

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

**for the**

**2021**

**National Survey of College Graduates**

## Contents

<b>B.</b>	<b>COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS.....</b>	<b>1</b>
1.	UNIVERSE AND SAMPLE DESCRIPTIONS.....	1
2.	INFORMATION COLLECTION PROCEDURES.....	2
3.	STATISTICAL ACCURACY OF THE COLLECTION.....	9
4.	TESTING PROCEDURES.....	11
5.	CONTACT INFORMATION.....	19

## **B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS**

### **1. UNIVERSE AND SAMPLE DESCRIPTIONS**

The target population for the National Survey of College Graduates (NSCG) includes individuals who meet the following criteria:

- Earned a bachelor's degree or higher,
- Are not institutionalized and reside in the United States or Puerto Rico as of the survey reference date, and
- Are younger than 76 years as of the survey reference date.

When the American Community Survey (ACS) replaced the decennial census long form beginning with the 2010 Census, the National Center for Science and Engineering Statistics (NCSES) within the National Science Foundation (NSF) began using the ACS as the sampling frame for the NSCG and implemented a rotating panel design. In this design, a new ACS-based sample of college graduates is selected and followed for four biennial cycles before the panel is rotated out of the survey. The use of the ACS as a sampling frame, including the field of degree questionnaire item included on the ACS, allows NCSES to efficiently sample subpopulations of interest (e.g., the science and engineering [S&E] workforce). The NSCG design oversamples cases in small cells of interest to analysts, including underrepresented minorities, persons with disabilities, and non-U.S. citizens. The goal of this oversampling effort is to provide adequate sample for NSF's congressionally mandated report on *Women, Minorities, and Persons with Disabilities in Science and Engineering*.

Under the rotating panel design, the 2021 NSCG production sample will include 164,000 sample cases which comprises:

- 1) Returning sample from the 2019 NSCG (originally selected from the 2013 ACS);
- 2) Returning sample from the 2019 NSCG (originally selected from the 2015 ACS);
- 3) Returning sample from the 2019 NSCG (originally selected from the 2017 ACS); and
- 4) New sample selected from the 2019 ACS.

About 74,000 new sample cases will be selected from the 2019 ACS. This sample size includes an increase of approximately 7,000 cases compared to the 2019 NSCG new sample to address lower expected response rates.

The remaining 90,000 cases will be selected from the set of returning sample members. While most of the returning sample cases are respondents from the 2019 NSCG survey cycle, about 15,000 nonrespondents from the 2019 NSCG survey cycle will be included in the 2021 NSCG sample. These 15,000 cases are individuals who responded in their initial NSCG survey cycle but did not respond during the 2019 cycle. As was the case in prior years, previous-cycle nonrespondents are being included in the 2021 NSCG sample in an effort to reduce the potential for nonresponse bias in the NSCG survey estimates. (Except for cases that are

temporarily ineligible, cases that do not respond during their first NSCG cycle are dropped from future cycles.)

The table below provides an overview of the NSCG sample sizes for the 2015 through 2021 NSCG survey cycles.

NSCG Year	Panel				Total
2015	Returning Sample			New Sample	2015 Total
	2010 NSRCG (n ≈ 13,000)	2009 ACS (n ≈ 23,000)	2011 ACS (n ≈ 57,000)	2013 ACS (n ≈ 42,000)	135,000
2017	Returning Sample			New Sample	2017 Total
	2009 ACS (n ≈ 23,000)	2011 ACS (n ≈ 28,000)	2013 ACS (n ≈ 25,000)	2015 ACS (n ≈ 48,000)	124,000
2019	Returning Sample			New Sample	2019 Total
	2011 ACS (n ≈ 27,500)	2013 ACS (n ≈ 24,500)	2015 ACS (n ≈ 29,000)	2017 ACS (n ≈ 66,000)	147,000
2021	Returning Sample			New Sample	2021 Total
	2013 ACS (n ≈ 24,000)	2015 ACS (n ≈ 28,500)	2017 ACS (n ≈ 37,500)	2019 ACS (n ≈ 74,000)	164,000

There are several advantages of this rotating panel sample design. It: 1) permits benchmarking of estimates to population totals derived from the sample using the ACS; 2) maintains the sample sizes of and allows cross-sectional estimation for small populations of scientists and engineers of great interest such as underrepresented minorities, persons with disabilities, and non-U.S. citizens; 3) provides an oversample of young graduates to allow continued detailed estimation of the recent college graduates population; and 4) permits longitudinal analysis of the retained cases.

Using the 2019 NSCG final response rate as a basis (65% unweighted, 68% weighted), NCSES estimates the final response rate for the 2021 NSCG to be 65 to 75 percent, unweighted.

## 2. INFORMATION COLLECTION PROCEDURES

### Sample Design and Selection

The 2021 NSCG target population includes all U.S. residents under age 76 with at least a bachelor’s degree prior to 1 January 2020. The new sample portion of the 2021 NSCG will be representative of this whole target population, for the most part.<sup>1</sup> The returning sample, on the

<sup>1</sup> The 2021 NSCG new sample is not completely representative of the population that first earned a degree during 2019, the ACS data collection year. For example, an ACS sample person that earned their

other hand, will be representative of a portion of the target population, namely, U.S. residents under age 76 with at least a bachelor's degree prior to 1 January 2018. The returning sample also does not include recent immigrants who have come to the United States since 2017.

As part of the 2021 NSCG sample selection, the returning sample portion of the NSCG sampling frame will be sampled separately from the new sample portion.

The 2021 NSCG returning sample will be selected with certainty from the eligible cases on the returning sampling frame, with one caveat: cases that were hard refusals in the previous cycle are not selected.

The sample selection for the 2021 NSCG new sample will use stratification variables similar to what was used in the 2019 NSCG. These stratification variables will be formed using response information from the 2019 ACS. The levels of the 2021 NSCG new sample stratification variables are as follows:

*Highest Degree Level*

- Bachelor's degree or professional degree
- Master's degree
- Doctorate degree

*Occupation/Degree Field*

*A composite variable composed of occupation and bachelor's degree field of study*

- Mathematicians
- Computer Scientists
- Biological/Medical Scientists
- Agricultural and Other Life Scientists
- Chemists
- Physicists and Other Physical Scientists
- Economists
- Social Scientists
- Psychologists
- Chemical Engineers
- Civil and Architectural Engineers
- Electrical Engineers
- Mechanical Engineers
- Other Engineers
- Health-related Occupations
- S&E-Related Non-Health Occupations
- Postsecondary Teacher, S&E Field of Degree
- Postsecondary Teacher, Non-S&E Field of Degree

---

first degree in May 2019, would be eligible for selection into the NSCG if their household was interviewed by ACS in July 2019 (i.e., after they earned their first degree). However, they would not be eligible for selection into the NSCG if their household was interviewed by ACS in March 2019 (i.e., before they earned their first degree). A weighting adjustment is implemented to correct for this undercoverage.

- Secondary Teacher, S&E Field of Degree
- Secondary Teacher, Non-S&E Field of Degree
- Non-S&E High Interest Occupation, S&E Field of Degree
- Non-S&E Low Interest Occupation, S&E Field of Degree
- Non-S&E Occupation, Non-S&E Field of Degree
- Not Working, S&E Field of Degree or S&E Previous Occupation (if previously worked)
- Not Working, Non-S&E Field of Degree and Non-S&E Previous Occupation (if previously worked) or Never Worked

Demographic Group

*A composite demographic variable composed of race, ethnicity, disability status, citizenship, and U.S.-earned degree status*

- U.S. Citizen at Birth (USCAB) or non-USCAB with high likelihood of U.S.-earned degree, Hispanic
- USCAB or non-USCAB with high likelihood of U.S.-earned degree, Black
- USCAB or non-USCAB with high likelihood of U.S.-earned degree, Asian
- USCAB or non-USCAB with high likelihood of U.S.-earned degree, AIAN or NHPI<sup>2</sup>
- USCAB or non-USCAB with high likelihood of U.S.-earned degree, disabled
- USCAB or non-USCAB with high likelihood of U.S.-earned degree, White or Other
- Non-USCAB with low likelihood of U.S.-earned degree, Hispanic
- Non-USCAB with low likelihood of U.S.-earned degree, Asian
- Non-USCAB with low likelihood of U.S.-earned degree, remaining cases

In addition, for the sampling cells where a young graduate oversample is desired, an additional sampling stratification variable will be used to identify the oversampling areas of interest. The following criteria define the cases eligible for the young graduate oversample within the 2021 NSCG.

- 2019 ACS sample cases with a bachelor's degree who are age 30 or less and are educated or employed in an S&E field,
- 2019 ACS sample cases with a master's degree who are age 34 or less and are educated or employed in an S&E field.

The multiway cross-classification of these stratification variables produces approximately 1,000 non-empty sampling cells. This design ensures that the cells needed to produce the small demographic/degree field groups for the congressionally mandated report on *Women, Minorities and Persons with Disabilities in Science and Engineering* (see 42 U.S.C. 1885d) will be maintained.

The 2021 NSCG reliability targets are aligned with the data needs for the NSF congressionally mandated reports. The sample allocation will be determined based on reliability requirements for key NSCG analytical domains. The 2021 NSCG coefficient of variation targets that drive the 2021 NSCG sample allocation and selection are included in Appendix D. Tables 1, 2, and 3 of Appendix D provide reliability requirements for estimates of the total college graduate population. Tables 4, 5, and 6 of Appendix D provide reliability requirements for estimates of young graduates, which are the target of the 2021 NSCG oversampling strata.

<sup>2</sup> AIAN = American Indian / Alaska Native, NHPI = Native Hawaiian / Pacific Islander

The 2019 ACS-based sampling frame for the 2021 NSCG new sample portion is expected to include more than 1 million cases representing the college-educated population of more than 69 million residing in the United States as of 2019.<sup>3</sup> From this sampling frame, 74,000 new sample cases will be selected based on the sample allocation reliability requirements discussed in the previous paragraph. Cases will be selected using systematic probability proportional to size sampling.

### **Weighting Procedures**

Estimates from the 2021 NSCG will be based on standard weighting procedures. As was the case with sample selection, the weighting adjustments will be done separately for the new sample cases and separately for each panel within the returning sample cases. The goal of the separate weighting processes is to produce final weights for each panel that reflect each panel's respective population. To produce the final weights, each case will start with a base weight that accounts for the probability of selection into the 2021 NSCG sample. This base weight reflects the differential sampling across strata.

#### *Weighting Adjustment for Initial Post-Stratification (New Cohort only)*

The variance of certain variables (sex, race, and age) that are used to sort the frame for systematic sampling is overestimated when using the base weights due to implicit stratification. To reduce this bias, we will perform an initial post-stratification adjustment to the base weights so that the weighted sample totals within each adjustment cell is equal to the weighted frame totals. The adjustment cells will be defined as the stratification cells broken down by sex (two levels), race (two levels: Black vs. non-Black), and age (two levels: under 40 vs. 40 and over). We will collapse adjustment cells with fewer than five sample cases.

#### *Weighting Adjustment for Survey Nonresponse*

Following the weighting methodology used in the 2019 NSCG, we will use propensity modeling to account and adjust for unit nonresponse. Propensity modeling uses logistic regression and characteristics available for all sample cases, such as prior survey responses and paradata, to predict likelihood of response. One advantage of this approach over the cell-collapsing approach used in the 1990 and 2000 decades of the NSCG is that it provides nuanced weights for each individual that better adjust for nonresponse bias. An additional advantage to using propensity modeling is the avoidance of creating complex noninterview cell collapsing rules.

Using the sampling frame variables that exist for both respondents and nonrespondents, a logistic regression model will predict response (dependent variable). The propensities output from the model will be used to categorize cases into cells of approximately equal size, with similar response propensities in each cell. The nonresponse weighting adjustment factor will be calculated as the ratio of the overall weighted population to the weighted number of respondents within each cell.

#### *Weighting Adjustment for First Post-Stratification*

---

<sup>3</sup> The 2017-ACS based sampling frame that was used for the new sample portion of the 2019 NSCG included 1,016,000 cases representing the college-educated population of approximately 69 million residing in the United States in 2017.

Following the development of nonresponse weights, post-stratification weights will be developed to ensure the population totals (according to the ACS, as the sample frame) are upheld. The cross of sex and race/ethnicity form twelve adjustment cells. The nonresponse weights will be ratio-adjusted so they sum to the ACS totals within each of the twelve adjustment cells.

#### *Weighting Adjustment for Extreme Weights*

After the completion of these weighting steps, some of the weights may be relatively large compared to other weights in the same analytical domain. (We use seven analytical domains: for example, sex, demographic group by broad occupation group, highest degree level by detailed occupation group, and young graduate oversample by demographic group by broad occupation group.) Because extreme weights can greatly increase the variance of survey estimates, weight trimming options will be implemented. When weight trimming is used, the final survey estimates may be biased. However, by trimming the extreme weights, the assumption is that the decrease in variance will offset the associated increase in bias so that the final survey estimates have a smaller mean squared error.

#### *Weighting Adjustment to Reallocate Weights*

Due to trimming, the weighted population totals within each of the analytical domains may be less than the pre-trimmed weighted population totals. This leads to biased estimates of totals. To reduce this bias, we will reallocate the trimmed weights through iterative raking to the pre-trimmed weight totals in the analytical domains.

#### *Weighting Adjustment for Second Post-Stratification*

An additional post-stratification procedure to control to ACS population totals within the twelve adjustment cells (defined by the cross of sex and race/ethnicity) will be performed.

#### *Degree Undercoverage Adjustment (New Cohort only)*

Given that individuals who earned a degree after their ACS interview date are not eligible for inclusion on the NSCG sampling frame, the 2021 NSCG has undercoverage of individuals with their first degrees earned in 2019. To ensure the 2021 NSCG provides coverage of all individuals with degrees earned during 2019, a weighting adjustment is included in the 2021 NSCG weighting procedures to account for this undercoverage. The Census Bureau conducted research on weighting adjustment methods during the 2019 NSCG cycle, and benchmarking to Integrated Postsecondary Education Data System (IPEDS) totals was the method chosen to adjust the weights for this undercoverage. Cases that first earn a bachelor's degree in the U.S. in 2019 will have their weights ratio adjusted to match IPEDS totals. The weights after the degree undercoverage adjustment serve as the final panel-level weights.

#### *Derivation of Combined Weights*

To increase the reliability of estimates of the small demographic/degree field groups used in the congressionally mandated report on *Women, Minorities and Persons with Disabilities in Science and Engineering* (see 42 U.S.C. 1885d), the new sample cases and returning sample cases will be combined. The combined weights will be formed by adjusting the new sample final weights and the returning sample final weights to account for the overlap in target population coverage. Specifically, we will use the Mecatti method to adjust the final weights from



each individual panel by dividing the final weights by the number of frames into which a sample person was eligible to be selected. The result will be a combined final weight for all 164,000 NSCG sample cases.

#### Replicate Weights and Standard Errors

Sets of replicate weights, using the successive difference method, will also be constructed to allow for separate variance estimation for the new sample, for each panel within the returning sample, and for the combined sample. The replicate weights will be used to estimate the standard errors of the 2021 NSCG estimates. This method requires that the sample selection and the estimation procedures be independently carried through (replicated) several times. The dispersion of the resulting replicated estimates then can be used to measure the variance of the full sample.

#### **Questionnaires and Survey Content**

As was the case in the 2019 NSCG, we will use three different versions of the 2021 NSCG questionnaire: 1) one for new sample members, 2) one for returning sample members who responded in 2019, and 3) one for returning sample members who did not respond in 2019. The main difference is that the questionnaires for returning sample members do not include questions where the response likely would not change from one cycle to the next. Specifically, the questionnaire for new sample members includes questions about one's degree history and demographic characteristics (e.g., race, ethnicity, and sex) that are not asked in the questionnaires for the returning sample members. If these items were not reported by the returning sample members during a prior NSCG data collection, the web and CATI instruments will attempt to collect the information this cycle. The two questionnaires for the returning sample members are similar to one another, with the exception of a slightly longer date range for the questions about recent educational experiences for the previous cycle nonrespondents.

The core items on the NSCG questionnaires focus on job characteristics, education activities, and demographics. These items are essential for sampling, respondent verification, basic labor force information, and/or robust analyses of the S&E workforce. They are asked of all respondents each time they are surveyed, as appropriate, to establish the baseline data and to update the respondents' labor force status and changes in employment and other demographic characteristics. Other questionnaire items provide the data needed to satisfy specific policy, research, or data user needs (e.g., federal support of work, job satisfaction, and immigration information).

The 2021 NSCG questionnaires will remain essentially the same as they were for the 2019 NSCG, aside from the following modifications:

1. Updating the reference dates.
2. Inserting a checkbox to gauge the willingness of respondents to receive text messages about the NSCG in future cycles. Specifically, where respondents are asked to provide their contact information (mobile phone number), a checkbox will be added to allow them to opt in to receiving text messages about the NSCG during the subsequent NSCG cycle, if Census Bureau policy permits text messaging in 2023.
3. Revising a few items and adding a few follow-up items to account for potential COVID-19 impacts. During 2020 the coronavirus pandemic affected the employment situation of

many individuals, and early 2021 may be similar. Where these effects could impact NSCG measures (e.g., employment status, part-time employment, job benefits, earnings, and conference attendance), revisions will be made to the questionnaire items to allow respondents to identify if the coronavirus pandemic was involved.

Appendix E includes a list of the questionnaire modifications being made to account for the coronavirus pandemic, along with a draft of the 2021 NSCG questionnaire for new sample members. The other two NSCG questionnaires (the questionnaire for previous cycle respondents and the questionnaire for previous cycle nonrespondents) both include a subset of the questions included on the new sample questionnaire.

### **Non-Sampling Error Evaluation**

In an effort to account for all sources of error in the 2021 NSCG survey cycle, the Census Bureau will produce a report that will include information similar in content to the 2017 NSCG Non-Sampling Error Report.<sup>4</sup> The 2021 NSCG Non-Sampling Error Report will evaluate two areas of non-sampling error – nonresponse error and error as a result of the inconsistency between the ACS and NSCG responses (measurement error). These topics will provide information about potential sources of non-sampling error for the 2021 NSCG survey cycle.

#### Nonresponse Error

Numerous metrics will be computed to motivate a discussion of nonresponse – unit response rates, compound response rates, estimates of key domains, and R-indicators.<sup>5</sup> Each of these metrics provides different insights into the issue of nonresponse and will be discussed individually and then summarized together.

Unit response rates are a simple method of quantifying what percentage of the sample population responded to the survey. We use AAPOR response rate equation three to calculate response rates, where we estimate what proportion of unknown eligibility cases are actually eligible.

The compound response rate looks at response rates over time and considers how attrition can affect the respondent population. Attrition can lead to biased estimates, particularly for surveys that do not continue to follow nonrespondents in later rounds, if respondents are different (e.g., would provide different information) from nonrespondents. The estimates become representative of the continually responding population over time, as opposed to the full target population.

To understand the potential for nonresponse bias, we will compare estimates using three sets of weights -- (1) base weights for the sample, (2) base weights for the respondents, and (3) nonresponse-adjusted weights for respondents. We will assume there is bias correction if, when comparing to the sample base weight estimates, the difference is smaller for respondents using nonresponse-adjusted weights rather than the base weights.

---

<sup>4</sup> Creamer, Selmin F. “Non-Sampling Error Report for the 2017 National Survey of College Graduates,” Census Bureau Memorandum from Tersine to Milan, March 2019. The 2019 NSCG non-sampling error report has not been completed yet.

<sup>5</sup> R-indicators are useful, in addition to response rates and domain estimates, for assessing the potential for nonresponse bias.

R-indicators and corresponding standard errors will be provided for each of the four originating sources of sample for the 2021 NSCG (namely, the 2013 ACS, 2015 ACS, 2017 ACS, and 2019 ACS). R-indicators are based on response propensities calculated using a predetermined balancing model (“balancing propensities”) to provide information on both 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. R-indicators close to 1 indicate close correspondence between the respondents and the sample and a lower risk for nonresponse bias.

#### *Error Resulting from ACS and NSCG Response Inconsistency*

Information from the ACS responses is used to determine NSCG eligibility and to develop the NSCG sampling strata. Because we use ACS responses to define the NSCG sampling strata, and we have different sampling rates in each of the strata, inconsistency with NSCG responses on the stratification variables leads to a less efficient sample design with increased variances.

### **3. STATISTICAL ACCURACY OF THE COLLECTION**

To maximize the overall survey response rate, NCSES and the Census Bureau will implement procedures such as conducting extensive locating efforts and collecting the survey data using three different modes (mail, web, and CATI). The contact information obtained for the sample members from the 2019 NSCG and the 2019 ACS will be used to locate the sample members in 2021.

#### **Respondent Locating Techniques**

The Census Bureau will refine and use a combination of locating and contact methods based on the past surveys to maximize the survey response rate. The U.S. Postal Service’s (USPS) automated National Change of Address (NCOA) database will be used to update addresses for the sample. The NCOA incorporates all change of name/address orders submitted to the USPS nationwide and is updated at least biweekly.

Prior to mailing the survey invitation letters to the sample members, the Census Bureau will engage in locating efforts to find good addresses for problem cases (“upfront locating”). The locating efforts will include using such sources as educational institutions, alumni associations, and other publicly available data found on the internet, Directory Assistance for published telephone numbers, and Accurint for address and telephone number searches. The mailings will use the “Return Service Requested” option to ensure that the postal service will provide a forwarding address for any undeliverable mail.

#### **Data Collection Methodology**

A multimode data collection protocol will be used to improve the likelihood of gaining cooperation from sample cases that are located. Using the findings from the 2010 NSCG mode effects experiment and the positive results of using the web-first approach in the 2013, 2015, 2017, and 2019 NSCG data collection efforts, the majority of the 2021 NSCG sample cases will initially

receive a web invitation letter encouraging response to the survey online. Nonrespondents will be sent a paper questionnaire mailing and will be followed up in CATI. The college graduate population is web-literate and 85% of the 2019 NSCG respondents completed the survey online.

Motivated by the findings from the incentive experiments included in the 2010 and 2013 NSCG data collection efforts and the positive results from the 2015, 2017, and 2019 NSCG incentive usage, NCSES is planning to use monetary incentives to offset potential nonresponse bias in the 2021 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 2021 NSCG data collection effort. “Highly influential” refers to the cases that had large sampling weights and a low response/locating propensity. We expect to offer \$30 debit card incentives to approximately 14,800 of the 74,000 new sample cases included in the 2021 NSCG. In addition, we will offer a \$30 prepaid debit card incentive to past incentive recipients at week 1 of the 2021 NSCG data collection effort. We expect to offer \$30 debit card incentives to approximately 14,400 of the 90,000 returning sample members. These debit cards will have a six-month usage period at which time the cards will expire and the unused funds will be returned to the Census Bureau and NCSES.

Within the 2021 NSCG data collection effort, the following steps will be taken to maximize response rates and minimize nonresponse:

- Providing “user friendly” survey materials that are simple to understand and use;
- Sending attractive, personalized material, making a reasonable request of the respondent’s time, and making it easy for the respondent to comply;
- Using priority mail for targeted mailings to improve the chances of reaching respondents and convincing them that the survey is important;
- Devoting significant time to CATI interviewer training on how to deal with problems related to nonresponse and ensuring that interviewers are appropriately supervised and monitored; and
- Using refusal-conversion strategies that specifically address the reason why a potential respondent has initially refused, and then training conversion specialists in effective counterarguments.

Please see Appendix F for the NSCG survey mailing materials from last cycle that will be updated for 2021. (The prenotice in Appendix F is from the 2017 NSCG because a prenotice was not used in 2019.) See Appendix G for the data collection pathway that provides insight on when the different survey mailing materials will be used throughout the data collection effort.

Each cycle, the NSCG asks respondents to report in which mode (web, paper, or CATI) they would prefer to complete future rounds of the survey. In past cycles, respondents who reported a preference for paper received a paper questionnaire in the first week of data collection. In the 2019 cycle, about 22% of returning respondents had indicated a preference for paper, but only half of them used the paper questionnaire to respond, while 43% responded online. Given the high rate of web response among people who reported a paper preference, we will be delaying the paper questionnaire until the

second week to see if we can maximize the number of web respondents. This delay will still provide respondents with their preferred mode but will attempt to move more people to the web, resulting in faster data processing and less postage paid in questionnaire returns. Additionally, if the web response rate among respondents with a paper preference increases significantly, it may be possible in future NSCG survey cycles to either move the paper questionnaire to week 5, where there will be a large reduction in the number of questionnaires mailed, or remove the mode preference altogether, which would simplify operations in the future.

#### **4. TESTING PROCEDURES**

##### **Survey Methodological Experiments**

Two survey methodological experiments are planned as part of the 2021 NSCG data collection effort. These experiments are designed to help NCSES and the Census Bureau strive toward the following data collection goals:

- Decrease potential for nonresponse bias in the NSCG survey estimates,
- Increase or maintain response rates,
- Lower overall data collection costs,
- Increase efficiency and reduce respondent burden in the data collection methodology.

The two methodological experiments are:

- Adaptive Survey Design Experiment,
- Prenotice Experiment.

An overview of ongoing and past experiments is provided in Supporting Statement A, Section 8; details on the two 2021 experiments can be found in Appendices H and I. Both experiments are planned for both the new sample and the returning sample data collection efforts. This section introduces the design for each experiment, describes the research questions each experiment is attempting to address, and includes information on the sample selection proposed for these studies.

##### *Adaptive Survey Design Experiment*

Beginning in 2013, the NSCG has included adaptive survey design (ASD) experiments in production data collection. Each survey round (2013, 2015, 2017, and 2019) has focused on different operational or methodological goals.

The 2019 NSCG ASD experiment intervened based on predicted survey status and response for a key variable, self-reported salary. The goal was to minimize data collection costs without sacrificing the quality of this key survey estimate. The work included the prediction of response propensity, expected costs, and both item- and unit-level imputation under different data collection scenarios to predict the estimate. These predictions allowed us to selectively apply

cost-saving data collection interventions to cases that would have a minimal impact on self-reported salary. At each intervention point, we followed these steps:

- (1) After each week of data collection, obtain posterior estimates of response propensity for all unresolved cases, using the methodology in Wagner and Hubbard (2014).<sup>6</sup>
- (2) Impute missing values for salary for all unresolved cases. The resulting estimate of the mean will be treated as the unbiased target parameters of interest.
- (3) For given response propensity cut-points identified using the propensity output in (1), assign cases to adaptive interventions (e.g., sending cases to CATI early, or replacing questionnaires with web invites) and re-compute the estimates of response propensity from step (1) under the new scenario to predict who will respond.
- (4) Re-compute estimates from step (2) after removing cases who are predicted to be non-respondents from the dataset (meaning fewer unresolved cases will be imputed). This will enable a comparison of the estimates in step (3) vs step (2). The root mean squared error (RMSE) will be estimated using the output of step (2) as the truth.
- (5) Compare the estimated cost of contacting all unresolved cases using the current data collection protocol versus only those retained in steps (3) and (4) using a new data collection protocol (e.g., introducing CATI early, withholding different contact strategies, etc.). Cost estimates will be obtained from prior rounds of the NSCG, incorporating information from the sampling frame, and costs by mode of response.
- (6) Repeat steps (3) - (5) throughout data collection, plotting the estimated costs versus the RMSE by cut-point. Times where there are sharp drops in cost with small increases in RMSE point to times where design features should be changed to reduce cost without increasing the expected RMSE.

Results from the 2019 ASD experiment are currently being analyzed; however, initial results are shown in the table below. In both the new sample and the returning sample, adaptive interventions were able to reduce the median cost-per-case by several dollars (approximately 16% in both cases). At the same time, the RMSE of self-reported salary increased by a small percentage in the treatment versus control groups (0.4% and 1.2% in the new and returning samples, respectively).

*Initial Results for the 2019 NSCG ASD Experiment*

	New Sample		Returning Sample	
	Treatment	Control	Treatment	Control
<i>Median Cost-per-Case</i>	\$31.16	\$36.94	\$28.69	\$34.14
<i>Unweighted RMSE of Salary</i>	2.1%	1.7%	3.2%	2.0%
<i>Unweighted Response Rate</i>	58.5%	59.6%	69.3%	69.6%

<sup>6</sup> Wagner, J., Hubbard, F. (2014). Producing Unbiased Estimates of Propensity Models During Data Collection, *Journal of Survey Statistics and Methodology*, 2:3, Pages 323–342, <https://doi.org/10.1093/jssam/smu009>

The response rates show that while the interventions generated cost savings by reducing effort on some cases, those reductions in effort did not cause large drops in response rates in the treatment groups versus control groups (1.1% drop in the new sample and 0.3% drop in the returning sample). Taken together, these initial results suggest that data collection interventions can be targeted to cases in an adaptive way to save data collection resources without causing large adverse effects on survey outcomes. Detailed analysis of the results, including weighted analyses and statistical significance testing, is still underway.

#### *Adaptive Survey Design in the 2021 NSCG*

The 2019 ASD experiment represented a significant innovation – intervening based on actual survey responses rather than proxies for data quality, like R-indicators. However, surveys are concerned with many estimates, not just one. Therefore, the goal of the 2021 adaptive design experiment is to identify cases for intervention that will minimize data collection costs without causing large increases in the RMSE for *several* key survey items.

While the steps carried out in 2021 will be the same as those described above in 2019, imputation for nonresponding cases (Step 2) will result in imputations for *several* key survey items. Then cases selected for cost-saving interventions will be identified based on the multivariate effect on key survey estimates. We are currently conducting a literature review and initial research into a method for making this decision. One straightforward method would be to identify the most appropriate cases for each survey item independently, and then intervene only on the cases that overlap across all the selected key survey items (intersection). Alternatively, a multivariate decision method, such as a Wald chi-square method described in Lewis (2019),<sup>7</sup> may be appropriate.

In addition to determining a method for making decisions in a multivariate fashion, Census will work with NCSES in order to determine an expanded list of key survey items, in order to develop imputation models for those additional items, similar to the way models were constructed for the 2019 experiment. Census will also update the data collection operations and intervention schedule used in 2019 to ensure that intervention time points are appropriate for the 2021 data collection operations plan.

As in 2019, we will use a Bayesian framework to leverage both historical and current information that can aid in predicting response propensity, survey outcomes and cost. Historical predictors are used to create *priors*, or initial predictive model coefficients, which are then updated with the current round of information as more data are collected to create *posteriors*. The posteriors are then used in the predictive models mentioned above. A benefit of these types of models, particularly early in data collection, is that the historical data can compensate for the data that have not been collected in the current survey round.

In preparation for the 2015 NSCG ASD experiment, and in response to a request from OMB, NCSES and the Census Bureau prepared a detailed adaptive design plan that included the time

---

<sup>7</sup> Lewis, T. (2019). Multivariate Tests for Phase Capacity. *Survey Research Methods*. 13:2, 153-165.

points where potential interventions would occur and the interventions that were available at each time point. This form of documentation will continue to serve as an overview of the adaptive design approach for the 2021 NSCG. Appendix H includes a table outlining the production data collection mailout schedule and the potential interventions available at each time point, along with the adaptive design goals and monitoring metrics.

After the 2021 NSCG experiment has ended, a review paper will be drafted summarizing the findings from all five adaptive design experiments (2013 – 2021).

### Prenotice Experiment

In the 2017 NSCG cycle, NCSES experimented with a new set of mailing materials, which were adopted for the 2019 cycle. The changes from 2017 to 2019 included eliminating some of the mailings (e.g., the prenotice letter), as well as updating the envelopes and letter content. Given all the changes to the materials in the 2017 experiment, it was not possible to measure the effect of any single change.

Lower response rates in 2019 led to the decision to reintroduce the prenotice for the 2021 cycle. However, the true impact of the prenotice, both in terms of cost and response, is not known for the NSCG. Research suggests that eliminating the prenotice can lead to slightly *slower* response but that it does not have an impact on overall response rates at the end of data collection.<sup>8,9,10</sup>

This experiment will have two conditions: the control group will receive a prenotice and the treatment group will not. Appendix I provides more detail on the rationale behind this experiment and the research questions being investigated.

### Designing the Sample Selection for the 2021 NSCG Methodological Experiments

Two methodology studies are proposed for the 2021 NSCG: the adaptive design experiment and the prenotice experiment. This section describes the sample selection methodology that will be used to create representative samples for each treatment group within the two experiments. The eligibility criteria for selection into each of the studies are as follows:

- Adaptive Survey Design Experiment: All cases are eligible for selection;
- Prenotice Experiment: Cases must not be selected in the adaptive design treatment or control group and must have a mailable address.

The sample for the adaptive design experiment will be selected independently of the sample for the prenotice experiment. Keeping the adaptive design cases separate from the other experiment will allow maximum flexibility in data collection interventions for these cases. In addition, the sample selection will occur separately for the new sample cases and the returning sample cases. This separation will allow for separate analysis for these two different sets of potential

---

<sup>8</sup> Dillman, D., Smyth, J., and Christian, L. (2014). *Internet, Phone, Mail and Mixed-Mode Surveys: The Tailored Design Method (4<sup>th</sup> edition)*. New York: Wiley & Sons.

<sup>9</sup> Murphy, P. and Roberts, A. (2015). 2014 Pre-Notice Test. American Community Survey Research and Evaluation Program.

<sup>10</sup> U.S. Census Bureau (2014). Memo: 2012 National Census Test Contact Strategy Results; Optimizing Self Response.



respondents. The main steps associated with the sample selection for the 2021 NSCG methodological studies are described below.

*Step 1: Identification and Use of Sort Variables*

Because the samples for the treatment and control groups within the methodological studies will be selected using systematic random sampling, the identification of sort variables and the use of an appropriate sort order is extremely important. Including a particular variable in the sort ensures similar distributions of the levels of that variable across the control and treatment groups.

Incentives are proposed for use in the 2021 NSCG. It has been shown in methodological studies from previous NSCG surveys that incentives are highly influential on response. An incentive indicator variable will be used as the first sort variable for both methodological studies. The 2021 NSCG sample design variables are also highly predictive of response and will also be used as sort variables in all studies. The specific sort variables used for each experiment are:

- Incentive indicator,
- 2021 NSCG sampling cell and sort variables.

*Step 2: Select the Samples*

For the adaptive design experiment, roughly the same sample sizes will be used in 2021 as were used in the 2015, 2017, and 2019 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, effect on key estimates, and cost.

For the new sample adaptive design experiment, a systematic random sample of approximately 8,000 cases will be selected for the treatment group and 8,000 cases will be selected for the control group. For the new sample prenotice experiment, the sample will be subset to the eligible population (see above for eligibility criteria) and a systematic random sample of approximately 10,000 cases will be selected for the treatment group (i.e., no prenotice). All eligible new sample cases not selected into the adaptive design treatment, adaptive design control, or prenotice treatment group will be assigned to the prenotice control group (approximately 48,000 cases).

For the returning sample adaptive design experiment, a systematic random sample of approximately 10,000 cases will be selected for the treatment group and 10,000 cases will be selected for the control group. For the returning sample prenotice experiment, the sample will be subset to the eligible population (see above for eligibility criteria) and a systematic random sample of approximately 10,000 cases will be selected into the treatment group. All eligible returning sample cases not selected into the adaptive design treatment, adaptive design control, or prenotice treatment group will be assigned to the prenotice control group (approximately 60,000 cases).

*Minimum Detectable Differences for the 2021 NSCG Methodological Experiment*

Appendix J provides information on the minimum detectable differences achieved by the sample sizes associated with the 2021 NSCG methodological experiment.

### Analysis of Methodological Experiments

In addition to the analysis discussed in the sections describing the experiment, we will calculate several metrics to evaluate the effects of the methodological interventions and will compare the metrics between the control group and treatment groups. We will evaluate:

- response rates (overall and by subgroup);
- R-indicators (overall R-indicators, variable-level partial R-indicators, and category-level partial R-indicators);
- mean squared error (MSE) effect on key estimates; and
- cost per sample case/cost per complete interview (overall and by subgroup).

The subgroups that will be broken out are the ones that primarily drive differences in response rates and include age group, race/ethnicity, highest degree, and young graduate groups.

### **Non-production Bridge Panel**

The nature of educational attainment, the labor market, and workforce-related training is changing. In response to this changing environment, and further motivated by a recommendation from the National Academies of Science, Engineering, and Medicine’s Committee on National Statistics,<sup>11</sup> NCSES would like to explore modifications to the NSCG survey content. Standard cognitive testing of question wording modifications provides a qualitative measure of quality and usability but does not assess or quantify the potential impact on survey estimates. Given the importance of maintaining the NSCG’s trend data, NCSES plans to include a small, representative, non-production sample (referred to as a bridge panel) to quantify the potential impact of question wording modifications on key survey estimates.

The bridge panel would allow NCSES to compare current NSCG survey estimates (using responses from the NSCG production sample) with estimates resulting from the modified questions (using responses from the bridge panel). Thus, the bridge panel would serve as a bridge to our current questions and could aid in the transition of our survey to possible question wording modifications. In future cycles, the bridge panel would provide NCSES the opportunity to assess and quantify the impact to survey estimates of potential methodological changes.

### Sample Design and Selection

The 2021 NSCG non-production bridge panel will include 5,000 sample cases selected from the 2019 ACS. As noted earlier, the 2021 NSCG target population includes all U.S. residents under age 76 with at least a bachelor’s degree prior to 1 January 2020. Like the new sample portion of the 2021 NSCG production sample, the bridge panel will be representative of this whole target population, for the most part.<sup>12</sup>

---

<sup>11</sup> At NCSES’s request, CNSTAT convened an expert panel to review, assess, and provide guidance on NCSES’s efforts to measure the S&E workforce population in the United States. Recommendation 5.2 of the [panel’s consensus study report](#) noted that NCSES “*should continue to monitor, and formally evaluate as needed, the content of its survey questionnaires to ensure that the concepts and terminology are up to date and familiar to respondents. Changes should be implemented with careful consideration of their impact on trend data.*”

As part of the 2021 NSCG sample selection effort, the 5,000-case non-production bridge panel will be sampled separately from the 164,000-case production sample. The sample selection for the bridge panel will use stratification variables similar to those used for the new sample cases, as discussed in Section B.2., but with aggregated categories. These stratification variables will be formed using response information from the 2019 ACS. The levels of the 2021 NSCG bridge panel stratification variables are as follows:

Highest Degree Level

- Bachelor's degree or professional degree
- Master's degree
- Doctorate degree

Occupation

- Science and Engineering (S&E) Occupations
- Health-related Occupations and S&E-Related Non-Health Occupations
- Postsecondary and Secondary Teachers
- Non-S&E Occupations
- Not Working or Never Worked

Demographic Group

*A composite demographic variable composed of race, ethnicity, disability status, citizenship, and U.S.-earned degree status*

- U.S. Citizen at Birth (USCAB) or non-USCAB with high likelihood of U.S.-earned degree, Hispanic
- USCAB or non-USCAB with high likelihood of U.S.-earned degree, Black
- USCAB or non-USCAB with high likelihood of U.S.-earned degree, Asian
- USCAB or non-USCAB with high likelihood of U.S.-earned degree, AIAN or NHPI<sup>13</sup>
- USCAB or non-USCAB with high likelihood of U.S.-earned degree, disabled
- USCAB or non-USCAB with high likelihood of U.S.-earned degree, White or Other
- Non-USCAB with low likelihood of U.S.-earned degree, all remaining cases

The multiway cross-classification of these stratification variables produces approximately 105 non-empty sampling cells. The 5,000 bridge panel sample cases will be allocated across the 105 sampling cells in a manner that aligns with how the 74,000 new sample cases were allocated across these same 105 sampling cells (recall, these 105 sampling cells were determined by aggregating the 1,000 new sample sampling cells). After determining the sample allocation per sampling cell, cases will be selected using systematic probability proportional to size sampling. The use of aggregated versions of the new sample stratification variables will enable comparison of key estimates between the bridge panel and the new sample.

---

<sup>12</sup> The 2021 NSCG non-production bridge panel will not be completely representative of the population that first earned a degree during 2019, the ACS data collection year. For example, an ACS sample person that earned their first degree in May 2019, would be eligible for selection into the bridge panel if their household was interviewed by ACS in July 2019 (i.e., after they earned their first degree). However, they would not be eligible for selection into the bridge panel if their household was interviewed by ACS in March 2019 (i.e., before they earned their first degree). A weighting adjustment will be implemented to correct for this undercoverage.

<sup>13</sup> AIAN = American Indian / Alaska Native, NHPI = Native Hawaiian / Pacific Islander

### Weighting Procedures, Replicate Weights, and Standard Errors

Estimates from the 2021 NSCG bridge panel will be based on standard weighting procedures. As was the case with sample selection, the weighting adjustments will be done separately for the bridge panel and production sample cases. The goal of the separate weighting processes is to produce final weights for the bridge panel. To produce the final weights, the bridge panel cases will follow the weighting methodology outlined in Section B.2. In addition, sets of replicate weights, using the successive difference method, will be constructed to allow for separate variance estimation for the bridge panel.

### Questionnaire and Survey Content

The 2021 NSCG bridge panel questionnaire will include content similar to the new sample questionnaire included in Appendix E with three modifications:

- 1) The educational history section will first ask respondents to provide a roster of each degree earned, and then use this roster to solicit their degree history.
- 2) For the questionnaire items that were modified for 2021 to include coronavirus pandemic response options (i.e., employment status, part-time employment, job benefits, earnings, and conference attendance), the question wording from 2019 without the coronavirus pandemic response options will be used.
- 3) The questionnaire item measuring gender will be modified to offer response options beyond the binary responses of male and female.

### Respondent Locating Techniques and Data Collection Methodology

As described in Section B.3, the Census Bureau will use a combination of locating and contact methods based on past NSCG surveys to maximize the survey response rate among the bridge panel cases. In terms of data collection methodology, the bridge panel will use a single-mode, web-based **data collection protocol**. The bridge panel cases will follow the ‘web first pathway’ outlined in Appendix G with two modifications given the single web mode: (1) Bridge panel cases will not receive a paper questionnaire as part of the contacts at weeks 8 and 20; (2) No outgoing telephone calls will be made to the bridge panel cases during weeks 12-22. The bridge panel will receive survey mailing materials similar to those planned for the production sample cases (see Appendix F).

Similar to the incentive approach planned for the production sample, we plan to offer a \$30 prepaid debit card to a subset of highly influential bridge panel cases at week 1 of the 2021 NSCG data collection effort. “Highly influential” refers to cases that have large sampling weights and a low response/locating propensity. We expect to offer \$30 debit card incentives to approximately 1,000 of the 5,000 bridge panel cases included in the 2021 NSCG. These debit cards will have a six-month usage period at which time the cards will expire and the unused funds will be returned to the Census Bureau and NCSES.

### Evaluation of Bridge Panel Question Modifications

As noted earlier, the bridge panel is designed to provide NCSES with an opportunity to compare current NSCG survey estimates (using responses from the NSCG production sample) with estimates resulting from the modified questions (using responses from the bridge panel sample).

This comparison is designed to aid in the NSCG transition to possible question wording modifications.

To determine if the modified questionnaire items should be included in subsequent NSCG survey cycles, NCSES will conduct an evaluation of the bridge panel question modifications at the completion of the 2021 NSCG data collection effort. The evaluation will include three components:

- 1) Comparison of survey estimates between the production sample and bridge panel sample for the modified questionnaire items. This comparison will be at both the total college-educated population level and the level of the bridge panel stratification variables.
- 2) Comparison of overall nonresponse rates and item-level nonresponse rates between the production sample and bridge panel sample.
- 3) Comparison of web instrument paradata (e.g., breakoff rates, changed answer rates, etc.) for both the production sample and bridge panel sample to assess the user experience associated with the modified questionnaire items.

This evaluation approach is designed to provide insight on both the estimation impact of these question modifications as well as any nonsampling error issues that would be introduced through the inclusion of these questions.

## **5. CONTACT INFORMATION**

At NCSES, the contacts for statistical aspects of data collection are Samson Adeshiyan, NCSES Chief Statistician (703)292-7769, and Lynn Milan, NSCG Project Officer (703)292-2275.

The U.S. Census Bureau will be responsible for collecting data for the 2021 NSCG, via an Interagency Agreement. Chief consultant on data collection and methodological issues at the Census Bureau is Stephen Simoncini, NSCG Survey Director, (301)763-4816. The Demographic Statistical Methods Division (DSMD) will manage all NSCG sample selection operations at the Census Bureau. Chief consultant on statistical issues at the Census Bureau is Aaron Gilary, DSMD Lead Scientist, (301)763-9660.