the data needs of decision makers; 5) identifying issues and trends that are important for maintaining the relevance of the data; and 6) proposing ways to enhance the content of the NCSES education and workforce surveys. The panel consists of a rotating

membership of between 12 and 15 individuals who represent the sciences, academia, business/industry, government, researchers and policy makers.

NSF/NCSES conducts institution site visits focused primarily on improving poor response rates and resolving data collection problems. The site visits also allow for a discussion of the uses of SED data.

Other Consultations

NSF/NCSES has numerous other contacts with the user community, including staff of organizations such as the National Postdoctoral Association, the Association of American Medical Colleges, the Association for Institutional Research, the Council of Graduate Schools, the American Association of Universities, and the Organization for Economic Co-operation and Development. Routine information requests also provide insight into the interests of the public.

A.9. Payment or Gifts to Respondents

No incentives in the form of payment or gifts to the doctoral graduates are used in the SED.

A.10. Assurance of Confidentiality

The SED is collected in conformance with the strict confidentiality requirements found in the NSF Act of 1950, as amended. The SED is also collected in conformance with the Privacy Act of 1974, including the section of the Privacy Act requiring notification of the respondent concerning the data uses and the voluntary nature of their responses. The confidentiality pledge to SED respondents follows. The third sentence (highlighted) is a recent addition resulting from the implementation of EINSTEIN.

> This information is solicited under the authority of the National Science Foundation Act of 1950, as amended. All information you provide is protected under the NSF Act and the Privacy Act of 1974, and will be used only for research or statistical purposes by your doctoral institution, the survey sponsors, their contractors, and collaborating researchers for the purpose of analyzing data, preparing scientific reports and articles, and selecting samples for a limited number of carefully defined follow-up studies. Per the Federal Cybersecurity Enhancement Act of 2015, your data are protected from cybersecurity risks through screening of the systems that transmit your data. The last four digits of your Social Security Number are also solicited under the NSF Act of 1950, as amended; provision of it is voluntary. It will be kept confidential. It is used for quality control, to assure that we identify the correct persons, especially when data are used for statistical purposes in Federal program evaluation. Any information publicly released (such as statistical summaries) will be in a form that does not personally identify you or other respondents. Your response is voluntary and failure to provide some or all of the requested information will not in any way adversely affect you.

The time needed to complete this form varies according to individual circumstances, but the average time is estimated to be 20 minutes. If you have comments regarding this time estimate, you may write to the National Science Foundation, 4201 Wilson Blvd., Arlington, VA 22230, Attention: NSF Reports Clearance Officer. A Federal agency may not conduct or sponsor a collection of information unless it displays a currently valid OMB control number.

Specific procedures for protecting both hard copy and electronic data are used by the survey contractor, and all project staff are required to sign confidentiality agreements before they first access any SED data, and on a yearly basis thereafter (see Attachment 7). Data files with personal identifiers are provided to NIH and its contractors, a collaborating federal agency, through an NCSES Restricted-Use Data Licensing agreement with all contractors signing data use agreements. As indicated explicitly in the confidentiality statement, the graduate dean of the respondent's institution may request data for respondents from that institution only with a written agreement to use such data for statistical and program evaluation purposes only. Lastly, SED data files with personal identifiers are provided to the SDR contractor under a signed data use agreement, to locate the SDR's selected sample members. No one outside of these groups can obtain data files with direct identifiers such as phone numbers and addresses. Qualifying researchers can obtain SED microdata (but no direct identifiers) only by executing an NCSES Restricted-Use Data Licensing Agreement with NSF.

A.11. Justification for Sensitive Questions

The SED recognizes the growing sensitivity of requesting respondents' Social Security numbers to an increasing segment of the population. The SED is allowed to collect respondent Social Security numbers under the NSF Act of 1950 (42 U.S.C. 1861 et seq.), as amended, and in accordance with the Privacy Act of 1974. However, the SED only collects the last four digits of the Social Security number to be used to ascertain the correct identity of survey respondents for survey operations and evaluation purposes.

A.12. Estimate of Respondent Burden

The SED is a census of all individuals receiving a research doctorate in the United States in an academic year. In 2018 approximately 58,000 individuals are expected to receive research doctorates from U.S. institutions. Using the past response rate (90%), the number of SED respondents in 2018 is estimated to be 52,200 (58,000 doctorate recipients × 0.90 response rate). Similarly, the number of individuals expected to earn research doctorates in 2019 is estimated to be about 59,000; hence, the number of respondents in 2019 is estimated to be 53,100 (59,000 × 0.90).

Using 2017 SED data, the average response time for the 2018 SED is estimated to be 20 minutes. This estimate includes an additional minute for proposed new questions and a loss of a minute for the removal of questions. Thus, the total respondent burden is estimated to be 17,400 hours

for completing the 2018 SED questionnaire (52,200 respondents × 20 minutes) and 17,700 hours for the 2019 SED questionnaire (53,100 respondents × 20 minutes).

The estimated cost to respondents for the 2018 data collection is $504,600, based on the estimated 17,400 hours of response burden at a time-cost of $29.00 per hour. Similarly, the estimated cost to respondents for the 2019 data collection is $513,300. The $29.00 per hour time-cost estimate is derived from results from the 2015 SED which indicate that the median income for doctorate recipients was $58,000. Assuming a 40-hour work week and 50 weeks of work per year, an annual salary of $58,000 equates to $29.00 per hour.

In addition to completing the SED, NSF/NCSES also requires that institutions collect administrative data. The IC at the institution helps distribute the registration URL or paper questionnaires, track them, collect them, and submit the completed questionnaires and graduation lists to the SED survey contractor. To fulfill their role, ICs must complete the following forms (see Attachments 8.2 through 8.6 for examples of the materials):

- Transmittal Form, to be included in the package of completed surveys or with a graduation list, requests the total number of graduates for a particular graduation date, the total number of surveys enclosed, and contact information for each nonrespondent.

- Interim Result Form reports the number of graduates currently accounted for on each graduation date.

- Address Roster Form requests a physical mailing and/or e-mail address for nonrespondents that were not already provided on the Transmittal Form.

- Missing Information Roster requests the critical items for any nonrespondents and the missing critical items for respondents.

Based on focus groups conducted with ICs, it is estimated that the SED requires no more than 1% of the IC's time, which computes to 20 hours per year (40 hours per week x 50 weeks per year x .01). At an estimated median hourly wage rate of $26.61 for ICs, and with approximately 590 participating programs (from approximately 450 institutions), the total estimated time-cost to ICs of administering the SED is $313,998 per year. The $26.61 median hourly wage estimate is derived from the Bureau of Labor Statistics "May 2015 National Occupational Employment and Wage Estimates" for a combination of Office and Administrative Support Occupations ($15.96 median hourly wage, representing 60% of ICs) and Education Administrators, Postsecondary ($42.59 median hourly wage, representing 40% of ICs).

The chart below summarizes the annual burden anticipated for all the tasks involved with conducting the SED:

Description	# of Responses (mean for 2018	Respondent Burden	Annual Burden Hours	Annual Cost Burden

	& 2019)			
Doctorate recipients completing the SED Questionnaire	52,650	20 minutes	17,550	$508,950
Institutional Contacts administering SED-related tasks	590	20 hours	11,800	$313,998
Total SED Annual Burden			**29,350**	**$822,948**

A.13. Cost Burden to Respondents

Respondents need not purchase, operate, or maintain capital equipment, software, or storage facilities. There is no actual cost to the SED respondents other than the burden hour cost noted in A.12.

A.14. Cost to the Federal Government

The cost to the Federal Government for this annual data collection is approximately $3.18 million per year. This amount is based on the contract cost for the 2017-2019 SEDs.

A.15. Program Changes or Adjustments

The only expected cost adjustment in the 2018 SED is from the increase in SED universe size, resulting in a higher number of respondents.

A.16. Tabulation and Publication Plans and Project Schedule

The results of the SED will be disseminated in a number of ways. To release the data, NSF/NCSES will publish a set of approximately 70 online Summary Report Data Tables. These tables will be descriptive in nature and will provide extensive information on the education and employment plans of doctoral graduates by field of study, doctorate granting institution, and demographic characteristics such as race/ethnicity, citizenship, sex, and disability. NSF/NCSES will also publish a Report Digest, with approximately 36 figures highlighting findings from key survey themes. The Digest will be available in both print and electronic formats. The printed Digest is provided to participating SED institutions and to individuals and institutions who have requested past survey results.

The SED data will also be used in the development of key NSF reports, including the Congressionally-mandated reports *Science and Engineering Indicators* and *Women, Minorities, and Persons with Disabilities in Science and Engineering*. Both of these publications, plus additional detailed tables, will be available on the NSF/NCSES website.

Aggregated data on selected SED variables are publicly available through the WebCASPAR database (http://webcaspar.nsf.gov). SED data for race/ethnicity, sex, and citizenship for 2007

and later years are no longer available in WebCASPAR. However, they are available in the SED Tabulation Engine, a tool that includes a disclosure control mechanism for protecting the identity of respondents when reporting sex, race/ethnicity, and citizenship. Additionally, SED data will be available to licensed researchers via the NSF/NCSES's data enclave, a secure environment that provides researchers remote access to microdata while still protecting respondent confidentiality.

The Scientists and Engineers Statistical Data System (SESTAT), available via NSF/NCSES's website, includes some SED variables that pertain to SDR respondents (e.g., past educational degrees). The SESTAT system provides a rich online resource for producing tabulations. As noted above, microdata are also disseminated to collaborating agencies and licensed researchers. Their results are expected to generate reports and other publications that further disseminate the data. Finally, it is anticipated that substantive SED data analyses will be presented at relevant professional meetings, such as the annual meetings of the Association for Institutional Research, the Council of Graduate Schools, the American Educational Research Association, the American Statistical Association, the American Economic Association, the American Sociological Association, the American Association for Public Opinion Research, etc.

The 2018 SED covers doctoral graduates from 1 July 2017 to 30 June 2018. The paper SED will be mailed to the graduate schools and the web version will be uploaded upon OMB approval for continuous distribution to graduate students as they complete their doctoral requirements. Returned paper questionnaires will be edited and coded until survey close-out in December 2018 for the 2018 academic year. After the survey close-out, data variables will be constructed, edited, evaluated, and reviewed for trend consistency in January 2019. In February 2019, the file will be further evaluated and quality control checks will be made. Data will be tabulated in April 2019 and prepared for publication by November 2019. Aggregate data will be made available to the public in November 2019 via the on-line Summary Report Data Tables on the NSF/NCSES website.

Project Schedule

The 2018 SED survey schedule follows. The 2019 SED survey schedule is expected to be similar except lagging by one year.

Phase	Time
Receive OMB clearance approval	May 2017
Survey instrument available to students	June 2017
Data collection close-out	December 2018
Preparation of data file	February 2019
Production of tabulations	April 2019
Release of data by NSF	November 2019
Interagency Report released	December 2019

A.17. Display of OMB Expiration Date

The OMB Expiration Date will be displayed, as indicated.

A.18. Exception to the Certification Statement

The 2018-2019 SEDs will comply with the certification statement on form OMB 83-1.