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). In this design, a new ACS-based sample of college graduates will be 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 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[20] of the previous design including the undercoverage of key groups of interest such as foreign-degreed immigrants with S&E degrees and individuals with non-S&E degrees who are working in S&E occupations.

The NSCG design for the 2010 decade oversamples cases in small cells of particular 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*.

The 2017 survey cycle marks the full implementation of the four-panel rotating panel design that began with the 2010 NSCG. Under this fully implemented rotating panel design, the 2017 NSCG will include 123,500 sample cases which includes:

1) Returning sample from the 2015 NSCG (originally selected from the 2009 ACS);
2) Returning sample from the 2015 NSCG (originally selected from the 2011 ACS);
3) Returning sample from the 2015 NSCG (originally selected from the 2013 ACS); and
4) New sample selected from the 2015 ACS.

About 48,000 new sample cases[21] will be selected from the 2015 ACS. The remaining 75,500 cases will be selected from the set of returning sample members. While most of the returning sample cases are respondents from the 2015 NSCG survey cycle, about 14,000 nonrespondents from the 2015 NSCG survey cycle will be included in the 2017 NSCG sample. These 14,000

[20] Prior to 2010, the NSCG selected its sample once each decade from the decennial census long form. NSCG respondents educated or working in S&E fields were then followed biennially throughout the decade. Since no additional NSCG sample was selected throughout the decade, the previous NSCG design suffered from undercoverage of immigrants who entered the U.S. during the decade and individuals who began working in an S&E occupation during the decade.

[21] The 48,000 case sample size for the 2017 NSCG new sample portion is an 8,000 case increase from the sample size for the 2015 NSCG new sample. This sample size increase is being implemented in response to the increased size of the college-educated population as well as lower response rates in recent survey cycles for new sample cases.

cases are individuals that responded in their initial survey cycle, but did not respond during the 2015 NSCG survey cycle. These previous-cycle nonrespondents are being included in the 2017 NSCG sample in an effort to reduce the potential for nonresponse bias in our NSCG survey estimates.

The 2017 NSCG survey target population includes all U.S. residents under age 76 with at least a bachelor's degree prior to January 1, 2016. The new sample portion of the 2017 NSCG will provide complete coverage of this target population. The returning sample, on the other hand, will provide only partial coverage of the 2017 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 prior to January 1, 2014.

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 the 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; and 4) provides an oversample of young graduates to allow continued detailed estimation of the recent college graduates population.

Using the 2015 NSCG final response rates as a basis, NCSES estimates the final response rate for the 2017 NSCG to be 70 to 80 percent.

2. SURVEY METHODOLOGY

Sample Design and Selection

As part of the 2017 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 2017 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 2009 ACS or the 2013 ACS. The only portion of the returning sampling frame that will have a sample reduction is the portion of cases that originated in the 2011 ACS. These cases will receive a 50% sample maintenance reduction as part of the planned implementation of the NSCG rotating panel design[22]. 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

[22] The NSCG began using the ACS as a sampling frame in the 2010 survey cycle. To help with the transition to the four-panel rotating panel sample design, the 2010 NSCG selected two panels of cases from the 2009 ACS and the 2013 NSCG selected two additional panels of cases from the 2011 ACS. As a result, the 2013 NSCG sample included four panels with two panels originating from the 2009 ACS and two panels originating from the 2011 ACS. Continuing the four-panel rotating panel design, the 2015 NSCG included a new panel selected from the 2013 ACS and removed one of the two panels that originated from the 2009 ACS. Finally, in the 2017 NSCG, we are fully implementing the four-panel rotating panel design with panels selected from four separate ACS years by including a new panel selected from the 2015 ACS and removing one of the two panels that originated from the 2011 ACS.

sample to allow for reliable estimation. Since the 2017 NSCG will include new sample selected from the 2015 ACS, a portion of the returning ACS-based sample is no longer needed. As a result, only 50% of the returning cases that originated in the 2011 ACS will be selected for the 2017 NSCG sample. This 50% maintenance cut will occur across all returning cases that originated in the 2011 ACS regardless of their 2015 NSCG final interview disposition.

The sample selection for the 2017 NSCG new sample will use stratification variables similar to what was used in the 2015 NSCG. These stratification variables will be formed using response information from the 2015 ACS. The levels of the 2017 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

- 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 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)

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
- 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[23], 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 2017 NSCG.

- 2015 ACS sample cases with a bachelor's degree who are ages 30 or less and are educated or employed in an S&E field
- 2015 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 2017 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 2017 NSCG coefficient of variation targets that drive the 2017 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 2017 NSCG oversampling strata.

In total, the ACS-based sampling frame for the 2017 NSCG new sample portion includes approximately 1,040,000 cases representing the college-educated population of 65 million residing in the U.S. as of 2015. From this sampling frame, 48,000 new sample cases will be

[23] Since the young graduate 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.

selected based on the sample allocation reliability requirements discussed in the previous paragraph.

Weighting Procedures

Estimates from the 2017 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 defined as the probability of selection into the 2017 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 2015 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 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 2017 NSCG nonrespondents when forming survey estimates. The weight of the nonrespondents will be redistributed to the respondents and ineligibles within the 2017 NSCG sample. 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 implement 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 2017 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.

Degree Undercoverage Adjustment

The 2017 NSCG new sample does not have complete coverage of the population that first earned a degree during 2015, the ACS data collection year. For example, an ACS sample person that earned their first degree in May 2015, would be eligible for selection into the NSCG if their household was interviewed by ACS in July 2015 (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 2015 (i.e., before they earned their first degree).

Given that individuals who earned a degree after their ACS interview date are not eligible for the NSCG, the 2017 NSCG has undercoverage of individuals with their first degrees earned in 2015. To ensure the 2017 NSCG provides coverage of all individuals with degrees earned during 2015, a weighting adjustment is included in the 2017 NSCG weighting processing to account for this undercoverage. The Census Bureau conducted research on weighting adjustment methods during the 2013 NSCG cycle and an iterative reweighting model-based method was chosen to adjust the weights for this undercoverage. 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), 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 123,500 NSCG sample cases.

Replicate Weights

Sets of replicate weights will also be constructed to allow for separate variance estimation for the new sample and for each panel within the returning 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 replicate weights will be used to estimate the standard errors of the 2017 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 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 2015 NSCG, we will use different versions of the NSCG questionnaire for new sample cases and for returning respondents. 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 respondents. If these items were not reported by the returning respondents during their initial NSCG interval, the web and CATI instruments will attempt to collect this information this cycle.

In addition to the new sample questionnaire and questionnaire for previous cycle respondents, the 2017 NSCG cycle will use a third questionnaire for previous cycle nonrespondents. This questionnaire will be very similar to the questionnaire for previous cycle respondents with the exception of a slightly longer date range for the questions about recent educational experiences and the inclusion of four questions on community college enrollment that were not captured from cases that have not responded to the survey since the 2010 survey cycle.

The survey questionnaire items for the NSCG can be divided into two types of questions: core and module. Core questions are defined as those considered to be the base for the surveys. 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. 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.

For the 2017 survey cycle, the NSCG questionnaires will include two new items on veteran status that will be added to the set of core questionnaire items. These items will use the question wording currently used on the American Community Survey. The inclusion of these veteran status items will allow analysts to investigate the relationship between education and career pathways for the college-educated veteran population.

To partially offset the burden associated with the addition of these two veteran status items, the 2017 NSCG questionnaires will no longer include the questionnaire item requesting the respondent provide contact information for two individuals to aid in the locating of the respondent for future survey cycles.

Appendix E includes the 2017 NSCG questionnaires for the new sample cases. The other two NSCG questionnaires (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.

Nonsampling Error Evaluation

In an effort to account for all sources of error in the 2017 NSCG survey cycle, the Census Bureau will produce a report that will include information similar in content to the 2015 NSCG Nonsampling Error Report[24]. The 2017 NSCG Nonsampling Error Report will evaluate two areas of nonsampling error – nonresponse error and error as a result of the inconsistency between the ACS and NSCG responses. These topics will provide information about potential sources of nonsampling error for the 2017 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 2015 NSCG new sample portion, the overall weighted response rate was 63.3%; however, age groups had weighted response rates ranging from 52% for younger age groups, versus 70% 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 weighted response rate of 98% in the ACS followed by a weighted response rate of 73% in the initial NSCG survey cycle results in a compound response rate of just 72%. This means that only 72% of the cases originally eligible and sampled for the NSCG through the 2009 ACS exist in the current NSCG sample, with most of the attrition occurring in the initial round of the NSCG. 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.

[24] White, Michael, "2015 NSCG Nonsampling Error Report," Census Bureau Memorandum from Treat to Finamore and Rivers, August 2016, draft.

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 sources of sample for the 2017 NSCG (namely, the 2009 ACS, 2011 ACS, 2013 ACS, and 2015 ACS). 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 2017 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[25]. To the extent ACS responses are inconsistent with NSCG responses, using the

[25] 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.

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.

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 2015 NSCG and the 2015 ACS for the sample members will be used to locate the sample members in 2017.

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.

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 and 2015 NSCG data collection effort, the majority of the 2017 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 potential respondents.

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 NSCG incentive usage, NCSES is planning to use monetary incentives to offset potential nonresponse bias in the 2017 NSCG. We plan to offer a $30 prepaid debit card incentive to a subset of highly influential new sample cases at week 1 of the 2017 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 10,000 of the 48,000 new sample cases included in the 2017 NSCG. In addition, we will offer a $30 prepaid debit card incentive to past incentive recipients at week 1 of the 2017 NSCG data collection effort. We expect to offer $30 debit card incentives to approximately 7,000 of the 75,500 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.

As part of the 2015 NSCG production data collection effort, we included a questionnaire impact methodology experiment for returning sample cases that examined the benefit of offering paper questionnaires as a response option. This experiment included a control group that used our past data collection methodology and received a paper questionnaire at two predetermined dates within the data collection effort (weeks 8 and 18). In addition, the experiment included three treatment groups that replaced one or both of the questionnaire mailings at these two data collection dates with a web invitation letter. As a result, the experiment groups for this study were:

- Control group – Paper questionnaires mailed at week 8 and week 18
- Treatment group #1 – Paper questionnaire at week 8; Web invitation letter at week 18
- Treatment group #2 – Web invitation letter at week 8; Paper questionnaire at week 18
- Treatment group #3 – Web invitation letter at week 8 and week 18

This experiment[26] resulted in no significant difference in response across the experimental groups, but noticeable cost savings for the treatment groups that had received a web invitation letter in place of one or both of the paper questionnaire contacts. In response to the findings from this experiment, NCSES will no longer use a paper questionnaire at week 18 as part of our standard data collection pathway for the returning sample cases. Instead, we will use a web invitation letter for the week 18 contact.

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;

[26] Simoncini, Stephen, "2015 NSCG Paper Questionnaire Impact Study," Census Bureau Memorandum from Treat to Finamore and Rivers, August 2016, draft.

- 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 F for survey mailing materials and 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.

4. TESTING OF PROCEDURES

Survey Methodological Experiments
Two survey methodological experiments are planned as part of the 2017 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 and reduce respondent burden in the data collection methodology

The two methodological experiments are:

- Adaptive Design Experiment
- Contact Strategies Experiment

Both experiments are planned for the 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 Design Experiment

2015 NSCG Adaptive Design Results
The 2015 Adaptive Design Experiment ("2015 Experiment") was an expansion of the 2013 NSCG Adaptive Design Experiment ("2013 Experiment"). The two major expansions that occurred between 2013 and 2015 were an increase in the sample size, and the inclusion of

returning sample cases as part of the experiment. In 2015, the sample size of the new sample experiment was approximately 8,000 cases, doubling the sample size of the experiment from 2013. The returning sample had an experimental sample size of approximately 10,000 cases. Both of these samples were representative, and had control groups of comparable size. These sample sizes were provided by NCSES to be able to detect statistical significance for reasonable differences in the treatment versus control comparisons.

The primary objectives of the 2015 Experiment were threefold. First, NCSES wanted to formalize and prioritize the goals of adaptive design in the NSCG and determine the most appropriate interventions available to reach those goals. The second objective was to expand the number of interventions available for the NSCG. Finally, we needed to execute interventions in a way that would allow us to evaluate their effect on our data collection goals. In other words, we wanted to make fewer, larger interventions so their effect on data quality metrics would be more possible to measure. This was different from the primary goal of the 2013 Experiment, which was more about testing the operational functionality of adaptive design, which led us to execute many small interventions.

Secondary objectives included increasing the number of monitoring metrics available for informing data collection interventions and beginning to develop thresholds for when interventions should occur to further formalize adaptive design in the NSCG.

The primary objectives of the 2015 Experiment were met successfully. In response to a request from OMB[27], NCSES and the Census Bureau prepared a detailed adaptive design plan that included the time points where potential interventions would occur, the interventions that were available at each time point, the goals that would be served by executing interventions, and the criteria that would be used to determine whether an intervention was executed. This documentation served us throughout 2015 NSCG data collection effort, and given the similarities between the 2015 and 2017 NSCG data collection pathway, this documentation will continue to serve us in 2017, with minimal changes.

[27] OMB approved the 2015 adaptive design experiment and the collection of the 2015 NSCG under the following terms of clearance:

Within four weeks, NCSES will submit as a nonsubstantive change request, a "decision tree" or other descriptive rendering of the way that criteria will be applied to determine which non-responding sample cases will be assigned which treatments when in the adaptive design experiment. It will be important for the record to reflect how this experiment will be carried out. OMB approves NCSES starting production activities in the interim.

NCSES met these terms for the 2015 NSCG through a nonsubstantive change request submitted on June 17, 2015 that included a description and graphic of the 2015 NSCG data collection flow as well as a listing of the specific criteria planned for use in determining when adaptive interventions would be implemented. For the 2017 NSCG adaptive design experiment, Appendix H includes similar information on the data collection flow and the intervention schedule and criteria.

Additionally, working with teams in the various data collection modes, we were able to increase the number of interventions available to us for the 2015 NSCG. The full list of interventions used in the 2015 Experiment included:

- Sending an unscheduled web invitation to sample persons;
- Sending an unscheduled questionnaire 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 mode rather than limiting contacts to self-response modes;
- 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 previously offered modes.

Finally, our third primary objective was to make fewer, larger interventions so that we could more readily attribute changes in monitoring metrics to specific interventions. In the 2013 experiment, we made nine interventions (among new sample cases only) affecting 1,785 cases in total. In the 2015 Experiment, we made only five major interventions across both new sample cases and returning sample cases, but affected over 4,500 cases. In addition, the cases affected by these interventions were more evenly spread across both overrepresented and underrepresented cases, whereas in 2013, the larger interventions affected primarily overrepresented cases. We are still evaluating the effects of specific data collection interventions, but preliminary analyses show that our interventions for the returning sample cases resulted in statistically significant (α=0.10) differences in the full sample R-indicator and the unconditional variable-level partial R-indicator that persisted through the end of data collection.

This means that we saw a statistically significant improvement in sample balance due to our data collection interventions. There were statistically significant differences in these two metrics in the new sample cases near the time of the interventions, but these statistically significant differences did not persist through the end of data collection. At the same time, response rates between the treatment and control groups were not statistically different.

Meeting our primary objectives confirmed that adaptive design is a successful framework for improving the balance of our respondent population while maintaining response rates in the NSCG and at the Census Bureau.

NCSES and the Census Bureau also made progress on our secondary objectives. In the 2013 Experiment, we used R-indicators exclusively for making determinations about data collection interventions. In the 2015 Experiment, we expanded that list to include the number of cases available for interventions in subgroups of interest, the number of trips to locating, overall response propensity, and response propensity by mode. The first three additional metrics were

used to help us understand how many cases would be affected by potential interventions. The response propensities by mode were used to determine which cases should receive off-path questionnaires. While all of these items represented an expansion of the metrics, there is more progress to be made on this front. Fully meeting these objectives would have required us to settle on all of our statistical models of interest in advance of the 2015 Experiment. Given that the returning sample cases were being included in the adaptive design experiment for the first time, this proved difficult.

The 2017 NSCG will include an adaptive design experiment that builds upon the 2013 and 2015 Experiments with the same sample sizes as the 2015 Experiment, but with more fully automated data processing, data management, and statistical capabilities to better understand what would be required if we were to consider adaptive design in a full production roll-out, as opposed to an experimental setting.

2017 NSCG Adaptive Design Experiment

All 2017 NSCG new sample and returning sample cases are eligible for the 2017 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 and systematic automation 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 representativity 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 J provides information on the minimum detectible differences achieved by these sample sizes.

<u>2017 NSCG Contact Strategies Experiment</u>

As described in section B.3., sample members receive a variety of contacts throughout the data collection period including a pre-notification letter, web invitation, paper questionnaire, CATI invitation, reminder email, phone calls, and postcard. The new sample cases and returning sample cases receive slightly different materials, but the same types of contact on a similar schedule. Appendix G provide the current contact strategy for the 2017 NSCG new and returning samples. At a high level, sampled persons can receive approximately 12 mailing pieces and numerous phone calls during the six-month data collection period. This contact strategy can lead to high response rates (approximately 70% for new and 80% for returning

sample cases on average over the recent survey cycle), but it also leads to high costs and potentially over-burdens and frustrates sampled persons due to its high volume.

The type, timing and number of contacts can influence a respondent's decision to participate in a survey. Research has shown that multiple contacts with sample cases improve survey response rates especially when contacts are unique and present only pertinent information related to the survey.[28] Survey administrators, however, walk a fine line between maximizing response and overburdening sampled cases. OMB defines respondent burden as the "estimated total time and financial resources expended by the survey respondent to generate, maintain, retain, and provide survey information".[29] While the NSCG presents no cost burden to respondents, there is a time and contact burden. Contact burden refers to the number of times a sampled person is contacted regarding the survey. The primary goal of this experiment is to study various contact strategy approaches (i.e., types of contacts, number of contacts, messaging, and mode of contacts) to improve sample representativeness and reduce burden. The end goal is to develop a robust contact strategy based on data-driven results that optimizes sample representativeness and response while minimizing respondent burden and costs.

Through research conducted over the past year, the Census Bureau's Demographic Statistical Methods Division and the Census Bureau's Center for Survey Measurement have investigated contact strategies for the NSCG. The research evaluated both the number of contacts and their content. To assess the number of contacts, the researchers plotted 2015 NSCG daily response rates, along with contact mailing dates and telephone call dates and times, ran simulations to hypothesize the response rates with fewer contacts, and tracked the outcome codes of telephone calls. The research also used feedback from focus groups and cognitive interviews to develop and assess the content of the messages. Specifically, the research sought to determine what messages motivate people to respond, what modes of communication are preferable, and how people react to multiple contacts. The result of both the quantitative and qualitative components of this research enabled a proposal with three different contact strategies for NSCG to experimentally test as part of the 2017 NSCG Contact Strategies Experiment:

- Using a revised mailout strategy and new survey contact materials;
- The inclusion of an infographic as a contact type; and
- Reducing the number of call attempts.

The examination of these three different contact strategies creates a total of seven treatment groups that will be included within this experiment as displayed in the following table.

---

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

[29] Graham, John D. (2006). Questions and Answers when Designing Surveys for Information Collections. Office of Information and Regulatory Affairs, Office of Management and Budget Available at: https://www.whitehouse.gov/sites/default/files/omb/inforeg/pmc_survey_guidance_2006.pdf.

Contact Materials	Infographic	Call Attempt Reduction	Experimental Group
New Materials	Yes	Yes	Treatment Group #1
		No	Treatment Group #2
	No	Yes	Treatment Group #3
		No	Treatment Group #4
Production	Yes	Yes	Treatment Group #5
		No	Treatment Group #6
	No	Yes	Treatment Group #7
		No	Control Group

The experiment will be conducted separately for the new sample cases and the returning sample cases. The sample size for the treatment groups in the contact strategies experiment will be:

- Contact strategies new sample experimental group – 10,625 cases (approximately 1,328 per group)
- Contact strategies returning sample experimental group – 18,875 cases (approximately 2,359 per group)

Appendix I provides more detail about different contact strategies, the rational for their inclusion in the contact strategies experiment, and the research questions we are attempting to answer through this experiment. Appendix J provides information on the minimum detectible differences achieved by the sample sizes.

Designing the Sample Selection for the 2017 NSCG Methodological Experiments

Two methodology studies are proposed for the 2017 NSCG: the adaptive design experiment and the contact strategies 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:

- Adaptive Design Experiment
    - All cases are eligible for selection
- Contact Strategies Experiment
    - Cases not selected in the adaptive design experiment

The sample for the adaptive design experiment will be selected independently of the sample selection for the contact strategies 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 respondents. The main steps associated with the sample selection for the 2017 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 2017 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 2017 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
- 2017 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. For the new sample contact strategies experiment, a systematic random sample of approximately 1,328 cases will be selected in each of seven treatment groups. All eligible new sample cases not selected into the adaptive design or contact strategies treatment groups will be assigned to the control group (approximately 30,000 cases).

For the returning sample adaptive design experiment, a systematic random sample of approximately 10,000 cases will be selected to the treatment group. For the returning sample contact strategies experiment, a systematic random sample of approximately 2,359 cases will be selected into each of seven treatment groups. All eligible returning sample cases not selected into the adaptive design or contact strategies treatment groups will be assigned to the control group (approximately 49,000 cases).

Minimum Detectable Differences for the 2017 NSCG Methodological Experiments
Appendix J provides the minimum detectable differences associated with the 2017 NSCG methodological experiments.

Analysis of Methodological Experiments
In addition to the analysis discussed in the sections describing the 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 Stephen Simoncini, NSCG Survey Director – (301) 763-4816. 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 Samson Adeshiyan, NCSES Chief Statistician – (703) 292-7769, and John Finamore, NSCG Project Officer – (703) 292-2258.