

B. COLLECTION OF INFORMATION EMPLOYING STATISTICAL METHODS

1. RESPONDENT UNIVERSE AND SAMPLING METHODS

In the past, NSF has drawn a completely new sample for each decade from the decennial census long form. Beginning with the 2010 Census, the ACS replaced the long form. The NSF will use the ACS as a sampling frame for the 2010 NSCG and beyond. After reviewing numerous sample design options proposed by the NSF, 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 will allow the NSF to more efficiently target the 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 will select more cases in small cells of particular interest to analysts, including underrepresented minorities, women, persons with disabilities and non-U.S. citizens. This will result 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, the 2013 NSCG sample will include returning sample from the 2010 NSCG and 2010 NSRCG, as well as new sample selected from the 2011 ACS. The majority of the 2013 NSCG sample, 83,000 cases, will be selected from the 2011 ACS. This group is referred to as the new cohort sample. The other portion of the 2013 NSCG sample, 61,000 cases, will be from the 2010 NSCG (originally selected from the 2009 ACS) and the 2010 NSRCG. This group is referred to as the old cohort sample.

The 2013 NSCG survey target population will consist of all U.S. residents under age 76 with at least a bachelor's degree as of January 1, 2011. The new cohort sample will provide complete coverage of this target population. The old cohort sample, on the other hand, will provide only partial coverage of the target population. Specifically, the old cohort will cover the population of U.S. residents with a bachelor's or master's degree in a SEH field earned in the U.S. as of June 30, 2009 and U.S. residents under age 76 with at least a bachelor's degree as of January 1, 2009. Unbiased population estimates will be possible with careful weighting.

There are several advantages of this rotating panel sample design. It will: 1) permit longitudinal analysis of the retained cases from the 2009 ACS-based sample; 2) permit benchmarking of estimates to population totals derived from the sample using ACS; 3) maintain 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) provide an oversample of young graduates to allow continued detailed estimation of the recent college graduates population; and 5) allow 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 ACS.

There are two different versions of the questionnaire, one for each cohort. The main difference in the two questionnaires is the degree history questionnaire grid and certain demographic questions (e.g., race, ethnicity, and gender) that are not asked in the old cohort questionnaire, since this information has been collected for all old cohort sample cases in past survey cycles.

The target response rate for the new cohort is approximately 80 percent. The target response rate for the old cohort is approximately 90 percent. NSF targets these response rates based on 2010 final response rates for the new and old cohorts.

2. SURVEY METHODOLOGY

The old cohort portion of the 2013 NSCG frame will be sampled separately from the new cohort portion. The old cohort sample will be selected with certainty from the eligible sampling frame.

The 2013 NSCG new cohort will use stratification variables similar to what was used in the 2010 new cohort. These stratification variables will be formed using response information from the 2011 ACS. The levels of the 2013 NSCG new cohort stratification variables are as follows:

Highest degree level

- bachelor's degree/professional degree
- master's degree
- doctorate degree

Occupation/Degree Field (a composite occupation variable (which is composed of S&E bachelor's field of degree and occupation))

- 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 desired⁸, an additional sampling stratification variable will be used to identify the oversampling areas of interest. As noted earlier in this document, we decided to use the following criteria to define the cases eligible for the young graduates oversample within the 2013 NSCG new cohort.

- ACS sample cases with a bachelor's degree who are ages 28 or less and are educated or employed in an SEH field
- ACS sample cases with a master's degree who are ages 32 or less and are educated or employed in an SEH field

The multiway cross-classification of these stratification variables produces approximately 900 possible 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.

Research on the panel design of NSCG sample allocation has shown that NSCG can produce the estimates for the key domain statistics that can support the reliability target in the report. As a result, the 2013 NSCG reliability targets are aligned with the data needs the NSF congressionally mandated reports. Final sample allocation for strata in both the old and new cohorts will be determined after detailed analysis of the 2011 ACS sampling frame. The sample allocation will be determined based on reliability requirements for key NSCG analytical domains provided by the NSF. The 2013 NSCG Coefficient of Variation Targets that will drive the sample allocation are included in Appendix C. Tables 1, 2, and 3 of Appendix C provide reliability requirements for estimates of the total college graduate population. Tables 4, 5, and 6 of Appendix C provide

⁸ 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.

reliability requirements for estimates of young graduates, which are the target of the 2013 NSCG oversampling strata.

In total, the ACS-based sampling frame for the 2013 NSCG new cohort includes over 800,000 cases representing the college-educated population of 56 million residing in the U.S. as of 2009. From this sampling frame, the 83,000 new cohort sample cases will be selected based on the sample allocation reliability requirements discussed in the previous paragraph.

Estimates from the 2013 NSCG will be based on standard weighting procedures. As was the case with sample selection, the weighting adjustments will occur separately for samples from each cohort. The goal of the separate weighting processing is to produce individual cohort final weights that reflect the respective cohort population. To produce the individual cohort final weights, each case will start with a base weight defined as the probability of selection into the 2013 NSCG sample. This base weight reflects the differential sampling across strata. Base weights will be then be adjusted to account for noninterviews.

Following the weighting methodology used in the 2010 NSCG, we will use propensity modeling as an alternative to the cell collapsing approach 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 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. This logistic regression model will use response as the dependent variable. The propensities output from the model will be used to categorize cases into noninterview cells of approximately equal size, with similar response propensities in each cell. The noninterview weighting adjustment factors will be calculated within each of the noninterview cells.

The noninterview weighting adjustment factor is used to account for the weight of the 2010 NSCG nonrespondents when forming survey estimates. The weight of the nonrespondents will be redistributed to the respondents and ineligibles within the 2010 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. In response to this weight distribution plan, two different noninterview adjustment factors will be used: one for respondents and one for ineligibles. Using these two factors ensures that the eligible nonrespondents' weight only goes to eligible respondents and not to the ineligible cases.

After the noninterview adjustment, weights will be raked to ACS population totals through an iterative raking procedure that ensures population totals are upheld.

After the completion of the weighting, 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, NSF 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 last step in the 2013 NSCG individual cohort weighting processing will be an additional execution of the iterative raking procedure. After the additional execution of the iterative raking procedure, the resulting weight will be the individual cohort final weight.

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), NSF will combine the new and old cohort together and form combined cohort weights. The combined cohort weights will be formed by adjusting the two sets of individual cohort final weights to account for the overlap in target population coverage. The result will be a combined cohort final weight for all 144,000 NSCG sample cases.

Replicate Weights. Two replicate weights will also be constructed separately to allow for separate variance estimation for the old cohort sample and the new cohort sample. The replicate weight for the combined cohort estimates will be constructed from these individual cohort replicated 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 2013 NSCG estimates as in the past. 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.

3. METHODS TO MAXIMIZE RESPONSE

Maximizing Response Rates

In order to maximize the overall survey response rate, NSF and the Census Bureau will implement procedures such as conducting extensive locating efforts and collect the survey data using three collection modes (mail, web, and CATI). The contact information obtained from the 2010 NSCG, 2010 NSRCG, and the 2011 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 2013.

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 of the 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, which 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 include 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), NSF 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.

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, the majority of the 2013 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, so the initial offering of a web response option is apt to be appealing to NSCG respondents (including the NSRCG panel sample members.)

Motivated by the findings from the late stage incentive included in the 2010 NSCG in combination with the desire to obtain a better understanding of optimal incentive usage in data collection efforts, the NSF is considering two monetary incentive experiments to examine potential nonresponse bias in the 2013 NSCG: an incentive timing study and an incentive conditioning study. The incentive timing study will examine the impact that the timing of when an incentive is offered has on response rate, sample representation, bias reduction, and cost. The incentive conditioning study will examine the impact that a previous incentive has on a sample case's propensity to respond in a subsequent survey cycle. In addition, the incentive conditioning study will examine the sample representation, bias reduction, and cost associated with different data collection techniques for interviewing previously incentivized cases.

The incentive in both studies will be a \$30 prepaid debit card incentive that is similar to the debit card incentive used in the 2010 NSCG survey cycle. 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 NSF (minus the predetermined per card fee).

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
- 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 the Appendix E for survey mailing materials.

4. TESTING OF PROCEDURES

Because data from the SESTAT surveys are combined into a unified data system, the surveys must be closely coordinated to provide comparable data from each survey. Most questionnaire items in the two surveys are the same.

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.

All content items in the SESTAT survey questionnaires had undergone an extensive review and testing before they were included in the 2013 questionnaire. Compared to the 2010 NSCG, the 2013 NSCG questionnaires will include five new items.

During our outreach efforts for redesigning SESTAT, we received feedback that adding the NSRCG questions on community colleges and financial aid to the NSCG would meet the content needs that might have been lost with discontinuing the NSRCG. Thus, with the discontinuation of the NSRCG and the NSCG oversampling young graduates as a substitute for the NSRCG, we will include the following questions on both the new cohort and old cohort questionnaires for the

2013 NSCG survey cycle. In future NSCG survey cycles, these questions will be included only on the NSCG new cohort questionnaire. These are questions that had been previously included on the NSRCG questionnaire and were deemed of high analytical interest by data users and policy makers.

The specific wording for the community college questions will be derived from the 2008 NSRCG questionnaire and the specific wording for the financial aid questions will be derived from the 2006 NSRCG questionnaire. Prior to the 2015 NSCG survey cycle, NSF plans to conduct testing of the question wording for both the community college and financial aid questions. This testing will be conducted in an effort to determine whether these questions remain appropriate from a content, cognitive, and usability perspective.

Community College Questions

- Did you take courses at a Community College during any of the following time periods?
- For which of the following reasons did you take Community College courses?
- Which two reasons were your most important reasons?

Financial Aid Questions

- What types of financial support have you received to finance your undergraduate or graduate degrees?
- What is the amount you have borrowed to finance your undergraduate and graduate degrees and how much do you still owe?

In addition to the questions on community college and financial aid, the 2013 NSCG questionnaire content has been revised from 2010 as follows:

- Changed the survey reference date from October 1, 2010 to February 1, 2013.
- Modified the response option wording for the residence question (E7 on the new cohort questionnaire and F6 on the old cohort questionnaire)
- Modified the response option wording for the employer size question (A11 on both questionnaires)
- Modified the response option wording for the federal agency support question (A37 on both questionnaires)

A complete list of questions proposed to be added, dropped, or modified in the 2013 NSCG questionnaires is included in Appendix D.

Survey Methodology Studies to be Undertaken

Four survey methodological studies are planned as part of the 2013 NSCG data collection effort. Together, these studies are designed to help the NSF and Census Bureau strive toward the following data collection goals:

- Lower overall data collection costs
- Decreased potential for nonresponse bias in the NSCG survey estimates
- Higher response rates
- Increased efficiency in the use of incentives as part of a data collection methodology

The four methodological studies are as follows:

- Incentive Timing Study
- Mode Switching Study
- Priority Mailing Study
- Incentive Conditioning Study

The first three studies listed above are planned for inclusion in the 2013 NSCG new cohort data collection effort and the fourth study, the Incentive Conditioning Study, is planned for the old cohort data collection effort. This section introduces the design for each study, describes the research questions each study is attempting to address, and includes information on the sample selection proposed for these studies. The information in this section is supplemented by the information in Appendix G for the Incentive Timing Study and the information in Appendix H for the Incentive Conditioning Study.

Incentive Timing Study

Findings from the 2010 NSCG Late-Stage Incentive Study

The 2010 NSCG included a late-stage incentive study targeting hard to enumerate cases that had not responded to the survey after multiple contacts. As part of the study, the hard to enumerate cases were allocated to three treatment groups:

- \$30 debit card incentive
- \$20 debit card incentive
- No incentive

Other than the use and amount of the debit card incentive, the three treatment groups received the same data collection contact strategy. At the conclusion of the study period (approximately six weeks), the response rate for the three treatment groups differed significantly. The \$30 incentive treatment group had a response rate of 29.5%, the \$20 incentive treatment group had a response rate of 24.1%, and the no incentive group had a response rate of 6.4%.

In addition to the expected increase in the response rate for the hard to enumerate cases that were targeted as part of this study, the use of the incentive also had a profound effect on the overall representation of the responding sample. As described in Appendix G, the incentive was successful in obtaining responses from respondents who are demographically different than the set of respondents prior to the incentive stage. This ability to increase the demographic diversity of our responding sample helps decrease the potential for nonresponse bias in our estimates.

At the conclusion of the evaluation period for the 2010 NSCG late-stage incentive study, we have concluded that the study was successful in increasing response and lowering the potential for nonresponse bias. However, despite the success, there remained numerous unanswered questions regarding the use of the incentives as a data collection strategy including:

- Would a definition of hard to enumerate that include both a response component and a bias component allow us to better address both our response and bias needs?
- What incentive timing approach provides the optimal balance of response impact, bias reduction, and cost minimization?

The desire to answer these questions led to the development of the Incentive Timing Study for the 2013 NSCG new cohort data collection.

Development of the 2013 NSCG Incentive Timing Study

As we consider the plans for interviewing the hard-to-enumerate cases in the 2013 survey cycle, questions exist about whether the late-stage timing is the optimal way to use the incentive in future survey cycles. By offering an incentive at an earlier stage of data collection, we could obtain responses for these hard to enumerate cases and decrease the need for nonresponse follow-up efforts. However, the question remains of whether the cost of offering the incentive to a larger number of cases exceeds the nonresponse follow-up costs had the incentive not been offered at an earlier stage.

Furthermore, in the 2010 late-stage incentive study, the hard-to-enumerate cases were defined by their data collection response mode to the American Community Survey⁹. As a result, questions were raised about whether this definition best met both the response and bias needs of the survey.

⁹ The hard to enumerate cases in the 2010 NSCG Late-Stage Incentive Study were the NSCG nonrespondents that had responded to the ACS through the telephone or personal visit data collection modes.

In order to explore the issue of optimal incentive timing and the desire to address both response and bias in our data collection strategy, we are planning an incentive timing study as part of the 2013 data collection effort. The 2013 NSCG incentive timing study will include cases identified as hard to enumerate in a manner that is different than the definition used in the 2010 NSCG. Once identified, these hard to enumerate cases will be randomly allocated to five treatment groups as part of the 2013 NSCG data collection effort:

- No incentive
- Incentive offered at week 1 of data collection
- Incentive offered at week 7 (coincides with the introduction of a mail response option)
- Incentive offered at week 12 (coincides with the introduction of a telephone response option)
- Incentive offered at week 25 (coincides with the final contact phase)

The incentive in this study will be a \$30 prepaid debit card incentive that is similar to the debit card incentive used in the 2010 NSCG survey cycle. 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 NSF (minus the predetermined per card fee).

The incorporation of this study will allow for the investigation of the two research questions raised in the previous section, namely:

- Would a definition of hard to enumerate that include both a response component and a bias component allow us to better address both our response and bias needs?
- What incentive timing approach provides the optimal balance of response impact, bias reduction, and cost minimization?

Model-based Definition of Hard to Enumerate

The 2013 NSCG incentive timing study identifies influential, hard to enumerate cases and makes them eligible for one of five incentive experiment groups. The hard to enumerate cases are identified by a model-based approach using the weighted response influence, which is the product of a sampled cases baseweight and predicted response propensity. The weighted response influence formula is calculated as follows:

$$W_i = \omega_i * \hat{\phi}_i, \text{ where } \hat{\phi}_i = \left(\frac{1}{\hat{\rho}_{Li}} \right) \left(\frac{1}{\hat{\rho}_{Ri}} \right).$$

The weighted response influence for a case, W_i , is the product of the base weight, ω_i , and the response influence, $\hat{\phi}_i$. The response influence is the inverse of the product of the locating propensity, $\hat{\rho}_{Li}$, and the response propensity, $\hat{\rho}_{Ri}$.

Both the locating propensity and response propensity will be determined using results from the 2010 NSCG survey cycle. For both of these propensity models, numerous variables from the ACS-based sampling frame were examined as potential predictors. For the locating propensity model, a locate indicator was the independent variable in the model. For the response propensity model, a response indicator was the independent variable in the model.

The variables considered as possible predictors for both models were as follows (variable names are included in parenthesis). All two-way interactions between variables were also considered as possible predictors in the models.

- Demographic Group (ACS_DEMGROUP)
- Highest Degree Level (ACS_HIDEG)
- Detailed Occupation Group (ACS_OCC_DETAIL)
- S&E Degree Field Indicator (ACS_SEDEG)
- Gender (ACS_SEX)
- Age Group (AGEGROUP)
- Census Division (DIVISION)
- Mobility Status (MIG)
- ACS Data Collection Mode (MODE)
- Number of Persons in Household (NP)
- Person-Level Income (PINC)
- Relationship to Householder (DETAIL_REL_NOGQ)
- Tabulation Date (Interview Month) (TABDT)
- Tenure (TEN)
- Urban/Rural Indicator (UR)

For the locating propensity model, the following variables were identified as significant predictors of the ability to locate a case. Please note that the terms with an asterisk refer to interaction terms that were significant.

- ACS Data Collect Model (MODE)
- Demographic Group (ACS_DEMGROUP)
- Age Group (AGEGROUP)
- Relationship to Householder (DETAIL_REL_NOGQ)
- Tenure (TEN)
- Highest Degree Level (ACS_HIDEG)
- Detailed Occupation Group (ACS_OCC_DETAIL)
- Gender (ACS_SEX)
- Mobility Status (MIG)
- Number of Persons in Household (NP)
- S&E Degree Field Indicator (ACS_SEDEG)
- Census Division (DIVISION)
- ACS_HIDEG*AGEGROUP
- MODE*DIVISION

For the response propensity model, the following variables were identified as significant predictors of response. Please note that the terms with an asterisk refer to interaction terms that were significant.

- Detailed Occupation Group (ACS_OCC_DETAIL)
- ACS Data Collect Model (MODE)
- Highest Degree Level (ACS_HIDEG)
- Age Group (AGEGROUP)
- Gender (ACS_SEX)
- Relationship to Householder (DETAIL_REL_NOGQ)
- Census Division (DIVISION)
- S&E Degree Field Indicator (ACS_SEDEG)
- Tenure (TEN)
- ACS_HIDEG*AGEGROUP
- Demographic Group (ACS_DEMGROUP)
- ACS_SEX*TEN
- MODE*DETAIL_REL_NOGQ
- ACS_OCC_DET*DIVISION

For the 2013 NSCG incentive timing study, the results from the 2010 NSCG locating and response propensity models will be used to assign cases both a response propensity score and locating propensity score. The inverse of the product of these scores will be the response influence value discussed in the formula above. Results from the 2010 NSCG survey cycle will be used to determine the hard to enumerate status since the cases will be assigned to the incentive timing study treatment groups prior to the 2013 NSCG survey cycle. At the conclusion of the incentive timing study, analysis will be conducted using 2013 data to examine the accuracy of the hard to enumerate definition.

A low response propensity will result in a high inverse response propensity, what Särndal calls the response influence of a case. For NSCG, that response influence incorporates both the locating propensity and response propensity. That factor multiplied by the base weight of the case represents the magnitude of the effect that nonresponse would have on a given case's nonresponse cell in the NSCG. Once the weighted response influence is calculated, a natural cut off would be used to determine which cases were "hard to enumerate".

A large weighted response influence value identifies a case that has the potential to cause weight variation thus increasing variances after propensity-based nonresponse adjustments. Additionally, a large weighted response influence value identifies cases which will cause larger variance in the propensity to respond. A larger variance in the propensity to respond can be a good predictor of nonresponse bias (see representative indicators, and balancing set literature). For the NSCG both the weight of the nonrespondents and the propensity to respond are directly linked to the unit nonresponse bias mitigation strategy since NSCG uses propensity-based nonresponse cell weighting adjustments.

Mode Switching Study

As we consider the data collection efforts for the 2013 NSCG new cohort, it is unclear whether repeated contact attempts in a particular data collection mode are worthwhile before switching a case to a different mode. In the 2010 NSCG, the default design included mail and web invite letters and automated phone calls. Production CATI did not begin until Week 15 of the survey. By incorporating an adaptive design data collection strategy and changing response mode for a case earlier in the data collection process, we could decrease the cost of getting a completed survey by decreasing the number of unsuccessful contact attempts that occur in other modes. In addition, we would collect new information earlier in the data collection period about a given case's propensity for responding in a given mode, possibly allowing us to inform further mode switching decisions later in the data collection process for other similar cases.

From an operational standpoint, this test would allow us to observe and document some impediments to widespread implementation of flexible mode switching, particularly at the Census Bureau's telephone centers and National Processing Center. In addition, the question remains as to whether the flexibility of mode switching will reduce the cost of successfully completing cases versus the existing structured mode assignments currently in production. Finally, there is a question of how well the propensity of a case to respond in a given mode could be modeled, and whether using the mode associated with the highest-propensity will actually elicit higher response rates.

The 2013 NSCG mode switching study will include a representative sample of cases. These cases will be actively managed throughout the data collection process. This will allow for *ad hoc* mode switching based on quality and cost measures. Cases in the test group will all initially start in the web mode (following the default web first data collection approach). From that point, weekly meetings would be held to determine which cases, if any, should be moved from one mode into another. Another option is to extend the time a case remains in a particular mode. These mode switching decisions will be made using propensity models based on 2010 NSCG data in combination with the incoming 2013 NSCG results. The planned daily processing of 2013 NSCG responses will allow the NSCG survey team to monitor several quality measures throughout data collection period, including R-indicators, benchmarking, stability of estimates, and response propensities by mode. This experiment is an attempt to allocate resources more efficiently in order to maximize survey quality while minimizing wasted funds and effort.

There are several operational difficulties associated with implementing a mode switching test during production data collection. These difficulties include moving a case into and out of CATI in a repetitive fashion, and removing a case from the mail mode once it has been included in a mail batch. To implement a mode switching study, therefore, cases will be managed through one of the following discrete paths:

- No treatment (continue in the default data collection flow)
- Extended time in web mode, receiving an additional web invite at week 7 (before moving to default data collection)
- Early transfer to the CATI mode

Regardless of the decisions made during the mode switching study, all cases would move to the default data collection path at week 12. The benefit of mode switching over the default data collection flow is that the timing of the mode switch is not fixed. For example, if a case were of less interest in the NSCG, it could remain in the web mode for longer, reducing the cost of contacting that case via other methods. Likewise, if a case were more likely to respond in CATI, we could move it to CATI at an earlier point in the data collection process, eliminating the need to contact that case via other methods. Finally, being able to move cases into new modes at different times would allow us to update propensity models based on past years of data with current information, informing mode switches for other cases later in data collection.

This study is designed to provide insight on the following questions:

- Can mode switching be implemented operationally, even for a small sample?
- If we switch a case into the mode that has the highest propensity for completion earlier in the data collection period, are we more likely to get a completed interview?
- Is there a tangible savings of cost when mode switching is used?
- How well can we predict the mode in which a case will eventually respond?
- Is mode switching more effective at increasing response than a well-timed incentive? (This is a secondary objective, particularly if the sample allocated to this test is small.)
- Does mode switching improve quality indicators, including measurement error, representativeness of sample, and reduction of bias?

The 2013 NSCG new cohort has sample size of 83,000 cases. Cases will be eligible for the mode switching study if they have enough contact methods to actually switch modes. This would require a mailable address and a valid contact phone number. From the cases that meet these eligibility criteria, a randomly selected set of cases will be included in this study and initially assigned to the web mode. Case will be moved based on information from a number of sources including the importance of a case, propensity for response, and accrued cost, amongst others. The expected treatment group sample size for the mode switching study is 2,500 cases.

Priority Mailing Study

In the 2010 NSCG mode effects experiment, sample cases were mailed a survey invitation packet at the beginning of data collection to introduce the survey and encourage response. At the fifth week of data collection, a replacement packet was mailed to nonrespondents. Following the data collection implementation best practice of varying the look of each contact attempt, the 2010 NSCG replacement packet was sent via priority mail.

This priority mailing allowed the replacement packet to be sent in an envelope that was different from the envelope used with the survey invitation packet. In addition, the priority mailing allowed a quicker delivery to the respondent. Past research has shown that the use of priority mail results in an increase in response in a data collection methodology that emphasizes response

by mail. However, the use of priority mailing is extremely expensive compared to first class mailing postage rates. Mailing a standard business envelope via first class mailing would cost approximately \$0.45 whereas the same envelope sent via priority mailing would cost approximately \$5.00. Furthermore, the research supports the use of priority mail when encouraging response by mail, but the literature is limited on the benefits of priority mail when attempting to encourage response by a web-based data collection mode.

In the 2013 NSCG, we are expecting to mail replacement packets via priority mail to approximately 70% of our sample cases. So, the postage associated with this replacement mailing will be approximately \$500,000. Given this extremely high cost, we are proposing a study as part of the 2013 NSCG to examine whether it is possible to achieve our NSCG response rate goals by using a first class mailing for the replacement packets rather than a priority mailing. If the first class mailing for the replacement packets is deemed successful in encouraging response, the potential cost savings for future NSCG survey cycles is approximately \$400,000.

As part of this priority mailing study, we will randomly allocate cases to two treatment groups:

- Priority mailing for the replacement packet
- First class mailing for the replacement packet (using an envelope different from the one used in the survey invitation packet)

Consideration is also being given to a third treatment group that will use a mailing that is closer in price to the first class mailing, but provides the impression of importance like a priority mailing (e.g., certified mailing).

This study is designed to provide insight on the following research questions:

- Does the priority mailing increase response compared to a first class mailing with a different envelope?
- Does the priority mailing result in a demographically different responding sample than the first class mailing with a different envelope? Investigating the demographic distribution of the responding sample will allow the examination of the potential bias reduction associated with the proposed treatments.

The eligible cases for this priority mailing study includes NSCG new cohort cases with a mailable address who are not included in another 2013 NSCG study. The eligible cases will be randomly allocated across the treatment groups with the majority of the cases to be included in the priority mailing group (following our past practices) and a sample allocation for the first class mailing group that is large enough to meet our statistical significance testing detection level requirements.

Incentive Conditioning Study

Findings from the 2010 NSCG Late-Stage Incentive Study and the 2010 NSRCG Incentive Usage

As noted earlier, the 2010 NSCG included a late-stage incentive study targeting hard to enumerate cases that had not responded to the survey after multiple contacts. At the conclusion of the study period (approximately six weeks), the response rate for the three treatment groups differed significantly. The \$30 incentive treatment group had a response rate of 29.5%, the \$20 incentive treatment group had a response rate of 24.1%, and the no incentive group had a response rate of 6.4%. In addition, the incentive was successful in obtaining responses from respondents who are demographically different than the set of respondents prior to the incentive stage.

A 2008 NSRCG incentive experiment found that offering incentives to the recent graduates population at an early point in the data collection effort was effective from a response and cost perspective. Furthermore, it concluded that the use of a differential incentive based on response mode (\$20 for response by mail and \$30 for response by web) encouraged response by the mode offering the higher incentive. Given these findings, the 2010 NSRCG data collection strategy called for a web initial invitation letter that offered \$20 incentive for a completed questionnaire and provided the sample member with his or her user ID and password for completing the survey online. No paper questionnaire was included in the initial mailing. The second mailing, about five weeks later, included a paper questionnaire and offered the differential incentive. At the end of the 2010 data collection, this design helped increase the percentage of web completions from 67.5% in 2008 to 91.4% in 2010. Moreover, the percentage of CATI completes was reduced from 17.5% in 2008 to 6.6%. The paper questionnaire completes dropped from 15.5% to 2%.

Development of the 2013 NSCG Incentive Conditioning Study

As we consider the plans for interviewing these previously incentivized cases in the 2013 NSCG survey cycle, questions exist about whether these cases require an incentive to respond in future survey cycles and the potential bias implications if these cases do not respond to the survey. To answer these questions we propose an incentive conditioning study as part of the 2013 data collection processing. The 2013 NSCG incentive conditioning study will include cases that received an incentive in the 2010 survey cycle and responded to the survey. These cases will be randomly allocated to three treatment groups as part of the 2013 NSCG data collection effort:

- No incentive
- Incentive offered at the beginning of data collection
- Incentive offered at a late stage of data collection

The incentive in this study will be a \$30 prepaid debit card incentive that is similar to the debit card incentive used in the 2010 NSCG survey cycle. 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 NSF (minus the predetermined per card fee).

This study is designed to provide insight on the following questions:

- Do previous incentive recipients require an incentive to encourage response in future survey cycles?
- Demographically, how do the following groups differ:
 - Previous incentive recipients that responded in a future survey cycle without an incentive
 - Previous incentive recipients that responded in a future survey cycle with an early incentive
 - Previous incentive recipients that responded in a future survey cycle with a late stage incentive
- Using the information gathered in the evaluation of demographic differences, what data collection strategy could be used in future survey cycles to optimize bias reduction for previously incentivized cases?
- If an incentive is required in future survey cycles for previous incentive recipients, what is the optimal strategy for the offering of the incentive? Optimal should be measured in terms of response, bias, timing, and cost.

Please note that Appendix G provides additional information associated with the implementation of this study.

Designing the Sample Selection for the 2013 NSCG Methodological Studies

As noted earlier, three of the methodological studies will be implemented within the 2013 NSCG new cohort (the incentive timing study, the mode switching study, and the priority mailing study). This section describes the sample selection methodology that will be used to create fully representative samples for each treatment group within each of the three new cohort studies. The new cohort cases leftover (i.e., those not selected for any of the studies) will serve as the control and will also be fully representative of the NSCG sample population. The next sections outline the main steps associated with the sample selection for the 2013 NSCG methodological studies within the new cohort.

It should be noted that the fourth study, the incentive conditioning study, will be implemented within the 2013 NSCG old cohort study and, as a result, its sample selection will be independent of the other three studies.

Step 1: Identification and Use of Sort Variables

Since the samples for the treatment and control groups within the new cohort methodological studies will be identified using systematic sampling, the identification of sort variables and the use of an appropriate sort order is extremely important.

- The incentive timing study is limited by the number of incentives that will be mailed. Since the study targets hard to enumerate cases, identifying whether a case is defined as

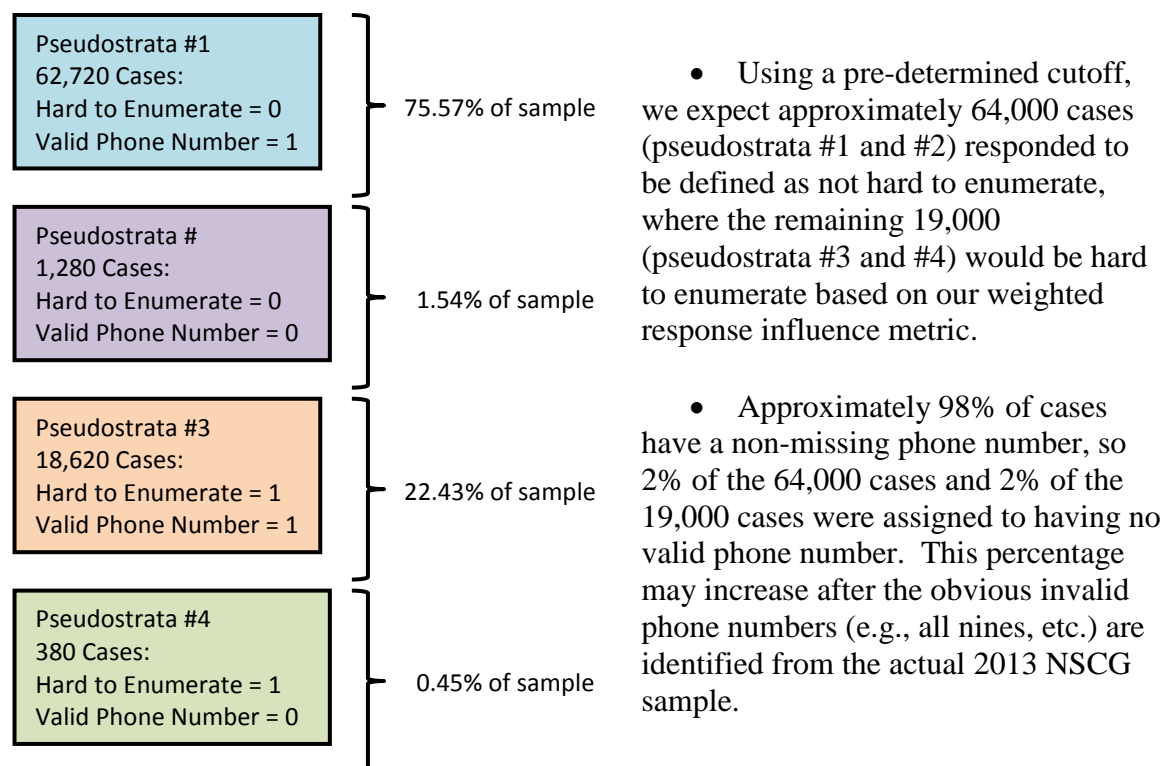
hard to enumerate is vital to keeping the study population representative of the sample population.

- The mode switching study requires contact information in two modes (mail and telephone), so identifying whether a case has a valid telephone number is necessary information for use in our sample selection.
- The priority mailing study requires nothing beyond a valid mailing address.

Therefore, our sort order will be:

- Hard to enumerate indicator
- Valid phone number indicator
- 2013 NSCG new cohort stratification variables

This sort will create four main pseudo-strata. Example strata sizes are estimated below. The use of the 2013 NSCG new cohort stratification variables allows for the selection of a representative sample on numerous demographic variables that are high interest to the NSCG.



Step 2: Select the Samples

For this example, we assume the following sample sizes:

- For the priority mailing study the treatment group (receiving first class mail) is 10,000.
 - This selection will be a systematic random sample.
- For the incentive timing study 3,000 hard to enumerate sample cases would be eligible to receive an incentive in each of 4 treatment groups
 - This selection will be a systematic random sample with cluster sizes of four, in order to select the four treatment groups simultaneously and help ensure the four treatment groups are as similar as possible.
 - Approximately 13,105 cases will be selected to each of the four treatment groups. Of these 13,105 cases, we only expect approximately 3,000 cases in each treatment group to be "hard to enumerate" and therefore receive an incentive. However, all of the 13,105 cases selected in each treatment group would maintain a representative sample in order to allow for comparisons with the control and other treatments.
- For the mode switching study 2,500 cases would be put in the flexible path
 - This selection will be a systematic random sample.

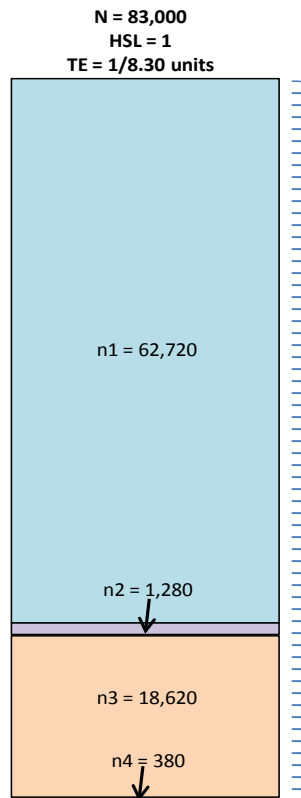
The following page has a visual step-by-step of the sample selection.

This sample selection methodology creates fully representative samples for each treatment group within each of the three new cohort studies. The cases leftover (i.e., those not selected for any of the studies) will serve as the control and will also be fully representative of the NSCG sample population. Another approach considered for the sample selection was a factorial design. However, the factorial design would not have allowed the ability to generalize the results to the level of the full survey population as easily as this proposed methodology. In addition, teasing out main effects could be very difficult in a factorial design due to the ability for a case to be in more than one methodological study. In other words, a factorial design makes it more difficult to determine the effect of a single treatment on survey parameters like cost or responses. The proposed sample selection methodology removes possible confounding effects by ensuring each case is in only one study.

Priority Mail Study:

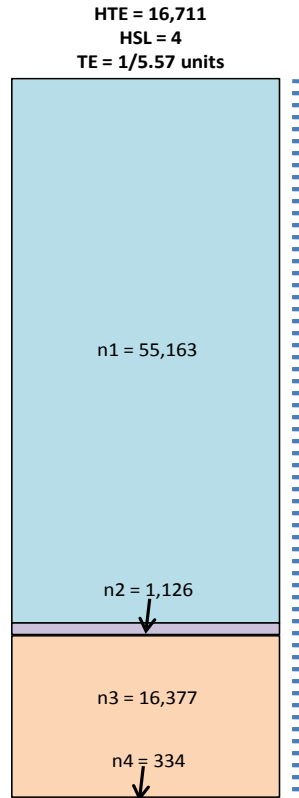
Select a representative sample of 10,000 out of 83,000 cases resulting in the following approximately strata breakdowns:

$n1_{PM} = 7557$
 $n2_{PM} = 154$
 $n3_{PM} = 2243$
 $n4_{PM} = 46$

**Incentive Timing Study (4 Groups):**

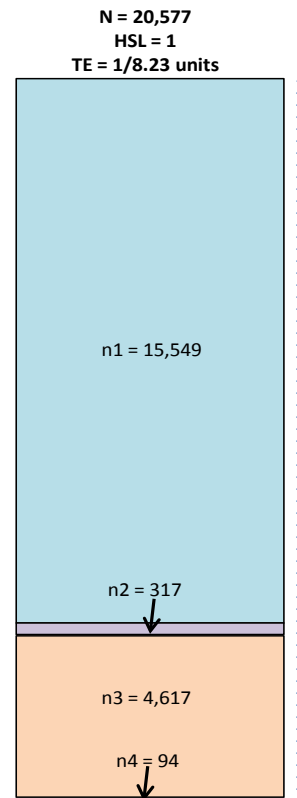
Select a representative sample which will identify a subset of 12000 (3000 clusters of 4) of the 16,711 available hard to enumerate (HTE) cases to send incentives:

$n1_{IT} = 39614$
 $n2_{IT} = 809$
 $n3_{IT} = 11760$
 $n4_{IT} = 240$
N = 73,000
HTE = 16,711
HSL = 4
TE = 1/5.57 units

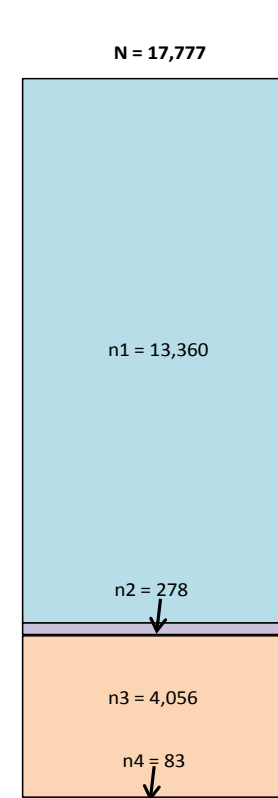
**Mode Switching Study:**

Select a representative sample of 2,500 out of 20577 cases resulting in the following approximately strata breakdowns:

$n1_{MS} = 1889$
 $n2_{MS} = 39$
 $n3_{MS} = 561$
 $n4_{MS} = 11$
(50 cases ineligible w/ no phone)

**Control Group for ALL Experiments:**

Full Leftover Sample



Power Analysis, Sample Sizes, and Hypotheses for the 2013 NSCG Methodological Studies

To aid in the determination of sample sizes for the four methodological studies proposed for the 2013 NSCG, Appendix I provides details of a power analysis associated with the 2013 NSCG methodological studies.

Using the findings from the power analysis, we are proposing the following sample sizes for treatment groups in the four methodological studies.

Incentive Timing Study:

- Incentive offered at week 1 of data collection (n = 3,000)
- Incentive offered at week 7 (n = 3,000)
- Incentive offered at week 12 (n = 3,000)
- Incentive offered at week 25 (n = 3,000)

Mode Switching Study:

- Mode switching study treatment group (n = 2,500)

Priority Mailing Study:

- First class mailing treatment group (n = 5,000)
- First class mailing that provides impression of importance (n = 5,000)

Incentive Conditioning Study:

- No incentive (n = 4,667)
- Incentive offered at the beginning of data collection (n = 4,667)
- Incentive offered at a late stage of data collection (n = 4,667)

For the three studies being conducted as part of the 2013 NSCG new cohort data collection (incentive timing study, mode switching study, and priority mailing study), after cases have been allocated to the treatment groups for these three studies, the balance of the new cohort cases will be allocated to the control group for all the studies as documented above in the section on Designing the Sample Selection for the 2013 NSCG Methodological Studies.

These sample sizes will allow the evaluation of the following hypotheses for the methodological studies:

Incentive Timing Study Hypothesis – The use of a cheaper, but equally effective web first data collection methodology will enable the later use of incentives for hard to enumerate cases to be an optimal data collection approach in terms of response, bias, and cost.

Mode Switching Study Hypothesis – The incorporation of mode switching techniques will enable the Census Bureau and NSF to create a more efficient data collection process that reduces the cost of data collection.

Priority Mailing Study Hypothesis – When used as part of the web first data collection methodology, the increased cost of the priority mailing reminder will not result in a practical increase in response or sample representation.

Incentive Conditioning Study Hypothesis – The use of a cheaper, but equally effective web first data collection methodology will enable the later use of incentives for previously incentivized cases to be an optimal data collection approach in terms of response, bias, and cost.

Post-Study Evaluations

As part of the long-term planning for the NSCG, the 2013 NSCG survey cycle will include post-study evaluations aimed at improving the data collection techniques used within the survey. These post-study evaluations are not part of a controlled experiment, but instead provide insight on some unanswered questions. Some of the planned post-study evaluations for the 2013 NSCG are as follows:

- Revaluation of the hard to enumerate definition – Using results from both the 2010 and 2013 survey cycles, we will reexamine the hard to enumerate definition. If another definition is deemed more appropriate, we will examine the impact this other definition would have had if it had been used as part of the 2013 NSCG data collection effort. The examination will pay special attention to the potential bias reduction associated with the cases identified as hard to enumerate. As noted earlier, the metric used to identify the hard to enumerate cases in the 2013 NSCG incentive timing study is the weighted response influence. This metric uses each case's base weight to estimate its potential for nonresponse bias. Examining the effectiveness of the sample weight in estimating the potential for nonresponse bias within this metric will be part of this analysis.
- Understanding sample representation throughout the data collection effort – As part of the daily processing (editing, imputation, weighting) being incorporated in the 2013 NSCG data collection effort, we will have daily information on the responding sample and the resulting survey estimates. Post-data collection, we plan to evaluate this information to better understand the relationship between the data collection procedures and the impact on both the responding sample and resulting survey estimates. By examining the responding sample on a daily basis, we will be able to identify the effect different data collection techniques have on the sample representation of the responding sample. A better representative sample (i.e., a sample that is demographically aligned with the target population) reduces the potential for bias. Thus, the goal of this analysis

is to identify data collection techniques that reduce the potential for nonresponse bias. Results from this evaluation may influence data collection decisions for future survey cycles.

- Determination of data collection stopping rules – Using the results from the daily processing (editing, imputation, weighting) of collected data, we will be able to investigate the possibility of using stopping rules in future survey cycles to identify when certain sample cases should no longer be targeted because reliability requirements have been met for key survey estimates. If such stopping rules were implemented in future survey cycles, it could result in a reduction in data collection costs and sooner dissemination of survey results. In addition, by discontinuing data collection for certain sample cases in response to having met reliability requirements, we are freeing resources to continue interviewing cases where reliability requirements have not been met. If we are able to refocus our resources in this manner, we will have shifting our data collection effort to the cases would best meet our estimation needs (rather than the cases that just increase our response rates). As a result, we would be reducing the potential for nonresponse bias in our estimates. Therefore, the goal of this post-study analysis is to develop a data collection methodology that uses adaptive design techniques to improve our survey estimates.
- Constant effort to meet our data collection goals (cost, bias, response, timing) – Implicit in the post-study analysis is our constant effort to meet the following data collection goals:
 - Lower overall data collection costs
 - Decreased potential for nonresponse bias in the NSCG survey estimates
 - Higher response rates
 - Increased efficiency in the use of incentives as part of a data collection methodology

In the post-study analysis examples listed above, we provide specific steps that will be taken in an effort to meet these data collection goals. This post-study analysis will ensure that any data collection methodology decisions that are made have considered all these goals. In the past, a main focus for the NSCG has been higher response rates. In this survey cycle and in future survey cycles, we will give equal, if not additional, emphasis to the other goals especially reducing the potential for nonresponse bias in our survey estimates.

Through the advances we are incorporating this survey cycle for the NSCG data collection (including daily processing and adaptive design techniques) and our increased emphasis on bias, cost, and timing through our methodological studies and post-study analysis, we are taking steps that will help the NSCG become a survey that is flexible, more responsive, and produces higher quality data.

5. CONTACTS FOR STATISTICAL ASPECTS OF DATA COLLECTION

Chief consultant on statistical aspects of data collection is Benjamin Reist (301) 763-6021, Demographic Statistical Methods Division, Census Bureau. The Demographic Statistical Methods Division will manage all sample selection operations at the Census Bureau. At NSF the contacts for statistical aspects of data collection are Stephen Cohen, NCSES Chief Statistician (703) 292-7769, and John Finamore, NSCG Project Manager (703) 292-2258.