

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

### **1. RESPONDENT UNIVERSE AND SAMPLING METHODS**

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) at the National Science Foundation (NSF) identified the ACS as the potential sampling frame for the National Survey of College Graduates (NSCG) for use in the 2010 survey cycle and beyond. After reviewing numerous sample design options proposed by NCSES, the Committee on National Statistics (CNSTAT) recommended a rotating panel design for the 2010 decade of the NSCG (National Research Council, 2008). The use of the ACS as a sampling frame allows NCSES to more efficiently target the science and engineering (S&E) workforce population. Furthermore, the rotating panel design planned for the 2010 decade allows the NSCG to address certain deficiencies of the previous design including the undercoverage of key groups of interest such as foreign-degreed immigrants with S&E degrees.

The NSCG design for the 2010 decade sample selects more cases in small cells of particular interest to analysts, including underrepresented minorities, women, persons with disabilities and non-U.S. citizens. This results in the surveys of the 2010 decade continuing to oversample underrepresented minorities, women, and persons with disabilities as in the 2000 decade design. 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*.

To continue the transition into the rotating panel design that began with the 2010 NSCG, the 2015 NSCG will include 135,000 sample cases which includes: 1) Returning sample from the 2010 NSCG (originally selected from the 2009 ACS); 2) Returning sample from the 2013 NSCG (originally selected from the 2011 ACS and 2010 NSRCG); and 3) New sample selected from the 2013 ACS.

About 42,000 new sample cases will be selected from the 2013 ACS. The remaining 93,000 cases will be selected from the set of returning sample members. While most of the returning sample cases are respondents from the 2013 NSCG survey cycle, about 7,000 nonrespondents from the 2013 NSCG survey cycle will be included in the 2015 NSCG sample. These 7,000 cases are individuals that responded in the 2010 NSCG survey cycle, but did not respond during the 2013 NSCG survey cycle. These 2013 NSCG nonrespondents are being included in the 2015 NSCG sample in an effort to reduce the potential for nonresponse bias in our NSCG survey estimates.

The 2015 NSCG survey target population includes all U.S. residents under age 76 with at least a bachelor's degree as of January 1, 2014. The new sample portion of the 2015 NSCG will provide complete coverage of this target population. The returning sample, on the other hand, will provide only partial coverage of the 2015 NSCG target population. Specifically, the returning sample will cover the population of U.S. residents under age 76 with at least a bachelor's degree as of January 1, 2012.

There are several advantages of this rotating panel sample design. It: 1) permits longitudinal analysis of the retained cases from the ACS-based sample; 2) permits benchmarking of estimates to population totals derived from the sample using ACS; 3) maintains the sample sizes of small populations of scientists and engineers of great interest such as underrepresented minorities, persons with disabilities and non U.S. citizens; 4) provides an oversample of young graduates to allow continued detailed estimation of the recent college graduates population; and 5) allows direct comparison of the estimation capabilities for recent graduates estimates derived from cases originally sampled in the 2010 NSRCG to recent graduates estimates derived from cases originally sampled in the 2011 or 2013 ACS.

There are two different versions of the NSCG questionnaire – a version for new sample cases and a version for returning sample cases. The main difference is that the questionnaire for returning sample cases does not include questions where the response likely will not change from one cycle to the next. Specifically, the questionnaire for new sample cases includes a degree history grid and certain demographic questions (e.g., race, ethnicity, and gender) that are not asked in the questionnaire for the returning sample. If these items were not collected from the returning sample cases during the initial NSCG survey round, the web and CATI instruments will attempt to collect this information this cycle.

The target response rate for the new sample is approximately 70 percent. The target response rate for the returning sample is approximately 80 percent. NCSES targeted these response rates based on 2013 final response rates.

## **2. SURVEY METHODOLOGY**

### **Sample Design and Selection**

As part of the 2015 NSCG sample selection, the returning sample portion of the NSCG sampling frame will be sampled separately from the new sample portion.

The majority of the 2015 NSCG returning sample will be selected with certainty from the returning sampling frame. This certainty sampling approach will apply to cases that originated from the 2011 ACS or the 2010 NSRCG. The only portion of the returning sampling frame that will have a sample reduction is the cases that originated in the 2009 ACS. These cases will receive a 50% sample maintenance reduction as part of the planned implementation of the NSCG rotating panel design. In the first two cycles of the NSCG rotating panel design (i.e., the 2010 and 2013 NSCG), additional sample was selected from the ACS to ensure enough cases were in sample to allow for reliable estimation. Since the 2015 NSCG will include new sample selected from the 2013 NSCG, a portion of the returning ACS-based sample is no longer needed. As a result, only 50% of the 2015 NSCG sampling frame cases that originated in the 2009 ACS will be selected for the 2015 NSCG sample. This 50% maintenance cut will occur across all sampling frame cases that originated in the 2009 ACS regardless of their 2013 NSCG final interview disposition.

The sample selection for the 2015 NSCG new sample will use stratification variables similar to what was used in the 2013 NSCG. These stratification variables will be formed using response

information from the 2013 ACS. The levels of the 2015 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 which is composed of occupation and bachelor's degree field of study*

- Mathematician
- Computer Scientists
- Life Scientists
- Physical Scientists
- Social Scientists
- Psychologists
- Engineers
- Health-related Occupations
- S&E-Related Non-Health Occupations
- Post Secondary Teacher, S&E Field of Degree
- Post Secondary Teacher, Non-S&E Field of Degree
- 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
- Not Working, Non-S&E Field of Degree

Demographic Group

*A composite demographic variable which is composed of race, ethnicity, disability status, citizenship, and foreign earned degree status*

- U.S. Citizen at Birth (USCAB), Hispanic
- USCAB, Non-Hispanic, Black
- USCAB, Non-Hispanic, Asian
- USCAB, Non-Hispanic, American Indian/Alaska Native or Native Hawaiian/Pacific Islander
- USCAB, Non-Hispanic, White or Other Race, Disabled
- USCAB, Non-Hispanic, White or Other Race, Non-Disabled
- Non-USCAB, Hispanic
- Non-USCAB, Non-Hispanic, Asian
- Non-USCAB, Non-Hispanic, Other Race

In addition, for the sampling cells where a young graduates oversample is desired<sup>17</sup>, 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 graduates oversample within the 2015 NSCG.

- 2013 ACS sample cases with a bachelor's degree who are ages 30 or less and are educated or employed in an S&E field
- 2013 ACS sample cases with a master's degree who are ages 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 2015 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 provided by NCSES. The 2015 NSCG coefficient of variation targets that drive the 2015 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 2015 NSCG oversampling strata.

In total, the ACS-based sampling frame for the 2015 NSCG new sample portion includes over 970,000 cases representing the college-educated population of 63 million residing in the U.S. as of 2013. From this sampling frame, 42,000 new sample cases will be selected based on the sample allocation reliability requirements discussed in the previous paragraph.

### **Weighting Procedures**

Estimates from the 2015 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 the returning sample cases. The goal of the separate weighting processes is to produce final weights for each sample portion (i.e., new sample and returning sample) that reflects each portion's respective population. To produce the final weights, each case will start with a base weight defined as the probability of selection into the 2015 NSCG sample. This base weight reflects the differential sampling across strata. Base weights will then be adjusted to account for unit nonresponse.

#### *Weighting Adjustment for Survey Nonresponse*

Following the weighting methodology used in the 2010 and 2013 NSCG, we will use propensity modeling to account and adjust for unit nonresponse. Propensity modeling uses logistic regression to determine if characteristics available for all sample cases, such as prior survey responses and paradata, can be used to predict response. One advantage to this approach over

---

<sup>17</sup> Since the young graduates oversample planned for the NSCG serves to offset the discontinuation of the NSRCG, the oversample will focus only on bachelor's and master's degree recipients as had the NSRCG.

the cell collapsing approach used in the 1990 and 2000 decades of the NSCG is the potential to more accurately reallocate weight from nonrespondents to respondents that are similar to them, in an attempt to reduce nonresponse bias. An additional advantage to using propensity modeling is the avoidance of creating complex noninterview cell collapsing rules.

We will create a model to predict response using the sampling frame variables that exist for both respondents and nonrespondents. A logistic regression model will use response as the 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 noninterview weighting adjustment factors will be calculated within each of the cells.

The noninterview weighting adjustment factor is used to account for the weight of the 2015 NSCG nonrespondents when forming survey estimates. The weight of the nonrespondents will be redistributed to the respondents and ineligibles within the 2015 NSCG sample. The weight of nonrespondent eligible cases will only go to the respondent cases. The weight of nonrespondent – eligibility unknown cases will go to both the respondent and ineligible cases. After the noninterview adjustment, weights will be controlled to ACS population totals through a post-stratification procedure that ensures the population totals are upheld.

#### *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. Since extreme weights can greatly increase the variance of survey estimates, NCSES will examine weight trimming options. 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 square error. Depending on the weighting truncation adjustment used to address extreme weights, it is possible the weighted totals for the key marginals will no longer equal the population totals used in the iterative raking procedure. To correct this possible inequality, the next step in the 2015 NSCG weighting processing will be an iterative raking procedure to control to pre-trimmed totals within key domains. Finally, an additional execution of the post-stratification procedure to control to ACS population totals will be performed, and the resulting weight will be the final weight for each sample portion (i.e., new sample and returning sample).

#### *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), NCSES will combine the new sample and returning sample together and will form combined weights to use in estimation for the combined set of cases. 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. The result will be a combined final weight for all 135,000 NSCG sample cases.

#### *Replicate Weights*

Sets of replicate weights will also be constructed to allow for separate variance estimation for the returning sample and the new sample. The replicate weight for the combined estimates will be

constructed from these sets of replicate weights. The entire weighting process applied to the full sample will be applied separately to each of the replicates in producing the replicate weights.

### Standard Errors

The replication weights will be used to estimate the standard errors of the 2015 NSCG estimates. The variance of a survey estimate based on any probability sample may be estimated by the method of replication. This method requires that the sample selection, the collection of data, 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.

### **Nonsampling Error Evaluation**

In an effort to account for all sources of error in the 2015 NSCG survey cycle, the Census Bureau will produce a report that will include information similar to the contents of the 2013 NSCG Nonsampling Error Report<sup>18</sup>. The 2015 NSCG Nonsampling Error Report will evaluate three areas of nonsampling error – nonresponse error, error as a result of the inconsistency between the ACS and NSCG responses, and measurement error due to the NSCG questionnaire design. These topics will provide information about potential sources of nonsampling error for the 2015 NSCG survey cycle.

### Nonresponse Error

Numerous metrics will be computed in order to motivate a discussion of nonresponse – unit response rates, compound response rates, estimates of key domains, item nonresponse rates, and R-indicators. 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. For example, in the 2013 NSCG new sample portion, the overall weighted response rate was 70.4%; however, age groups had weighted response rates ranging from 64% for younger age groups, versus 80% for the oldest age groups. Some variation in response is expected due to random variation; however, large variations in response behavior can be a cause for concern with the potential to introduce nonresponse bias. Assuming we are measuring different subgroups of the target population separately because we are interested in the different response data they provide, then having differential response rates across subgroups may mean we are missing information in the less responsive subgroups. This is the driving force behind nonresponse bias – a relationship between the explanatory variables and the outcome variables. If the explanatory variables are also related to the likelihood to respond, resulting estimates may be biased.

The compound response rate looks at response rates over time, and considers how attrition can affect the respondent population. Attrition is important when considering the effect of nonresponse in longitudinal surveys like the NSCG. As an example, for the returning sample cases that originated in the 2009 ACS, a response rate of 98% in the ACS followed by two rounds of NSCG with weighted response rates around 80% results in a compound response rate

---

<sup>18</sup> Zotti, Allison, “Nonsampling Error Report for the 2013 National Survey of College Graduates,” Census Bureau Memorandum from Reist to Finamore and Rivers, June 2014 draft.

of just 63%. This means that only 63% of the cases originally eligible and sampled for the NSCG through the ACS have responded in the current round, with most of that attrition arising in the NSCG itself. Attrition can lead to biased estimates, particularly for surveys that do not continue to follow nonrespondents in later rounds. This is because weighting adjustments and estimates are based on a dwindling portion of the population. This can lead to weight inflation and increased variances, which may make significant differences more difficult to detect in the population. Further, if respondents are different (e.g., would provide different information) from nonrespondents, excluding the nonrespondents effectively excludes a portion of valuable information from the response and the resulting estimates. The estimates become representative of the continually responding population over time, as opposed to the full target population.

Examining the estimates of key domains provides insight on whether the potential for bias due to nonresponse error is adversely impacting the survey estimates. In order to account for nonresponse, and ensure the respondent population represents the target population in size, nonresponse weighting adjustments are made to the respondent population. Following the nonresponse adjustment, post-stratification is employed to ensure the respondent population represents not just the size of the target population, but also the proportion of members in various domains of the population. In order to estimate the effect of these adjustment steps, estimates of various domains within the NSCG target population will be calculated from the frame, from respondents, after the nonresponse adjustment, and after final adjustments. This examination will provide insight on whether the NSCG weighting adjustments are appropriately meeting the NSCG survey estimation goals.

In order to examine item nonresponse, response rates for all questionnaire items will be produced. In addition, to examine the impact of data collection mode on item nonresponse, item response rates by response mode also will be produced. Like the unit response rates, the item response rates can be used as an indicator for potential bias in our survey estimates.

R-indicators and corresponding standard errors will be provided for each of the four originating surveys that make up the 2015 NSCG. R-indicators are useful, in addition to response rate and domain estimates, for assessing the potential for nonresponse bias. R-indicators are based on response propensities calculated using a predetermined balancing model (“balancing propensities”) to provide information on 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.

#### *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. Inconsistency between ACS responses and NSCG responses has the potential to inflate non-sampling error in multiple ways and will be investigated as part of the 2015 NSCG nonsampling error evaluation. Since 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. For example, we sample non science and engineering (non-S&E) occupations at much lower rates than S&E occupations which leads to large weights for non-S&E cases and small weights for S&E cases. If a case is identified as non-S&E on the ACS, but

lists an S&E occupation on the NSCG, then this case with a large weight is introduced into the S&E domain thus increasing the variance of estimates for the S&E domain. The mixing of cases from different sampling strata due to ACS/NSCG response inconsistency thus leads to an inefficient design and contributes to larger variances.

Another opportunity for ACS/NSCG inconsistency leading to non-sampling error is with off-year estimation<sup>19</sup>. To the extent ACS responses are inconsistent with NSCG responses, using the ACS data to produce estimates for the college-educated population will lead to biased estimates. Therefore, consistency between the ACS and NSCG responses is very important if we want to consider the possibility of producing off-year estimates with smaller bias.

### *Measurement Error*

Measurement error due to a questionnaire's design can occur in three different areas: the word choices for each question, how the question is structured, and the order in which questions are presented. The 2015 NSCG nonsampling error report will discuss the impact of measurement error in these three areas of questionnaire design. The word choice or vocabulary used for each question can be a source of measurement error because respondents can interpret words differently. How a question is structured can also lead to measurement error. The length of a question, type of question (e.g. open-ended vs. multiple-choice answer), and what response options are available can all have an effect on a respondent's understanding of the question. This section of the 2015 NSCG nonsampling error report will discuss the findings from the cognitive interviews that occurred prior to the 2015 NSCG data collection effort and the potential impact of these findings on the 2015 NSCG survey estimates.

## **3. METHODS TO MAXIMIZE RESPONSE**

In order 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 from the 2013 NSCG and the 2013 ACS for the sample members and for the people who are likely to know the whereabouts of the sample members will be used to locate the sample members in 2015.

### **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 Census Bureau will utilize all available locating tools and resources to make the first contact with the sample person. The Census Bureau will use the U.S. Postal Service (USPS)'s automated National Change of Address (NCOA) database 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.

---

<sup>19</sup> Off-year estimation would provide estimates for the college educated population, using only ACS data, in the years where the NSCG is not in the field. For example, as the NSCG is conducted in 2013, 2015, and 2017, off-year estimation would produce estimates for the college-educated population in 2014 and 2016.



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. The mailings will utilize the “Return Service Requested” option to ensure that the postal service will provide a forwarding address for any undeliverable mail. For the majority of the cases, the initial mailing to the NSCG sample members will be a letter introducing the survey and inviting them to complete the survey by the web data collection mode. For the cases that stated a preferred mode for use in future survey rounds (e.g., mailed questionnaire or telephone), NCSES will honor that request by contacting the sample member using the preferred mode to introduce the survey and request their participation.

The locating efforts will include using such sources as educational institutions and alumni associations, Directory Assistance for published telephone numbers, Phone Disc for unpublished numbers, FastData for address searches, and local administrative record searches such as researching motor vehicle department records. Private data vendors also maintain up to 36-month historical records of previous address changes. The Census Bureau will utilize these data vendors to ensure that the contact information is up-to-date.

### **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 NSCG data collection effort, the majority of the 2015 NSCG sample cases will initially receive a web invitation letter encouraging response to the survey online. Nonrespondents will be given a paper questionnaire mailing and will be followed in CATI. The college graduate population is mostly web-literate and, as shown in the 2010 mode effects experiment, the initial offering of a web response option appeals to NSCG respondents (including the NSRCG panel sample members.) In addition, an adaptive design experiment will be incorporated into the 2015 NSCG data collection efforts that allows for contact tailoring to encourage response in subgroups with lower response propensities.

Motivated by the findings from the incentive experiments included in the 2010 and 2013 NSCG data collection efforts, NCSES is planning to use monetary incentives to offset potential nonresponse bias in the 2015 NSCG. We plan to offer a \$30 prepaid debit card incentive to a subset of highly influential new sample cases at week 1 of the 2015 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 8,000 of the 42,000 new sample cases included in the 2015 NSCG. In addition, we will offer a \$30 prepaid debit card incentive to past incentive recipients at week 1 of the 2015 NSCG data collection effort. We expect to offer \$30 debit card incentives to approximately 14,500 of the 93,000 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 Census and NCSES.

In addition to these procedures, the following steps will be taken to maximize response rates and minimize nonresponse:

- Developing “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 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 E for survey mailing materials.

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

##### **Questionnaire Construction**

Because data from the SESTAT surveys are combined into a unified data system, the two SESTAT surveys must be closely coordinated to provide comparable data from each survey. As a result, there are similarities in the questionnaire items between the two surveys.

The SESTAT survey questionnaire items are divided into two types of questions: core and module. Core questions are defined as those considered to be the base for the SESTAT surveys. These items are essential for sampling, respondent verification, basic labor force information, and/or robust analyses of the science and engineering workforce in the SESTAT integrated data system. 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. Module items are defined as special topics that are asked less frequently on a rotational basis of the entire target population or some subset thereof. Module items tend to provide the data needed to satisfy specific policy, research, or data user needs.

As part of the 2015 NSCG planning effort, NCSES conducted developmental work on new questionnaire items to capture information on alternative credentials including industry-recognized certifications, occupational licenses, and educational certificates. These are concepts that have been added to other federal surveys in recent years and were deemed of high analytical interest by data users and policy makers. After evaluating these questions by conducting stakeholder outreach and two rounds of cognitive interviews, NCSES decided to add a new NSCG questionnaire section to collect information on certifications and licenses.

In addition to considering the possible inclusion of questions on the attainment of certifications, licenses, and educational certificates, NCSES asked the Census Bureau’s Center for Survey Measurement to conduct an expert review and cognitive interviews for the full set of NSCG

questionnaire items. The expert review and additional cognitive interviews resulted in minor question wording revisions to numerous items throughout the NSCG questionnaire.

Appendix F includes the 2015 NSCG questionnaires for the new sample and returning sample. The questionnaires in Appendix F include coloring to identify changes to the questionnaire from the 2013 NSCG survey cycle.

### **Survey Methodological Experiments**

Three survey methodological experiments are planned as part of the 2015 NSCG data collection effort. Together, these experiments are designed to help NCSES and the Census Bureau strive toward the following data collection goals:

- Lower overall data collection costs
- Decrease potential for nonresponse bias in the NSCG survey estimates
- Increase or maintain response rates
- Increase efficiency in the use of incentives as part of a data collection methodology

The three methodological experiments are:

- Adaptive Design Experiment
- Paper Questionnaire Impact Experiment
- Email Reminder Experiment

The Adaptive Design Experiment, is planned for the both the new sample and the returning sample data collection efforts and the other two experiments listed above are planned for inclusion in the 2015 NSCG returning sample data collection effort. 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 Design Experiment

##### *2013 NSCG Adaptive Design Results*

The 2013 Adaptive Design Experiment (“2013 Experiment”) consisted of a 4,000 case representative sample selected from the 2013 NSCG new sample which was in turn selected out of respondents to the 2011 ACS. A representative control group that followed the standard data collection pathway was also identified for comparative purposes.

The primary objective of the 2013 Experiment was to evaluate whether data collection interventions could be implemented at the Census Bureau in the modes used by the NSCG: web, paper, and computer assisted telephone interviewing (CATI). Secondary objectives included implementing data monitoring to inform mode switching in a data-driven way and identifying the impact of mode switching on data quality in the form of response rates, R-indicators, cost, and effect on key estimates.

The primary objective of the 2013 Experiment was met successfully with data collection interventions occurring between weeks 4 and 23 of data collection. Data collection interventions included:

- Sending an unscheduled mailing to sample persons;
- Sending cases to CATI prior to the start of production CATI non-response follow up (NRFU), to target cases with an interviewer-assisted method rather than limiting contacts to self-response methods;
- Putting CATI cases on hold, to reduce contacts in interviewer-assisted modes, while still requesting response in self-response modes;
- Withholding paper questionnaires while continuing to encourage response in the web mode in order to reduce the operational and processing costs associated with “overrepresented” domains; and
- Withholding web invites to discourage response in “overrepresented” domains, while still allowing these cases to respond using previous invitations.

Meeting the primary objective confirmed that mode switching during data collection is an operational possibility for the implementation of adaptive design in the NSCG and at the Census Bureau.

The secondary objectives were met by the 2013 Experiment. R-indicators (measures of representativeness) were actively monitored throughout the data collection process in order to inform data collection interventions. R-indicators identify over- and under-represented domains using the within-domain variance of response propensities allowing the incorporation of data-driven interventions to improve the representativeness of the NSCG respondent population. Cases in under-represented subgroups, directly identified by the R-indicator, were moved to CATI prior to the start of production CATI NRFU to increase the variety of contact attempts made on these cases. Cases in over-represented subgroups were subject to one or several interventions that decreased the number of contact attempts, or otherwise restricted the response modes available to these cases in order to save on data collection resources or processing costs.

After the 2013 NSCG data collection period, we compared the adaptive design cases to the control cases on response rates, R-indicators, cost, and effect on key estimates. These comparisons were made at the full group level as well as at various subgroup levels. However, small sample sizes produced large confidence intervals around estimates, resulting in only one significant difference (overall weighted response rate for the adaptive design group versus the control) out of the many metrics and indicators compared. The secondary objectives required a larger sample size to effectively test. The 2015 NSCG will include an adaptive design experiment that builds on the 2013 Experiment with an increased sample size and a broader scope.

#### *2015 NSCG Adaptive Design Experiment*

All 2015 NSCG new sample and returning sample cases are eligible for the 2015 NSCG Adaptive Design Experiment. We require a representative sample of cases with multiple contact types (address, telephone number, email, etc.) and cases needing future research because they only have one contact type. This representative sample is necessary to make generalizations about how implementing adaptive design in the NSCG would affect the entire sample. The incorporation of adaptive design techniques creates the potential for NCSES and the Census

Bureau to develop a more efficient data collection process that reduces the cost of data collection and increases representivity of the responding sample cases.

Appendix H discusses the adaptive design goals, the interventions, and the monitoring metrics for the experiment.

The sample size for the treatment groups in the adaptive design experiment will be:

- Adaptive design new sample treatment group – 8,000 cases
- Adaptive design returning sample treatment group – 10,000 cases

Appendix I provides information on the minimum detectible differences achieved by these sample sizes.

#### Paper Questionnaire Impact Experiment

Mailing information to respondents is costly and questionnaire packets add additional expense since they are larger and heavier than a letter. The postage for sending a questionnaire via first class mail (\$1.08) is more than double the postage for a web invite letter (\$0.45). In addition to increased postage and the printing costs, paper questionnaires incur additional cost upon their return due to check-in, scanning and keying, and possible pre-key editing and failed edit follow up.

Prior to the 2013, the paper questionnaire was the data collection mode by which the largest percentage of respondents completed the NSCG. The 2013 survey marked the first cycle where more respondents completed the NSCG using the web rather than a paper questionnaire. Given this trend toward increased survey response by the web, the effectiveness of the paper questionnaire is an issue that deserves further attention. Thus, the 2015 NSCG will include an experiment to examine the impact of using a paper questionnaire in the data collection effort.

Currently, the NSCG default data collection pathway includes sending a paper questionnaire at week 7 at the beginning of the alternative mode phase. Then, another paper questionnaire is again sent at week 18 via priority mail during the telephone nonresponse follow-up phase. See Appendix G for more information on the 2015 NSCG default data collection pathway

Based on past response patterns, we expect to mail questionnaires to approximately 70,000 cases at week 7 and 30,000 cases at week 18. Given the extremely large number of cases that receive questionnaires during these weeks, we propose an experiment to examine whether we can change the way we use paper questionnaires in our NSCG data collection strategy without any adverse impact on cost, data quality, and response. The specific data collection strategy that we will test is to use a web invite in place of a planned questionnaire mailing. The research goals are:

- to understand the role that questionnaire mailings play in our overall contact strategy;
- to examine whether the questionnaire mailing and web invite mailing result in demographic differences in the responding sample. Investigating the demographic

distribution of the responding sample will allow the examination for potential bias reduction associated with the proposed treatments;

- to determine the cost, data quality, and response impact of using a web invite in place of a questionnaire mailing; and
- to inform data collection decisions for the 2017 NSCG.

All 2015 NSCG returning sample cases with a valid address on file that were not selected for the adaptive design experiment are eligible for the questionnaire impact experiment. To test the effectiveness of the paper questionnaire, our experiment will include a control group and three treatment groups. The eligible cases will be randomly allocated across the treatment groups with the majority of the cases in the control group (following our past practices).

<b>Experiment Groups</b>	<b>Sample Size</b>	<b>Week 7 (first class mailing)</b>	<b>Week 18 (priority mailing)</b>
Control (default path)	60,000	Paper questionnaire with web invite	Paper questionnaire with web invite
Treatment Group #1 (questionnaire at week 7)	3,500	Paper questionnaire with web invite	Web invite
Treatment Group #2 (questionnaire at week 18)	3,500	Web invite	Paper questionnaire with web invite
Treatment Group #3 (no questionnaires)	3,500	Web invite	Web invite

Appendix I provides information on the minimum detectable differences achieved by these sample sizes.

#### Email Reminder Experiment

When examining data collection costs, sending an email is more cost efficient than any type of postal contact since an email contact does not incur any postage or printing cost. In addition to cost savings, email has the added benefit that it is a different type of contact, contrasting against traditional postal mail that is used for the majority of the NSCG contacts. Moreover, an email contact can include a link to the web survey that allows the respondent to click or copy and paste the link into their web browser. The same can be done for the userid and password, essentially eliminating some of the user error with typing. While there are both cost and ease-of-use advantages to including email reminders as a contact type, limited research exists on how to incorporate email reminders into a full-scale data collection effort. In response to this limitation, the 2015 NSCG will include an experiment to examine the effectiveness of email reminders and letter reminders at different points in the data collection effort.

In the NSCG, email addresses are only available for returning sample cases that provided the email address while responding to the NSCG during a previous survey cycle. Currently, the NSCG default data collection pathway includes email reminders at week 16 and week 20. This experiment will examine both the appropriateness of using email reminders during these two contacts and the appropriateness of email reminders at two other weeks during the data collection effort – weeks 5 and 24. The specific data collection strategy that we will test is to use an email

reminder in place of planned postal letter reminders at weeks 5 and 24, and using a postal letter reminder in place of planned email reminders at weeks 16 and 20. The research goals are:

- to understand the role that email reminders play in our overall contact strategy;
- to examine whether an email reminder results in demographic differences for the responding sample than postal letter reminders. Investigating the demographic distribution of the responding sample will allow the examination for potential bias reduction associated with the proposed treatments;
- to determine the cost, data quality, and response impact of using an email reminder in place of a postal letter reminder; and
- to inform data collection decisions for the 2017 NSCG.

All 2015 NSCG returning sample cases with both a valid address and valid email address on file that were not selected for the adaptive design experiment and were not included in the CATI first data collection pathway are eligible for the email reminder experiment. To examine the appropriateness of email reminders, our experiment will include a control group and three treatment groups.

<b>Experiment Groups</b>	<b>Sample Size</b>	<b>Week 5</b>	<b>Week 16</b>	<b>Week 20</b>	<b>Week 24</b>
Control (default path)	45,000	Letter	Email	Email	Letter
Treatment Group #1 (all emails)	3,500	Email	Email	Email	Email
Treatment Group #2 (all letters)	3,500	Letter	Letter	Letter	Letter
Treatment Group #3 (opposite of default)	3,500	Email	Letter	Letter	Email

Appendix I provides information on the minimum detectable differences achieved by these sample sizes.

To complement the 2015 NSCG email reminder experiment, NCSES has implemented an experiment to test email contact strategies in the 2015 Survey of Earned Doctorates (SED). NCSES plans to use the results from the two studies to aid in the development of more comprehensive contact strategies that include email contacts in future NCSES surveys. While both experiments are designed to help NCSES better understand the role of emails in our overall survey contact strategy, there are unique differences between the two experiments that justify the implementation of both studies.

Major differences between the studies include the survey target population, contact history of the cases eligible for each study, and the manner in which email addresses are obtained within each survey.

- Survey target population – While the NSCG is attempting to collect information from the entire college-educated population in the United States, the SED is focused on a unique subset of this population – doctorate recipients that earned their degree from a U.S. educational institution in the previous academic year.
- Contact history – For the NSCG email reminder experiment, returning sample members that completed the NSCG in a previous survey cycle are eligible for the experiment. The SED email experiment is examining the impact of emails on late-stage nonrespondents who have never completed the SED. Since an individual’s past contact history is a strong predictor of response propensity, it is possible that the eligible cases selected for the two surveys will react differently to the experimental treatment.
- Email address source – In the NSCG, returning sample cases provided their email address to the NSCG survey contractor while responding to the NSCG during a previous survey cycle. In the SED, educational institutions provide the SED survey contractor with email addresses as a locating tool. It is possible that obtaining the email addresses directly from the individuals may result in the NSCG sample cases having a more favorable attitude to email contacts than the SED sample cases.

### Designing the Sample Selection for the 2015 NSCG Methodological Experiments

Three methodology studies are proposed for the 2015 NSCG returning sample portion: the adaptive design experiment, the paper questionnaire impact experiment, and the email reminder experiment. This section describes the sample selection methodology that will be used to create representative samples for each treatment group within the three experiments.

The eligibility criteria for selection into each of the studies are:

- Adaptive Design Experiment
  - All cases are eligible for selection
- Questionnaire Impact Experiment
  - Returning sample case
  - Has a valid address on file
  - Not selected in the adaptive design experiment
- Email Reminder Experiment
  - Returning sample case
  - Has a valid address on file
  - Has a valid email address
  - Not selected in the adaptive design experiment
  - Not included in the CATI-first data collection pathway



The sample for the adaptive design experiment will be selected independently of the sample selection for the other two experiments. Keeping the adaptive design cases separate from the other experiments will allow maximum flexibility in data collection interventions for these cases. The adaptive design experiment will select its control and treatment sample using a systematic random sample selection approach. This approach ensures that the adaptive design control group, the adaptive design treatment group, and the set of cases not selected for the adaptive design experiment each provide an unbiased representation of the returning sample population. Unbiasedness for each of these groups enables the results from all three methodological experiments to be generalized.

A *nested* experimental design will be used for the questionnaire impact and email reminder experiments. A nested design allows evaluation of the robustness of each intervention in the presence of other data collection strategies. Within the nested design, controls will be used to ensure similar proportions of questionnaire impact control and treatment group cases are selected into each email reminder treatment group. A systematic random sample selection approach will be used to select the samples for the questionnaire impact and email reminder studies.

The main steps associated with the sample selection for the 2015 NSCG methodological studies are described below.

*Step 1: Identification and Use of Sort Variables*

Since 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 2015 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 all three methodological studies. The 2015 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:

- Adaptive Design Experiment sort variables
  - Incentive indicator
  - Valid phone number indicator
  - 2015 NSCG sampling cell and sort variables
- Questionnaire Impact Experiment sort variables
  - Incentive indicator
  - Eligibility for email reminder experiment indicator
  - 2015 NSCG sampling cell and sort variables
- Email Reminder Experiment sort variables
  - Incentive indicator
  - Questionnaire impact experiment eligibility indicator
  - Questionnaire impact experiment control/treatment group indicator
  - 2015 NSCG sampling cell and sort variables

### *Step 2: Select the Samples*

For the new sample adaptive design experiment, a systematic random sample of approximately 8,000 cases will be selected to the treatment group. All cases not selected for sample into the treatment group will be assigned to the control group (approximately 34,000 cases).

For the returning sample adaptive design experiment, we will select a systematic random sample to ensure that both the control and treatment groups are representative of the returning sample population. In total, approximately 10,000 cases will be selected for the returning sample adaptive design control group and approximately 10,000 cases will be selected for the treatment group.

For the questionnaire impact experiment, we will select a systematic random sample to ensure the treatment groups are representative of the population eligible for the questionnaire impact experiment. Approximately 3,500 cases will be selected for each of the three questionnaire impact treatment groups. All eligible cases not selected for sample into the treatment groups will be assigned to the questionnaire impact control group (approximately 60,000 cases).

For the email reminder experiment, we will select a systematic random sample to ensure the three treatment groups are as similar as possible and representative of the population eligible for the email reminder experiment. Approximately 3,500 cases will be selected for each of the three email reminder treatment groups. All eligible cases not selected for sample into the treatment groups will be assigned to the email reminder control group (approximately 45,000 cases).

### Minimum Detectable Differences for the 2015 NSCG Methodological Experiments

Appendix I provides the minimum detectable differences associated with the 2015 NSCG methodological experiments.

### Analysis of Methodological Experiments

For all three experiments, 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 square 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 hard-to-enumerate.

## **5. CONTACTS FOR STATISTICAL ASPECTS OF DATA COLLECTION**

Chief consultant on statistical aspects of data collection at the Census Bureau is Benjamin Reist, NSCG Survey Director – (301) 763-6021. The Demographic Statistical Methods Division will manage all sample selection operations at the Census Bureau.

At NCSES, the contacts for statistical aspects of data collection are Jeri Mulrow, Acting NCSES Chief Statistician – (703) 292-4784, and John Finamore, NSCG Project Officer – (703) 292-2258.