National Science Foundation
WHERE DISCOVERIES BEGIN
Division of
SCIENCE RESOURCES STATISTICS

Design Options for SESTAT for the Current Decade: Statistical Issues

Working Paper | SRS 07-201 | June 2007

by Ronald S. Fecso (formerly Division of Science Resources Statistics, National Science Foundation), Robert Baskin, Adam Chu, Cindy Gray, Graham Kalton, and Richard Phelps (Westat, Inc.)

Introduction

The Scientists and Engineers Statistical Data System (SESTAT) is a database of the employment, education, and demographic characteristics of a sample of scientists and engineers in the United States. SESTAT is maintained by the National Science Foundation to provide data for policy analysis and general research. In the 1990s the database was compiled biennially in 1993, 1995, 1997, and 1999. At each round, the SESTAT integrated database was constructed from data collected in three separate surveys: the National Survey of College Graduates (NSCG), the National Survey of Recent College Graduates (NSRCG), and the Survey of Doctorate Recipients (SDR).

Under the SESTAT definition of scientists and engineers, in 1995 approximately 12 million scientists and engineers were in the United States, compared with a total resident population of 197 million who were age 18 years and older. Sampling this rare population presents a challenge. In the 1990s the starting point for the SESTAT database was to sample college graduates identified in the 1990 census. This sample of college graduates was surveyed in the NSCG in 1993; individuals identified as scientists and engineers became part of a panel that was surveyed for subsequent rounds of SESTAT. To represent the flow of new scientists and engineers after 1990, a survey of bachelor's and master's degree recipients who obtained their degrees in the previous 2 academic years—the NSRCG—was conducted at the same time as the NSCG.[1] Subsamples of the NSRCG respondents from each round were added to the panel for subsequent rounds of the SESTAT surveys. For each round, the panel of the census-based sample and the NSRCG samples from previous rounds is known as the NSCG. The third survey in the SESTAT system is the SDR, which is a panel survey that represents individuals who earned doctorates in the United States.

As the NSCG surveys during the 1990s progressed, the NSCG nonresponse and undercoverage rates increased. To address this problem, the sampling plan called for a "refreshing" when data from the next census became available. The Division of Science Resources Statistics (SRS) investigated various potential redesign alternatives to the 1990s plan to guide in creating the design for the current decade.

This report describes and evaluates the designs proposed.[2]

With the 2000 census, different sample design options became available for the NSCG in the following decade. An earlier study of potential sampling frames found that the only functional sampling frames for the SESTAT surveys in the 2000s are the census and a continuation of the existing panels. The most basic approach is to continue with the existing panels and supplements for new entrants to the population of scientists and engineers. An alternative approach is to start afresh with a sample for the NSCG from the 2000 census, repeating the design used in the 1990s. This leads to a consideration of combinations and refinements of these approaches. This report addresses some of the statistical issues related to the four design options considered by the National Science Foundation (NSF) for the SESTAT surveys for the current decade.

The report is organized as follows. "Overview of SESTAT Design of the 1990s" provides an overview of the 1990s SESTAT system and identifies some of the limitations of that system, particularly with respect to issues of undercoverage and nonresponse. "SESTAT Redesign Options" describes the four alternative designs considered and assesses their advantages and limitations. "Implications for Variance Estimation" discusses issues related to variance estimation and "Use of Web-based Data Collection" discusses the use of Web-based data collected for the SESTAT surveys. A summary of the main findings is contained in "Summary." The appendix to the report contains a memo discussing response issues for the SESTAT redesign options. This memo was revised to include comments from NSF.

Footnotes

[1] The first NSRCG of the decade included some scientists and engineers who graduated immediately after the census as well as those who graduated in the 2 following academic years.

[2] Although SRS's small technical staff often propose technical changes to surveys (in this case, SRS Chief Statistician Ronald S. Fecso proposed the design options forming the core of this report), it is SRS's practice to obtain the advice of other highly regarded methodological professionals before implementing major changes. Such reviews, generally done through contracts (as is the case with this report) help ensure that methodological proposals represent best practices as viewed by a range of methodologists. This report documents design change proposals and the review of such material that took place in the time period during which decisions were being made for the NSCG design for the 2000s. As such, data for 1999 were not available and are not included in this report.

Overview of 1990s SESTAT Design

This section describes the SESTAT suite of surveys of the 1990s. It starts with a description of the target population and the three survey components and then reviews the coverage gaps and response rates that have been achieved. The section closes with a brief review of the weighting and variance estimation procedures used in SESTAT.

Target Population

The target population for the SESTAT database of the 1990s is residents of the United States who have at least a bachelor's degree and who, as of a specified reference date, were noninstitutionalized, age 75 or younger, and either had a degree in science and engineering or were working as a scientist or an engineer. S&E is defined as the broad categories of computer and mathematical sciences, life and related sciences, physical and related sciences, social and related sciences, and engineering (for a complete definition, see NSF/SRS 1999). The SESTAT definition includes the following two groups:

- individuals with S&E degrees, regardless of their occupation
- college graduates (i.e., individuals with at least a bachelor's degree) who do not have S&E degrees but who are working in an S&E occupation

The definition of a scientist or engineer requires a college degree (i.e., a bachelor's degree or higher). Those working in S&E occupations who do not have college degrees (e.g., individuals with associate's degrees in any field) are not covered in the SESTAT surveys and database.

Some analysts may prefer to restrict their definition to a subset of the target population, such as individuals with S&E degrees. As discussed below, coverage gaps will be less of a concern when the definition is restricted to those with S&E degrees. The occupational mobility of college graduates without S&E degrees into and out of S&E occupations causes significant coverage problems.

Components of SESTAT

The SESTAT database includes three components, each designed to represent different parts of the target population. The NSCG represents U.S. scientists and engineers existing at the previous round of data collection, the NSRCG represents new S&E bachelor's and master's graduates from U.S. institutions since the last round of data collection, and the SDR represents the population of those who earned U.S. doctorates. The three components are described briefly in the following sections. Detailed descriptions can be found on the NSF/SRS website (http://www.nsf.gov/statistics/survey.cfm).

National Survey of College Graduates

The largest component of the SESTAT database in the 1990s was the NSCG, which was conducted in 1993. The mode of collection for the NSCG was mostly mail with computer-assisted telephone interviewing (CATI) and personal interviewing followup. The sample for the 1993 NSCG was drawn from the 1990 census long form records, which contained information on degree level, field of occupation, country of origin, and date of entry into the United States, along with demographic information. The sample from the long form was restricted to individuals who had at least a bachelor's degree. However, because the long form did not include a question about degree field, the 1993 NSCG sample included individuals with S&E degrees and those with only non-S&E degrees. The sample thus consisted of three components:

those with S&E degrees, whether or not they were working in an S&E occupation; those with degrees only in non-S&E fields who were working in S&E occupations; and those with degrees only in non-S&E fields who were not working in S&E occupations. The third group is not part of the S&E target population. In the 1993 NSCG, there were 148,932 respondents plus an additional 19,244 cases deemed to be ineligible (e.g., deceased, over 75 years old, or no longer in the United States) out of a total sample of 214,643. Of the 148,932 responses, 74,693 (50%) fit the definition of the S&E target population.

The NSCG was also conducted in 1995, 1997, and 1999 (the NSCG was not fielded in 2001). The 1995 NSCG was administered to all 1993 NSCG respondents who were classified in 1993 as being in the target population and to a sample of individuals from the 1993 NSRCG sample (as described below) to represent recent graduates. Individuals without S&E degrees who were working in non-S&E occupations were eliminated from the 1995 NSCG sample along with individuals who aged out of the sample when they became older than 75 and those who could have been selected for the current NSRCG or SDR components of SESTAT—the latter called overlap cases. The 1997 and 1999 NSCG samples were similar to the 1995 sample except they consisted of only a sample of the respondents from the previous NSCG cycle.[3]

National Survey of Recent College Graduates

The NSRCG is designed to identify recent college graduates who are not represented in the previous round of the SESTAT surveys. The NSRCG was conducted in 1993, 1995, 1997, 1999, and 2001. The sample for each cycle covered individuals who graduated in the previous 2 academic years with a bachelor's or master's degree and a major in an S&E field. Subsamples of NSRCG respondents from a given round were carried forward for inclusion in the following NSCG survey. The mode of collection for the NSRCG was mostly CATI with mail followup.

The NSRCG employed a two-stage sample design for the 1993 through 2001 surveys. The first stage was a stratified sample of colleges and universities in the United States that awarded bachelor's and/or master's degrees in science or engineering. The institutions were selected with probability proportional to the number of graduates in the S&E reporting categories from a sampling frame constructed from the Integrated Postsecondary Education Data System (IPEDS), which is maintained by the National Center for Education Statistics. The 1993 NSRCG institution sample design was based on the institution sample used during the previous decade. The sample design was revised for the 1995 cycle; it used a new institution sample selected using the 1991–92 IPEDS Completions file. The 1995 first-stage sample was used again in 1997. During the 1999 survey cycle, this sample was evaluated and supplemented with additional institutions from the 1994–95 IPEDS. Similarly, in the 2001 survey cycle, the sample was supplemented with institutions from the 1996–97 IPEDS. The second sampling stage involved the selection of S&E bachelor's and master's graduates from lists provided by the sampled institutions for each cycle of the survey.

Survey of Doctorate Recipients

The SDR represents individuals with earned doctorates from a U.S. degree-granting institution. Doctorate recipients are sampled separately in the SESTAT surveys because of a desire to increase the sample of earned doctorates and to maintain comprehensive information on this group. The primary source of information for the frame of doctorate recipients is the Survey of Earned Doctorates (SED), which is a census of newly granted research doctorates in the United States that has been collected each year since 1957. Before 1957, the National Academy of Sciences

maintained a register of highly qualified scientists and engineers assembled from a variety of sources, and such recipients are represented in the SDR panel until they age out of the sample. During the 1990s, the SDR sample included prior samples (with small maintenance reductions each cycle) and a stratified sample of recent doctorate recipients. The SDR was conducted in 1993, 1995, 1997, 1999, and 2001. The mode of collection for the SDR has mostly been mail with CATI followup.

Coverage Issues

The SESTAT target population includes residents of the United States who, as of the survey reference period, were noninstitutionalized, age 75 or younger, had at least a bachelor's degree, and either had a bachelor's degree or higher in an S&E field or were working as a scientist or an engineer. However, certain groups that are intended to be in the target population are either covered only partially or not covered at all in the SESTAT database.

One main group only partially covered is referred to as the "foreign degreed." Those not covered include individuals who were not residents of the United States as of 1 April 1990 (except those serving in the U.S. Armed Forces overseas) and who received a degree from a foreign degree-granting institution but not from a U.S. institution. Also not covered are those who were residents of the United States at the time of the 1990 decennial census and at that time had no degree but later received a degree from a foreign institution. Under the 1990s design, the foreign degreed are included in the SESTAT database only if they were included in the 1990 decennial census and already had at least a bachelor's degree. The undercoverage of the foreign-degreed group increased over the decade because of immigration and because of individuals who received foreign degrees after the census was conducted. However, some of these individuals then obtain U.S. S&E degrees and become part of the sampled population in the NSRCG or the SDR.

Another group that is only partially covered comprises individuals working in an S&E occupation who do not have an S&E degree. Individuals working in S&E occupations who initially graduated after 1 April 1990 and who have only non-S&E degrees are not covered. Also, among those who obtained a degree before 1 April 1990 and who have only non-S&E degrees, only those who were working in an S&E occupation in 1993 are covered. Again, for this group, the undercoverage increased over the course of the decade.

In general, the populations that are partially covered by the SESTAT database fall into one of the two groups described above. Additional details about the undercoverage can be found on the SESTAT website (http://sestat.nsf.gov/docs/techinfo.html#targetpop). These details are briefly summarized in the following sections.

Groups Not Covered in 1993

Within the coverage defined for the SESTAT integrated database, the following individuals with bachelor's and master's degrees were not included in the 1993 surveys.

- Those living abroad at the time of the decennial census (1 April 1990) who later came to live in the United States. However, if these individuals earned an S&E degree from a U.S. institution between the time of the census and 30 June 1992, they were covered. Members of the U.S. Armed Forces stationed abroad at the time of the 1990 census were covered by SESTAT.

- Those living in the United States at the time of the census who earned an S&E degree from a foreign institution after 1 April 1990. However, if they also held a U.S. degree at the time of the census, they were covered in SESTAT.
- Those who did not hold a degree at the time of the census and who obtained only non-S&E degrees between the time of the census and 1993 or received their first S&E degree after June 1992.

Doctorate-level individuals with S&E degrees who were not surveyed in 1993 were predominately U.S. residents who (1) received an S&E doctorate after June 1992 or (2) earned that degree at a foreign institution and

- had no degree in any field as of 1 April 1990 and had not received a bachelor's or master's S&E degree from a U.S. institution between 1 April 1990 and 30 June 1992 or
- resided outside the United States on 1 April 1990 and were not part of the U.S. Armed Forces stationed abroad.

Groups Not Covered in 1995

The following individuals with bachelor's and master's degrees were not surveyed and therefore are not represented in the 1995 SESTAT integrated database.

- U.S. residents whose S&E bachelor's and/or master's degrees were either received before April 1990 or were earned at a foreign institution, and who resided outside the United States on 1 April 1990 but were not part of the U.S. Armed Forces stationed abroad.
- U.S. residents who had no degree of any kind in any field as of 1 April 1990 and who were awarded an S&E degree after June 1994 by a U.S. institution or after April 1990 by a foreign institution.
- Individuals not previously eligible for NSCG sampling who obtained only non-S&E degrees between the time of the census and 1995.

Doctorate-level individuals with S&E degrees who were not surveyed in 1995 were predominately U.S. residents who (1) received an S&E doctorate either after June 1994 or (2) earned that degree at a foreign institution and

- had no degree of any kind in any field as of 1 April 1990 and did not receive a bachelor's or master's S&E degree from a U.S. institution between 1 April 1990 and 30 June 1994 or
- resided outside the United States on 1 April 1990 but were not part of the U.S. Armed Forces stationed abroad.

Groups Not Covered in 1997

The following individuals with bachelor's and master's degrees were not surveyed and therefore are not included in the 1997 SESTAT integrated database.

- U.S. residents whose S&E bachelor's and/or master's degrees were either received before April 1990 or were earned at a foreign institution, and who resided outside the United States on 1 April 1990 but were not part of the U.S. Armed Forces stationed abroad.
- U.S. residents who had no degree of any kind in any field as of 1 April 1990

and who were awarded an S&E degree after June 1996 by a U.S. institution or after April 1990 by a foreign institution.

- Individuals not previously eligible for NSCG sampling who obtained only non-S&E degrees between the time of the census and 1997.

Doctorate-level individuals with S&E degrees who were not surveyed in 1997 were predominately U.S. residents who (1) received an S&E doctorate either after June 1996 or (2) earned that degree at a foreign institution and

- had no degree of any kind in any field as of 1 April 1990 and did not receive a bachelor's or master's S&E degree from a U.S. institution between 1 April 1990 and 30 June 1996 or
- resided outside the United States on 1 April 1990 but were not part of the U.S. Armed Forces stationed abroad.

Response Rates

It is important to consider the impact of attrition on sample size when designing the surveys that populate the SESTAT database. Even when there is a relatively high response rate in any given survey year, the cumulative response rates of a longitudinal survey will deteriorate over time. Table 1 summarizes unweighted published response rates for the three survey components. The response rates in the table are not directly comparable. The NSCG response rate for 1993 is the response rate for the initial (full coverage) sample as selected from the census long form records, whereas the response rates for the later years of NSCG are "conditional" response rates pertaining to the sample of respondents from previous cycles (including supplemental cases from the NSRCG). The response rates for the NSRCG, on the other hand, are "unconditional" response rates pertaining to the cross-sectional samples that were selected for the particular cycles (years). The response rates for the SDR are also unconditional but, as noted in the table, a subsample of prior nonrespondents was selected for followup in the 1995 SDR (unlike the procedures for other years, in which all nonrespondents were followed up).

TABLE 1. Selected response rates for SESTAT components
(Percent)

	Conditional	Unconditional	
Year	NSCG	NSRCG	SDR
1993	78[a]	84	87
1995	95	83	77 (85)[b]
1997	94	81	84

NSCG = National Survey of College Graduates.
NSRCG = National Survey of Recent College Graduates.
SDR = Survey of Doctorate Recipients.

[a] Overall (unconditional) response rate for initial (1993) NSCG sample as reported in sample design section of website below.

[b] In 1995, a subsample of mail nonrespondents was selected for CATI followup; 77% is unweighted response rate using total sample (including those subsampled out) as base. Weighted response rate of 85% is weighted response rate that takes into account subsampling.

SOURCE: National Science Foundation, Division of Science Resources Statistics, Scientists and Engineers Statistical Data System (SESTAT), Design and Methodology report, http://sestat.nsf.gov/docs/techinfo.html.

The three components show somewhat different patterns of response. The initial response rate for the NSCG was the response rate for the sample selected from the 1990 census long form. However, in succeeding years the NSCG response rate shown

in the table is the rate of response among those who responded in the previous cycle. In the three NSCG cycles shown in the table, assume the response rates are multiplicative. Then with an initial response rate of 80%, the overall response rate in the second year would be 76% (80% times 95%). After another cycle, the overall response rate would decrease to 71% in 1997 and, carrying the computations forward, would continue to decrease in subsequent rounds.

The NSRCG response rates in Table 1 , which are based only on the new sample of recent graduates for a given cycle, declined slightly from 1993 to 1997 but were 82% or higher in all years surveyed. Because this sample "feeds" into the NSCG sample in the following round, the response rate for the NSRCG has a direct bearing on the overall response rate for the NSCG.

The SDR response rates in Table 1 are the overall (unweighted) response rates for the given cycle. In 1993, the unweighted SDR response rate was 87%. The unweighted SDR response rate of 77% in 1995 was due to subsampling of nonrespondents as the end of the survey approached. The 1995 SDR subsampling took about 60% of the mail nonrespondents prior to CATI followup. Thus, the unweighted response rate for 1995 was calculated using the original sample size as a base, even though 40% of the mail nonrespondents were not included in the CATI data collection. The weighted response rate that reflects the subsampling of nonrespondents was 85%. This is consistent with the 84% response rate achieved in the 1997 SDR. Thus, it may be reasonable to speculate that SDR response rates will remain at about 80% in subsequent rounds. Such subsampling was conducted in the 1995 cycle only.

Although response rates are generally high for both the NSRCG and the SDR, the NSCG is the largest component of SESTAT and thus has the greatest impact on overall response rates. Despite the high conditional response rates achieved for the NSCG, over time the cumulative effects of attrition will inevitably lead to very low overall response rates.

Weighting in SESTAT

The initial purpose of weighting is to compensate for differential probabilities of selection. This compensation is achieved through the "base weight," which is defined as the reciprocal of the probability of selecting a person for the study. The final weights used for the survey analysis may include one or more adjustments. For example, nonresponse-weighting adjustments are often used to inflate the base weights to compensate for unit nonresponse. Poststratification adjustments are used to make the weighted sample counts conform to known population totals, and thus provide a way of adjusting for possible undercoverage of the target population.

Although some aspects of weighting have varied from year to year, the main features of the weighting procedures used in SESTAT can be summarized as follows.

- In the 1993 NSCG, the base weights were adjusted to compensate for nonresponse and a poststratification adjustment was used to bring the weighted distribution for the sample into agreement with the 1990 census distribution for certain characteristics.
- In the 1993 NSRCG, two separate adjustments were made. The first one was a poststratification adjustment that was applied to the (first-stage) institution weight. For this adjustment, a ratio was calculated using IPEDS data in each of the 12 ratio-adjustment strata based on degree level and major field. For each ratio, the numerator was the sum of the number of degrees awarded over all

institutions in the universe (i.e., in IPEDS) and the denominator was the weighted sum of degrees awarded in the sampled responding institutions as reported in IPEDS, using the institution nonresponse-adjusted weight. The resulting (poststratified) institution weight was then used to develop an initial person-level weight, which was subsequently adjusted for survey nonresponse within designated weighting classes.

- In the 1993 SDR, the base weights were adjusted for nonresponse within specified weighting classes. There was no additional poststratification of the nonresponse-adjusted weights.
- The final step in the weighting process was to adjust for multiple chances of selection both within a given survey and across the three survey components. For example, a person who received a bachelor's degree in June 1995 and a master's degree in June 1996 could be selected twice in the 1997 NSRCG. As a final within-survey adjustment, the weights of all individuals in the NSRCG who could have been selected twice were halved. For the overall SESTAT weights, to address multiple selection across components, the sampled individuals were linked to only one survey and made ineligible if selected for the other component.
- In general, similar weighting procedures, except for postratification adjustments, were employed in subsequent survey years.

Variance Estimation in SESTAT

Replication methods have been used to produce estimates of sampling errors for all three components of SESTAT. Because the variance estimates were developed by different survey organizations, different methods have been used. For example, for the NSCG, balance repeated replication (BRR) was used for variance estimation. For the NSRCG, a jackknife method for a paired selection sample design, referred to as "JK2," was used (see Westat 2000); for the 1993 NSRCG 50 replicate weights were produced for variance estimation, whereas for the 1995–99 NSRCG 86 replicate weights were produced. Additional details of the methods used in the NSRCG are given in Westat (1999). For the 1993 SDR, BRR was used with 16 replicate weights for variance estimation. For later SDR samples, the number of replicate weights was increased to 48. Details of the methods used in the SDR for variance estimation are given in U.S. Census Bureau (2001).

The variance of an estimated total obtained by pooling the various SESTAT samples can be calculated by simply adding the variance of the individual components. As long as the individual variance estimates are approximately unbiased (and the samples can be assumed to be roughly independent), the resulting total variance is also approximately unbiased no matter what methods are used to obtain the individual variances. Because there is a large number of questionnaire items in the components of SESTAT, generalized variance curves are also used to provide approximate variances for broad classes of statistics based on observed relationships between the weighted estimate and the calculated sampling error.

Footnote

[3] The U.S. Census Bureau was required to perform the data collection for the NSCG sample derived from the 1990 census. However, for the 1997 and 1999 SESTAT cycles, the NSCG panel subsamples originally selected as part of the NSRCG were not collected by Census but as part of the NSRCG panel survey. The

data for these cases were integrated with the NSCG and SDR to form the full SESTAT integrated database.

SESTAT Redesign Options

NSF proposed four options for the redesign of the SESTAT suite of surveys. Each option is outlined in this chapter. One of the options was essentially a repeat of the design implemented during the 1990s. For all options, the target population is theoretically the same—residents of the United States with at least a bachelor's degree and who, as of the survey reference period, were noninstitutionalized, age 75 years or younger, and either have an S&E degree or are working in an S&E occupation. Each option has some coverage gaps; these gaps are different for each option.

Figure 1 is the options chart, which graphically shows information about each option for the 2003 data collection. The chart has a separate row for each population component included in the target population. The last three rows are for population components that are currently undercovered or not covered in the SESTAT surveys but may be part of the target population. The first three columns indicate the survey that currently covers each population component, the population component involved, and the frame from which the survey sample is currently selected. For example, the first row shows that the doctorate population is currently surveyed by the SDR, which uses the SED as the sampling frame. The population components currently included in the NSRCG are described by the academic year of degree receipt. For example, "1991–92 Bac/Master" includes degrees received from July 1990 through June 1992 (2 academic years). The next set of columns shows how each population component would be sampled and surveyed under each of the four options. Thus, in option 1, the doctorates are covered by the SED/SDR, the bachelor's/master's from pre-1990 through March 2000 are included in a 2000 postcensal survey, the bachelor's/master's from April through June 2000 are a panel selected from the 2001 NSRCG survey, and the bachelor's/master's from July 2000 through June 2002 are included in the 2003 NSRCG.

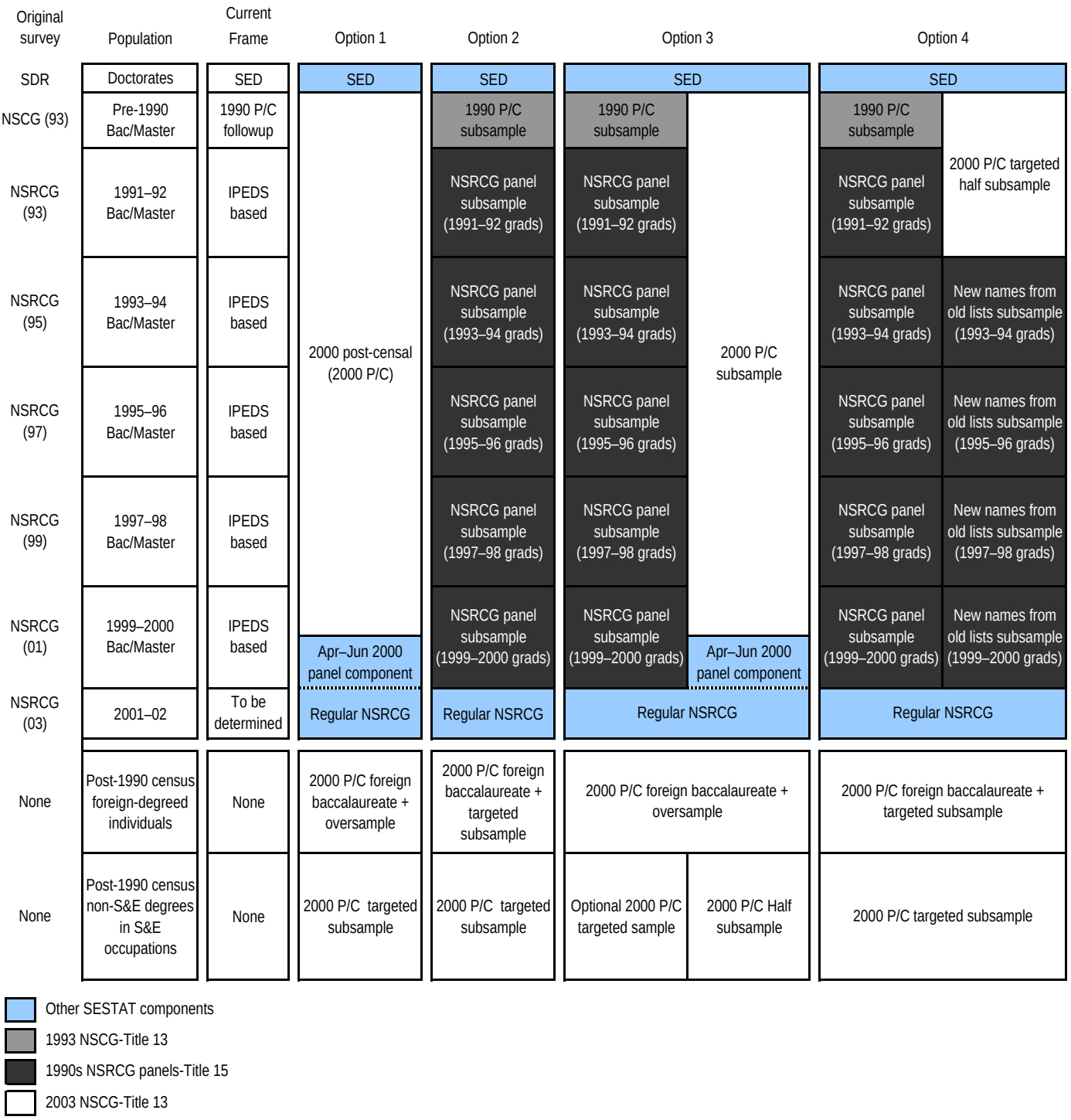
The four options have some characteristics in common. First, the SDR would continue its current system of sampling and data collection for the U.S. doctorate population under all of the options because a more efficient frame has not been identified. Furthermore, the design of the SDR would not be affected by the choice among the four options. Second, the NSRCG would not change. Third, under all four options the sample would be updated for new entrants into the S&E population every 2 years using new S&E graduates from the NSRCG. Fourth, the SESTAT survey would be conducted about every 2 years from 2003 to 2009 under each of these options.

Figure 1 also discusses the differences between the options with regard to coverage gaps, sample attrition, and screening requirements.

Option 1: A Replication of Design of the 1990s

Option 1 is a replication of the 1990s design. Under this option, the Census Bureau would conduct a postcensal survey in 2003 based on the 2000 census. This survey would include college graduates who, in April 2000, had received at least a bachelor's degree and were age 72 years or younger (and thus would be age 75 years or younger in 2003), noninstitutionalized, and living in the United States or serving in the Armed Forces overseas. In 2003, this sample would be contacted and interviewed for the NSCG. On the basis of the interview, those individuals in the 2003 sample who (a) were college graduates with S&E degrees or (b) had a college degree and were working in S&E occupations would be screened into the 2005 NSCG sample. Foreign-degreed individuals who were in the 2000 census as well as those with

FIGURE 1. 2003 SESTAT data collection



non-S&E degrees but who were working in S&E occupations in 2003 would be included in the followup NSCG sample.

Under this option two major groups in the target population would not be covered or would be poorly covered in 2003: (a) individuals eligible for the SESTAT integrated database who lived abroad as of the 2000 decennial census who later came to live in the United States and who did not earn a bachelor's or higher S&E degree from a U.S. institution after April 2000 and (b) individuals with only non-S&E degrees obtained after April 2000 who held S&E occupations in the survey reference period. In addition, individuals with only non-S&E degrees (who had obtained at least one degree before April 2000) who did not hold S&E occupations in 2003 but held such occupations in a later survey reference period would not be covered in the followup NSCG surveys after 2003.

As in 1993, this option for the NSCG is essentially a large screening effort. As an indication of how much screening might be needed, in the 1993 NSCG 214,643 individuals were selected from the 4,728,000 people who were reported on the 1990 census long form as meeting the degree and eligible age requirements. There were 148,932 eligible respondents for the 1993 NSCG from this sample and 19,224 others who were classified as ineligible for reasons such as being deceased, over age 75, or no longer living in the United States. The NSCG eligible respondents yielded a total of 74,693 respondents who were eligible for inclusion in the SESTAT database. About half of those who were interviewed in the 1993 NSCG were not eligible for inclusion in SESTAT because they did not have an S&E degree and were not working in an S&E occupation. This would imply that if the same response and eligibility rates occur in 2003, less than 35% of those individuals selected from the 2000 census would respond to the 2003 NSCG and be eligible for inclusion in SESTAT. This corresponds to a screening rate of almost 3 to 1 (i.e., the ratio of the initial sample size to the number of SESTAT-eligible respondents), which is the maximum screening ratio among the four options considered.

Under option 1, and indeed for each option, the coverage of foreign-degreed individuals and those with non-S&E degrees who have S&E occupations in 2003 will obviously be much greater than in the 1999 SESTAT database because there were no attempts during the 1990s to update the original NSCG sample with these subgroups. This can cause a "discontinuity" when comparing results from the new design with those from the previous design. However, if the subgroups of foreign-degreed individuals and those with non-S&E degrees are removed from the analysis, then the resulting population of inference will be similar to that represented by the 1999 SESTAT sample and time trends will be maintained—although differences will be subject to larger sampling errors than if the samples were overlapping. It should also be noted that time-trend comparisons may be affected by differential nonsampling errors, such as nonresponse and panel conditioning, in the surveys being compared. In addition, small domain estimates (e.g., demographic groups with small numbers in particular degree fields) may be unstable.

Because of the declining unconditional response rates that can be expected in the NSCG at each successive round (see the section "Response Rates"), the sample under option 1 (which is a "new" sample) will begin with an estimated 15% smaller nonresponse than the existing SESTAT sample, provided the 1993 initial response rate is maintained in the 2003 postcensal survey. Also, to the extent that any conditioning biases are introduced into the panel over time, the new sample will be subject to smaller panel effects than the existing sample.

A methodological approach that NSF also considered for option 1 is the use of decentralized CATI collection for followup of mail nonrespondents in the postcensal survey. This approach could be more cost effective and may result in higher response rates than centralized CATI collection for this survey. (The decentralized CATI option, which was named option 5 by NSF, is not considered separately in this report because any of the four options discussed here could have a decentralized CATI version.)

Option 2: A Continuation of Current Panels

Under option 2, the current sample based on the 1990 census would continue with some attempt to fill gaps in coverage, where possible and cost effective. Targeted samples would be screened from the 2000 census to update the sample of foreign-degreed individuals and college graduates with only non-S&E degrees who work in S&E occupations. To decrease the nonresponse rate, NSF could also go back to the original samples from 1993 as well as the later panels and try to trace the nonrespondents.

Individuals with at least a bachelor's degree who were born and educated abroad can be sampled directly from the 2000 census by stratifying according to country of origin and year of entry into the United States, but the foreign degrees would need to be screened to identify those in S&E fields. U.S. citizens educated abroad because the 1990 census would not be covered. Individuals with only non-S&E degrees who are working in S&E occupations cannot be identified in advance of sampling. Thus, if these individuals are to be included in the sample at the same rate as in option 1, the initial sample size from the census would have to be the same as in option 1. In this case, option 2 would offer no advantage over option 1 in terms of reducing screening costs.

On the other hand, if it is possible to identify occupational groups in which large numbers of individuals with only non-S&E degrees work (e.g., the computer science field), then it may be possible to either target (restrict) the sample of individuals with non-S&E degrees to selected occupations (however, this would result in some undercoverage in the occupations not sampled) or, alternatively, to sample the occupations differentially to reduce the total screening effort. NSF could conduct a study of S&E occupational fields to identify those occupations in which individuals with non-S&E degrees represent a sizable proportion of workers, and this information could then be used for sample design purposes to determine how to sample the various occupational groups.

Under this option (where the 2000 census would only be used to sample individuals born and educated abroad and those targeted as having non-S&E degrees), the coverage gaps would grow through the decade similarly to option 1. Although the existing SESTAT surveys do not cover foreign-degreed individuals and those with only non-S&E degrees after April 1990, "refreshing" the NSCG sample with the foreign degreed and those with only non-S&E degrees from the 2000 census would produce coverage similar to option 1.

A variant of the above approach would be to use the 2000 census to sample only the foreign-degreed (who can more readily be identified from the census long form than can those working in S&E occupations without S&E degrees). Individuals with only non-S&E degrees in S&E occupations would not be covered except, if desired, to the extent that they are represented in the 1990 census sample. This variant would allow inferences to be made to the restricted population of individuals with S&E degrees. However, even with this restricted definition there would still be some

undercoverage, namely individuals born in the United States who received only foreign degrees between April 1990 and April 2000.

The level of effort for screening under option 2 depends on the desired sample size for individuals with non-S&E degrees screened from a sample of college graduates identified in the 2000 census. If the intent is to maintain the same expected sample size for this subgroup as in option 1, then the level of screening would be roughly the same as in option 1. On the other hand, if the SESTAT population was restricted to individuals with S&E degrees (so those without S&E degrees would be excluded from sampling), then screening the census population would essentially be eliminated. Between the two extremes are intermediate positions in which the population of individuals with only non-S&E degrees who are working in S&E are (a) restricted to certain occupations or (b) sampled at varying rates depending on occupational group. Position (a) would lead to undercoverage bias whereas position (b) would avoid undercoverage but lead to increased sampling errors. Either of these positions could plausibly lead to screening rates that are one-half to three-quarters the size of those for option 1. Further analysis would be required to determine the approximately optimum sampling rate under the various scenarios.

This option allows for only limited flexibility in terms of allocating sample size in the old sample. Furthermore, unless the tracing efforts are successful, the response rate that is a concern in the current design would not be improved. Weighting adjustments needed to compensate for nonresponse tend to increase the variation in the weights, which in turn tends to increase sampling errors.

The primary advantage of this option (apart from the cost savings associated with retaining previous respondents) is that it maintains the longitudinal aspects of the original SESTAT design and small domain estimates are more stable over time. Specifically, option 2 will permit analysis of change among individuals who were in the study for two or more rounds in different decades. The potential for longitudinal analysis would be lost under option 1 until subsequent data collection rounds are completed, and would be nonexistent for comparisons involving a time interval of more than 10 years (e.g., spanning different decades).

Option 3: A Split Design Combining Options 1 and 2

This dual-frame option combines features of option 1 and option 2. In this dual-frame design, a portion of the sample would be selected using the frame from option 1 and the remainder using the 1999 NSCG panel frame from option 2. A 50-50 split of the sample is roughly optimum for making comparisons between two options in most such designs; however, a 50-50 split is not necessary and in this case is not the optimum allocation for producing estimates from the combined sample.

Under this option, the sample drawn from the 2000 census would be a smaller sample of college graduates than if the whole sample was used for the sample under option 1. An advantage of using the 2000 census for part of the sample is that the subpopulations with coverage problems would be represented. These subpopulations consist of foreign-degreed individuals and those with non-S&E degrees who have moved into S&E occupations since the 1990 census. The remaining portion of the option 3 sample would be selected from the existing NSCG panel. The advantage of using the 1999 NSCG panel for part of the sample is that the screening for S&E cases has already been done, so each sample case yields an eligible S&E case for analysis, whereas samples from the 2000 census yield only about one eligible S&E case for every three selected.

It should be noted that the 1999 NSCG panel (option 2) has only about 46,000 cases, thus limiting the option 2 allocation. Therefore, if the total number of samples approaches 200,000, as it did in past redesigns, the 1999 panel samples would make up the minority of cases, even if all 46,000 were used. The remaining sample of over 150,000 cases would be drawn from the 2000 census. Assuming further that the one-in-three screening rate for finding S&E cases obtained in past postcensal surveys still holds, the 2000 census (option 1 portion) would yield about 50,000 in-scope cases. Thus, the overall S&E sample from the option 3 design would be about 96,000 cases (46,000 + 50,000) compared with 67,000 cases (one-third of 200,000) from the option 1 design. Subsampling of the old panel cases could be used to bring the "old" sample size to a smaller portion of the total sample size if desired.

The coverage gaps under option 3 would be the same as those under option 1. However, the sample sizes for the undercovered subgroups may not be the same. For example, because it is not possible to identify individuals with non-S&E degrees who are working in S&E occupations in the census frame, the screening required for the option 1 portion of the sample in the hypothetical example above will yield about three-quarters of the sample size for the subgroup of individuals in S&E without an S&E degree (assuming a roughly 50,000/150,000 sample split and no other targeting, as described below). However, if it is cost effective to target selected occupations related to workers of interest with a non-S&E degree (as described previously under option 2), the sample for the targeted occupations can be increased accordingly. On the other hand, because it is probable that foreign-degreed individuals can be identified in the census files in advance of sampling, it will be feasible to sample this subgroup at the full rate without unduly increasing screening levels.

Option 3 permits an assessment of possible "panel effects" in the existing NSCG sample. If the two samples produce comparable results, they can be combined with relatively little loss in efficiency and potentially a net increase in the size of the S&E sample. On the other hand, if the comparisons indicate that estimates from the existing panel are markedly different from those based on the new sample, then it may be presumed that there is differential bias. In this case, the new sample can be used to make cross-sectional estimates as in the past (although with slightly reduced precision due to the smaller sample). Analysis of the differences might also provide improved nonresponse adjustment methodology that would bring the estimates from both frames back into effective joint use.

Option 4: A Variant of Option 2 Supplemented by the NSRCG

Option 4 is a variation of option 2 supplemented by the NSRCG. The old panels would continue as in option 2. Assuming that it is feasible to select new samples from the old NSRCG panel lists, additional samples would be taken to supplement the old panels. There could also be an attempt to recontact old nonrespondents. The difficulty is that there is no viable frame for the 1993 postcensal sample or for the 1993 NSRCG sample (the 1993 NSRCG frame is no longer available).

As in option 2, the 2000 census would be used for the limited purpose of refreshing the sample with the "missing" subpopulation of the foreign degreed and, perhaps, those with non-S&E degrees who have moved into S&E occupations.

The coverage gaps would be the same as in option 2. Sampling from the old NSRCG frames would not cover any missing subgroup but would only be an attempt to increase sample size in response to attrition. The resampling will not eliminate any existing nonresponse bias in the old cohorts but will merely provide another independent sample for comparison. As with option 3, the nonresponse bias, if any,

can be studied.

This option would require extensive tracing to locate individuals selected from the old NSRCG lists and may not have any cost advantage. It adds very little as compared with option 2.

Other Issues Pertaining to Design Options

Other issues that affect decisions about the SESTAT redesign include the following:

- frequency of the surveys
- relationships to other surveys
- methods of sample selection
- maintenance of trends with earlier data

There appears to be very little difference between the four options on most of these issues. All four options would allow the same frequency of surveying. All four options have exactly the same relationships to other surveys (e.g., Current Population Survey and American Community Survey). The four options are based on similar data sources and there is nothing in any option that would give it an advantage in terms of sample selection methodology.

The maintenance of trends within a longitudinal sample is the one area where some difference exists. Because the NSCG is a fresh sample in option 1, there may be a discontinuity in the trends from the 1990s and there is no basis for estimating individual gross changes across the decades. The other three options all allow some overlap of sample for investigation of gross differences over time. Each of the four options continues to have gaps in coverage; the types of gaps are conceptually the same but the extent varies. The greatest difference is in the way in which the options attempt to address screening levels (and costs), attrition, and nonresponse bias.

Implications for Variance Estimation

An important feature of probability sampling methods is that they permit the calculation of the sampling errors associated with the survey estimates. None of the four design options has a significant advantage over the others in regard to variance estimation. As long as appropriate variance stratum and unit identifiers are available in the data and related survey control files, reasonably good variance estimators can be developed using any of a number of well-known techniques. Because the four options all involve the same basic design elements (i.e., use of unclustered census samples plus essentially independent samples of recent college graduates and doctorate recipients), the same general approach can be used (with minor modifications) for all four options. The important aspects to capture in the variance estimator are (a) the relevant features of the designs used to select the various samples and (b) the weighting and estimation procedures used to develop estimates from the integrated survey data. Although the mechanics of weighting and variance estimation would be more complicated under option 3 (because it involves "parallel" samples of slightly different design that must be weighted separately and then combined using composite estimation), the additional processing burden associated with this option should not be a major factor in choosing among the alternative designs.

In general, sampling errors may be estimated by using analytical variance formulas based on a Taylor Series approach or by using replication techniques such as jackknife repeated replication or BRR. The Taylor Series approach is straightforward for simple linear estimates such as the expansion estimate of a population total, but the variance formulas can be complex for nonlinear statistics. Replication methods (e.g., see McCarthy 1966 or Wolter 1985) provide a relatively simple way of calculating variances and have some advantages over other variance estimation methods. In particular, the impact of weighting adjustments can be reflected approximately in the variance estimates obtained by replication methods.

As indicated in the section "Variance Estimation in SESTAT," the SESTAT surveys of the 1990s employed a variety of replication techniques for variance estimation. The BRR method was used for the NSCG and the SDR, whereas a jackknife approach was used for the NSRCG. Although valid variance estimates are generated when different variance estimation methods are used for the different components of the SESTAT integrated database (in fact, it would be possible to use Taylor Series approximations for one component and replication for another), using the same technique for all components would provide analysts with a unified approach for variance estimation. Unfortunately, this may be impractical because it would require recalculating existing replicate weights for one or more components of the SESTAT database.

For example, suppose that it is desired to use jackknife replication for variance estimation. Under option 1, a completely new census sample would be selected for fielding in 2003. For this component of the SESTAT database, it would be straightforward to develop the required jackknife replicate weights. This is also true for the NSRCG component (where jackknife replication is currently being used). However, the SDR has used BRR since the beginning of the integrated database with the 1993 surveys. To develop the required set of jackknife replicate weights for the SDR, it would be necessary to first construct jackknife replicates for each existing "panel" (cohort) in the SDR. Once the jackknife replicates had been constructed, all of the weighting adjustments applied to the full sample would have to be repeated for each replicate (separately for each panel within the SDR). Although it is theoretically

possible to construct the jackknife replicate weights in this manner, the work involved would be difficult, time consuming, and prone to error. Thus, it may be preferable to simply continue with the current BRR approach for the SDR.

As a simplification, under the assumption that the total sample can be assigned to appropriate variance units and variance strata, it is possible to replicate the current (existing) full sample weights (without replicating the weighting adjustments that have occurred in the previous rounds). The resulting weights will not reflect all of the adjustments that have been made, but may nonetheless provide a reasonably good approximation of the variance. If this simplification does not seriously affect the integrity of the variance estimates, it may provide a practical solution to the problem of retroactively creating replicate weights for an existing sample.

Use of Web-based Data Collection

Web-based techniques offer a promising new mode of data collection for the SESTAT surveys. Both the general availability of Web access and the public's familiarity with it continue to grow at a rapid rate. Web-based methods can provide superior capabilities over other methods in some aspects of data collection (Dillman 2000, pp. 352–6, 372–401; Poynter 2001). Furthermore, SRS has a long and successful history of moving toward Web-based data collection (Meeks et al. 1998). For example, edit and logic checks can be built into the software—these checks verify responses to ensure they agree with earlier responses and ask for clarification if they do not, respondents can complete the survey at a convenient time, and experience shows that a large proportion of total responses can arrive within days of the instrument's deployment on the Web.

Because of low response rates from general populations, however, it is unlikely that Web-based data collection will completely replace the traditional modes of data collection used in the SESTAT surveys. However recent experience suggests that it may provide a useful way to supplement traditional methods in the future. Such use is explored further in the sections below. The first section discusses recent research in Web collection and the second section discusses specific issues related to the three surveys.

Web Collection Research

During the 1999 National Survey of Recent College Graduates Followup Survey (NSRCG Panel), Westat conducted an experiment to test Web-based collection (Collins and Tsapogas 2000). The experimental effort was focused on the panel survey, rather than the NSRCG baseline (new graduate) survey, because the panel members are easier to locate and the 1997 cycle provided information on their access to the World Wide Web and their willingness to respond using such a method. A targeted sample rather than a representative sample of 3,500 panel members was selected for the experiment. Only panel members who completed the 1997 survey, said they had Web access, said they would be willing to respond to a Web-based survey, and had a "mailable" address (an address that was complete and had not been identified as invalid during 1997) were eligible for the experiment. Experimental sample members received two mailings, about a month apart, asking them to complete the Web survey. The mail packages included a letter from NSF and a question and answer sheet, both designed to encourage response, ensure confidentiality, and provide the instructions needed to complete the Web survey, including the sample member's personal identification number (PIN) and individual password. If no completed survey was received via the Web after about 2 months, then the experimental case was sent to CATI followup. Respondents who indicated during CATI followup that they wanted to complete the survey on the Web rather than on the telephone were given about a week to do so. If the survey was not completed within a week, then followup calls were resumed. Of the experimental sample of 3,500 panel members, about 27% used the Web application to respond; an additional 60% of these cases were completed through telephone interviewing and (in a few cases) by mail.

The experience with the NSRCG Panel appears to be in line with what other survey researchers reported at the May 2000 Annual Conference of the American Association of Public Opinion Research in a seminar entitled "Facing the Challenges of the New Millennium." At that conference, several major findings regarding the use of Web-based data collection became apparent. First, regardless of the population being surveyed (college students, marketing managers, businesses and professionals,

federal employees, or teachers), researchers employing Web-based data collection were obtaining response rates of about 25% or less. Second, most of the participants who responded to the Web-based survey did so within a very short period of time after the start of data collection. After that initial flurry of response, the number of respondents dropped off dramatically. A second mailing appeared to have little effect on substantially increasing the response rate after the initial flurry, even if the data collection period was extended. Third, several of the researchers noted that item nonresponse rates were lower for the Web-based survey as compared with some other modes of data collection.

Several conference sessions addressed Internet usage in general. Researchers reported that the number of individuals with access to the Internet is increasing, the population of users is ethnically diverse, and the number of low income and less educated Internet users is increasing. For example, for the first time a substantial number of high school graduates are online. A number of researchers reported that they expect to see response rates for Web-based surveys increase substantially over the next few years.

Westat researchers attended conferences (sponsored by Fed-CASIC in Washington, DC and by the Association for Survey Computing in London) that echo these trends (e.g., see Couper 2001 and Flatley 2001). These impressions also match Westat's past and current experience with dozens of Web-based data collection efforts on a variety of topics and with a variety of respondent populations.

Other reviews of research on Web-based data collection affirm that it is growing substantially in popularity, but mostly in two general situations: (1) surveys that are not particularly concerned about achieving a high response rate from a representative sample or (2) surveys in which all respondents are part of a single organization, association, or network that communicates through Web-based interfaces regularly (Dillman 2000, pp. 354–5; Deutschmann and Faulbaum 2001). Because neither of these situations applies to the SESTAT surveys, it appears that for the time being Web-based surveys will at best provide an alternate mode of data collection to supplement traditional methods.

Moreover, other problems cited in recent literature might be expected to mitigate the potential advantages of Web-based data collection. For example, differences in Web browser capabilities and line transmission speeds necessarily limit creativity and flexibility in survey design. Ironically, this could mean that the Web-based version of a survey instrument might need to be of a simpler design than the paper questionnaire. Also, some respondents' unfamiliarity with the technical capabilities of their computer equipment may result in inaccurate responses or complete failure to respond. Furthermore, e-mail addresses tend to have more multiplicity and instability than street addresses or residential telephone numbers. Finally, Web-based data collection can raise privacy concerns that may affect the quality of the data being collected (Dillman 2000, pp. 352–8, 372–6).

Web Collection Issues for SESTAT Surveys

The most effective use of Web collection for the three SESTAT surveys is expected to vary by survey, just as the use of other data collection modes varies by survey. Web collection is dependent on contacting sample members by mail or e-mail to send the request to complete the survey (the Web address, PIN, and password are needed for the survey to be completed). Therefore, the quality of the addresses available for sample members early in the data collection period is critical for this approach. All three SESTAT surveys have been using mail preceded by address updating activities

to contact sample members. The NSCG and the SDR collect data by mail with CATI followup. These studies use the addresses obtained from previous survey cycles as well as address updating activities to reach a large proportion of their samples by mail. Historically, each survey completed over 60% of the sample by mail prior to telephone followup, with additional mail responses received after telephone contacts. In contrast, the NSRCG has historically collected most data with CATI but has used mail to help locate sample members. The NSRCG has used a mail flier designed to introduce the study and to request the return of an address/telephone update form by mail. These fliers were mailed to addresses obtained from the colleges and universities and from address update activities conducted prior to and during mail flier update collection. In the 1997 NSRCG baseline, initial and followup flier mailings were conducted prior to CATI collection. Completed fliers were received from 26% of those mailed, which was 24% of the sample (fliers could not be mailed to some graduates because no locating information was available for them without extensive CATI tracing activities). The early NSRCG surveys that used mail collection with telephone followup experienced difficulty in contacting recent graduates by mail; final response rates of 68% and 73% were achieved in 1988 and 1990, respectively. By comparison, response rates of 82% to 86% were achieved during the 1993 through 1997 cycles using CATI collection.

In addition to postal mail, e-mail contacts can be used to request completion of a Web survey. E-mail messages could be sent either as the initial contact or as reminders. Using e-mail as an initial contact causes some confidentiality concerns about sending respondents' PINs and passwords by e-mail. There is less control over who reads e-mail messages than who reads sealed letters. Although it is a federal offense to tamper with postal service mail, employers may legally monitor e-mail messages sent to work addresses. Although most postal addresses used in the past have been home addresses, a number of survey e-mail messages will likely reach sample members at work. The availability of e-mail addresses also varies by survey. The continuing panel components of the SDR and the NSCG have e-mail addresses that are available from the previous survey cycle, but the NSRCG does not have access to these e-mail addresses. The NSRCG faces some of the same difficulties in obtaining current e-mail addresses as in obtaining current postal addresses early in the data collection. At this time, there do not appear to be any searches for updated e-mail addresses that can be done in "batch mode" in the way that the National Change of Address can be used for postal addresses. Instead, searches for e-mail addresses are usually done on a case-by-case basis, which can be time consuming and expensive. In the 2001 NSRCG, colleges and universities were asked to supply e-mail addresses along with other contact information on the graduate sampling lists. E-mail addresses were obtained from colleges for about 22% of the sample and were used to send e-mail fliers. In the first batch of e-mail fliers, about 2,500 messages were sent; about 40% were returned as undeliverable. Completed responses were received from 5% of the messages that were sent or about 8% of those not returned as undeliverable.

These experiences demonstrate that the ability to reach sample members by mail and e-mail is very different for the NSRCG than for the other two SESTAT surveys. Because of these differences as well as population differences, Web collection will be discussed separately for each survey.

Survey of Doctorate Recipients

The SDR survey population can be expected to have wide access and familiarity with the Web, especially because 40% of the sample members are employed in academic institutions. This feature, combined with the ability to reach a large portion of the

sample by mail and e-mail, makes Web collection very attractive for this survey. The Web collection test planned by NSF for the 2003 SDR will provide important information for this survey. Additional tests may be needed to determine the most efficient approach for including the Web collection with the other data collection modes. For example, the following questions need to be answered: (1) should the letter requesting completion of the Web survey be sent before the mail questionnaire or with the questionnaire, giving the respondent a choice of response mode? and (2) should the mail questionnaire collection be scaled back drastically (i.e., is it expected that the Web will replace mail collection in the future)? These types of questions may be answered through observations of response patterns and through formal experiments. It is expected that CATI telephone followup will still be needed to reach the response rate goal for the survey.

National Survey of College Graduates

The NSCG survey has also reached a large portion of the sample through mail, but this population may have a smaller proportion of Web users than the SDR survey. Requests to complete the Web survey will probably reach most sample members, but the proportion that will complete the survey on the Web is unknown. The experience of the 1999 NSRCG panel Web survey experiment may provide some guidance, although there are differences in the populations. As discussed earlier, that experiment obtained a 27% Web response rate among those who said they had Web access, said they would be willing to respond to a Web-based survey, and had a "mailable" address from the previous cycle. Although this experience is not directly applicable to the NSCG, it would be most relevant for the redesign option 2, in which the same panel of NSCG cases selected in the 1990s is continued into the 2000 decade. The Web collection issues that are dependent on the redesign options are discussed below.

With redesign option 1, a new sample would be selected for the NSCG from the 2000 census data. A large screening effort would be needed to identify the eligible sample members. A Web survey has several desirable characteristics to assist with this screening effort. A significant portion of the sample could be reached by mail using the census addresses with appropriate address updating; however, this would be a smaller proportion than that for the current NSCG panel. The Web survey could be designed to ask the screening questions early in the questionnaire and then follow different paths based on the responses to the screening questions. Because the respondent would not know the path his or her response would follow, the potential for bias (compared with a similarly designed mail survey) would be reduced. However, the ability to change responses for these questions would need to be limited. There are also issues related to field of study, which is one of the primary screening items. Respondents do not always self-select the most appropriate code for field of study and their eligibility could change if the education code is corrected. One way to address this issue is to inform respondents that they might be contacted for more information later and to collect all locating information at the end of the Web survey or screener. Completed NSCG Web screeners/surveys are expected to have lower processing costs than mail surveys and lower collection costs than CATI surveys, assuming that a large enough number of Web responses are collected to offset the Web development costs.

Redesign options 2, 3, and 4 all involve the possibility of including individuals who were either dropped due to nonresponse during the 1990s or were never in the sample (in the case of option 4). In general, the more extensive telephone followup needed to locate and contact these sample members, the less efficient Web collection becomes.

If extensive telephone contacts are needed, then it is likely to be most efficient to complete the survey by telephone.

National Survey of Recent College Graduates

The NSRCG survey population consists of recent graduates who are expected to use the Web frequently. However, telephone contacts are needed to reach a large proportion of the sample. During the 1999 NSRCG, 62% of the new graduates required some tracing and 43% required intensive tracing. Identifying cases that need tracing early in the data collection period improves the efficiency and results of the tracing activities. Unlike mail or Web collection, telephone collection provides immediate feedback on the graduate's location, thus allowing efforts to be directed to the most useful collection activities for each case—either tracing, questionnaire collection, or refusal conversion. Collection of referral information is another important aspect of tracing in which the telephone is more effective than mail or the Web. Referrals include locating information supplied by the graduates' relatives, friends, or other contacts. During the 1997 cycle, referrals were the second most productive source of addresses and telephone numbers for new graduates. Twenty-four percent of final survey responses were completed at referral telephone numbers. These statistics, however, do not reflect the full impact of referral information. Contacts who cannot provide the graduate's exact telephone number or address may provide other important information such as the city of residence, employer, or college/university that the graduate attends. This information can then be used to obtain the address and telephone number through other sources, such as directory assistance and college contacts. Although these other sources then appear as the source of the final response, the referral information provides a critical link in the tracing chain.

Web Collection Summary

Assuming that issues concerning comparability of data can be resolved, it appears that a Web-based survey may provide a useful way of augmenting the traditional modes of data collection used in the SESTAT surveys, especially the SDR. The NSCG is also expected to benefit from the addition of Web collection, but this will vary depending on the redesign option chosen. The benefits to the NSRCG are less certain due to the extensive telephone contacts needed for tracing graduates.

Summary

The target population for the SESTAT surveys includes residents of the United States who, as of the survey reference period, were noninstitutionalized, age 75 years or younger, had at least a bachelor's degree, and either had a bachelor's degree or higher in an S&E field or were working as a scientist or an engineer. However, certain groups that are intended to be in the target population are only partially covered in the current designs for the SESTAT surveys. The two main groups that are partially covered by the SESTAT surveys are those referred to as the "foreign degreed" and individuals who are working in an S&E occupation but do not have an S&E degree. Although these groups were mostly included in the NSCG, the sample has not been refreshed for these groups in later cycles.

Another important consideration for the design of the SESTAT surveys is the impact of attrition on sample size. Even with relatively high response rates in any given survey year, the cumulative response rates will decline over time. The three SESTAT surveys show somewhat different patterns of response. For the NSCG, the initial response rate was 80% in 1993 and the overall response rate in the second cycle was 76% (80% times 95%). The overall response rate decreased to 71% in 1997 and would be expected to decrease similarly in further rounds of the survey. The NSRCG response rates have declined slightly from 1993 to 1997, but were 82% or more in all years surveyed. In 1993, the unweighted SDR response rate was 87%. The unweighted SDR response rate of 77% in 1995 was due to the subsampling of nonrespondents, which was conducted in that cycle only. The 1995 weighted response rate that reflects the subsampling of nonrespondents was 85%. This is consistent with the 84% response rate achieved in the 1997 SDR. Thus, it may be reasonable to speculate that SDR response rates will remain at 80% or higher in subsequent rounds.

Although response rates generally have been high for both the NSRCG and the SDR, the NSCG is the largest component of the SESTAT integrated database and thus has had the greatest impact on overall response rates. Despite the high conditional response rates achieved for the NSCG, over time the cumulative effects of attrition inevitably have led to low overall response rates.

NSF has proposed four options for the redesign of the samples for the SESTAT surveys. Each option has strengths and weaknesses with regard to coverage gaps, sample attrition, and screening requirements. The four options also have some characteristics in common. First, the SDR would continue its current system of sampling and data collection under all of the options and the design of the SDR would not be affected by the choice among the four options. Second, under all four options the sample of experienced scientists and engineers would be updated every 2 years with new graduates from the NSRCG. Third, the NSRCG would not change. Fourth, the SESTAT surveys would be conducted about every 2 years from 2003 through 2009 regardless of the option chosen.

An important feature of probability sampling methods is that they permit the calculation of the sampling errors associated with the survey estimates. None of the four design options has a significant advantage over the others in regard to variance estimation. As long as appropriate variance stratum and unit identifiers are available in the data and related survey control files, reasonably good variance estimators can be developed using any of a number of well-known techniques. Because the four options all involve the same basic design elements (e.g., use of unclustered census samples plus what are essentially independent samples of recent college graduates

and doctorate recipients), the same general approach can be used (with minor modifications) for all four options.

Option 1 is a repetition of the current design. Under this option, the Census Bureau would conduct a postcensal NSCG survey in 2003 based on the 2000 census. Because the census data do not include the level of detail necessary to identify graduates eligible for the SESTAT database, a significant screening process would be needed. If the same response and eligibility rates occur in 2003 as occurred in 1993, less than 35% of those individuals selected from the 2000 census long form records would respond to the 2003 NSCG and be eligible for inclusion in the SESTAT database. This is the maximum screening rate among the four options considered.

Individuals with foreign degrees who were included in the 2000 census as well as those with non-S&E degrees who are working in S&E occupations at the time of the NSCG would be included as eligible sample members. Under this option, coverage of the foreign degreed and individuals with only non-S&E degrees who hold S&E occupations in 2003 would obviously be much greater than in the 1999 NSCG because there have been no attempts to update the original NSCG sample with these subgroups.

Under this option three major groups would be not covered or poorly covered in the SESTAT surveys: (a) eligible individuals who lived abroad as of the 2000 decennial census who later came to live in the United States and did not earn a bachelor's or higher S&E degree from a U.S. institution after April 2000; (b) individuals with only non-S&E degrees obtained after April 2000 who hold S&E occupations in the survey reference period; and (c) individuals with only non-S&E degrees, with at least one degree obtained before April 2000, who did not hold an S&E occupation in 2003 but held such an occupation in a later survey reference period.

If a new NSCG sample is drawn under option 1 and the response rate in 2003 is about the same as the 1993 response rate, that rate would be an estimated 15% higher than the response rate for the existing NSCG sample, which declined during the decade. Also, if any conditioning biases have been introduced into the panel over time, the new sample will be subject to smaller panel effects than the existing sample.

Under option 2, the current NSCG sample based on the 1990 census would continue; some attempt will be made to try to update gaps in coverage where possible and cost effective. To decrease the nonresponse rate, NSF would go back to the original samples from 1993 as well as the later panels and try to trace the nonrespondents.

The 2000 census would be used in a limited way to augment the sample with the "missing" subpopulation of the foreign degreed and, perhaps, those with non-S&E degrees in April 2000 who held S&E occupations in the survey reference period in 2003. An efficient screening could be done for individuals who were born and educated abroad, who can be sampled almost directly from the census by stratifying by country of origin and year of entry into the United States, but the foreign degrees would need to be screened to identify S&E degrees. U.S. citizens educated abroad since the 1990 census would not be covered. Individuals with only non-S&E degrees who are working in S&E occupations cannot be identified in advance of sampling, so a screening activity would be needed. The level of effort for screening under option 2 depends on the desired sample size for individuals with non-S&E degrees screened from a sample of college graduates identified in the 2000 census. If the expected sample size for this subgroup was the same as for option 1, then the level of screening would be roughly the same as for option 1. On the other hand, if the SESTAT

population was restricted to individuals with S&E degrees (so those without S&E degrees would be excluded from sampling), then screening the census population would essentially be eliminated. Between the two extremes are intermediate positions in which the population of individuals with only non-S&E degrees who are working in S&E occupations are (a) restricted to certain occupations or (b) sampled at varying rates depending on occupational group. Position (a) would lead to undercoverage bias and position (b) would avoid undercoverage but lead to increased sampling errors. Either of these positions could plausibly lead to screening rates that are one-half to three-quarters of the rates for option 1. Further analysis would be required to determine the approximate optimum sampling rate under the various scenarios.

Under this option (in which the 2000 census would only be used to sample the foreign degreed and individuals with only non-S&E degrees), the coverage gaps would grow throughout the decade in the same way as with option 1. The refreshment of the NSCG sample with the foreign degreed and individuals without S&E degrees from the 2000 census would produce coverage similar to option 1.

This option allows limited flexibility in terms of allocating sample size in the old sample. Furthermore, unless the tracing efforts are successful, the response rate that is a concern in the current design would not be improved. Weighting adjustments needed to compensate for nonresponse tend to increase the variation in the weights, which in turn tend to increase sampling errors.

The primary advantages of this option (apart from the cost savings associated with retaining previous respondents) are (1) it maintains the longitudinal aspects of the original SESTAT design and (2) small domain estimates are more stable over time. Specifically, option 2 would permit analysis of change among individuals who were in the study for two or more rounds in different decades. This potential for longitudinal analysis would be lost under option 1 until subsequent data collection rounds are completed, and would be nonexistent for comparisons involving a time interval of more than 10 years (e.g., spanning different decades).

Option 3 combines features of options 1 and 2. Part of the NSCG sample would be selected using option 1 and the remainder would be selected using option 2. Under this option, the 2000 census would be used to draw a sample of college graduates that is larger than half the size of the sample planned under option 1. The 2000 census sample size can be set to yield sufficient S&E cases to feed into the followup panel about 2 years later with the balance available from samples from the 1999 NSCG panel. The subpopulations consisting of the foreign degreed and those with non-S&E degrees who have moved into S&E occupations since the 1990 census would be represented in this part of the total sample (with this subpopulation oversampled to provide about the same amount of coverage as in option 1). The remaining part of the sample would be derived from a large subsample of the 1999 NSCG panel. The coverage gaps under this option would be the same as in option 1. However, the sample sizes for the undercovered subgroups may not be the same.

This option permits an assessment of possible "panel effects" in the existing NSCG sample. If the two samples produce comparable results, then they can be combined with relatively little loss in efficiency. On the other hand, if the comparisons indicate that estimates from the existing panel are markedly different from those based on the new sample, then differential bias may be presumed. In this case, the new sample can still be used to make cross-sectional estimates (but with reduced levels of precision). Analysis of the differences might also provide improved nonresponse adjustment methodology that would bring the estimates from both frames back into effective

joint use. This option would produce the most S&E cases and has the best control for rare demographic S&E groups if the samples can be combined.

Option 4 is a variant of option 2 with supplementation from the NSRCG. The old panels would continue as in option 2. Assuming that new samples can be drawn from the old NSRCG panel lists, these samples could be used to supplement the old panels. There could also be an attempt to contact old nonrespondents. As in option 2, the 2000 census would be used for the limited purpose of refreshing the sample with the "missing" subpopulation of the foreign degreed and, perhaps, those with non-S&E degrees who held S&E occupations in 2003.

The coverage gaps would be the same as in option 2. Sampling from the old NSRCG frames would not cover any missing subgroup but would only address the decline in sample size through attrition. It should be noted that the resampling would not reduce the existing nonresponse bias but would merely provide another independent sample for comparison. As with option 3, the nonresponse bias, if any, could be studied. This option would require extensive tracing to locate individuals selected from the old NSRCG lists, and adds very little as compared with option 2.

Appendix: Discussion of Response Issues for SESTAT Redesign Options

TO:	Ron Fecso
FROM:	Sestat Redesign Team
DATE:	March 14, 2001, Revised with 3/27/2001 comments from NSF
SUBJECT:	Discussion of Response Issues for Sestat Redesign Options

The purpose of this document is to discuss some of the issues related to the four possible SESTAT redesign options that Westat was asked to investigate for the Division of Science Resources Studies of the National Science Foundation (NSF/SRS). It is not intended to provide a complete description of all considerations for each option, but instead focuses on operational issues related to locating and contacting sample members and obtaining sufficient response rates.

Redesign Option 1

Option 1 replicates the design used in the decade of the 1990s. It involves drawing a new sample of college graduates from the 2000 Census for the National Survey of College Graduates (NSCG). This would be supplemented throughout the decade with samples from the National Survey of Recent College Graduates (NSRCG). Response patterns and problems in the 2000 decade are expected to be similar to those experienced in the 1990 decade. Although many surveys have experienced declining response rates in recent years, the positive image of the NSF and Census Bureau is expected to help offset this trend. Since this option is expected to be very similar to the past decade. It will not be discussed further here, except to note that SRS could modify the sampling strategy based on the 1993 experience to more carefully target the population of interest, and may change the target population definition to reduce the required sample.

Redesign Options 2 and 3

Option 2 involves continuing with the current panels, with the NSRCG continuing to contribute the new domestic S&E bachelor's and master's degree population each cycle. The SDR will provide new domestic S&E doctorate earners, and the decennial census will contribute individuals with foreign-earned S&E degrees at all levels. Option 3 is a combination design, with half the sample following the option 1 design and half following the option 2 design. For both options 2 and 3, NSF is considering whether to include members of the original samples that were dropped due to nonresponse during the 1990s. This would be used as a way to reduce the potential nonresponse bias. Some of the issues involved in including nonrespondents from previous cycles are discussed below.

First, it is helpful to review the design of the studies involved. The original design for the 1990s involved adding a sample of respondents from the previous NSRCG New Graduate survey to the NSCG to represent that cohort in the followup survey. For example, in the 1993 cycle, the NSRCG consisted of a sample of individuals who earned new S&E bachelor's and master's degrees in the spring 1990, 1991, and 1992 academic years. In the 1995 survey cycle, these 1993 sample cases were moved into the NSCG sample frame; the 1995 NSRCG includes only recent U.S. S&E bachelor's and master's degree earners from the 1993 and 1994 academic years. This procedure

of first contacting individuals with the NSRCG, and then later moving them to the panel sample frame was followed throughout the 1990s. These cases that were originally part of the NSRCG and later moved to the panel are collectively referred to as the NSRCG Panel.

During the 1990s, the NSCG was generally conducted only with sample members who responded to previous survey cycles. This procedure was not always followed for the cases that were part of the NSRCG Panel. Individuals in the NSRCG Panel had to respond to the NSRCG baseline survey to be included in the Panel frame the next year, but did not need to respond to the Panel followup survey to continue to remain in the frame. That is, once a sample member responded to the baseline, he/she was included in both Panel followup cycles. After two NSRCG followup cycles, the sample member became part of the NSCG and followed the rules for that survey.

When considering the tracing of nonrespondents from previous survey cycles, it is helpful to look at the cases that were sampled from the 1990 Census separately from those added from the NSRCG survey. There are three main differences between the two groups that have a significant effect on tracing. First, there are population differences. Most of those sampled from the Census are older and a greater fraction of these individuals are likely to have completed their education than those sampled as new graduates. As such, they may be easier to locate than recent graduates who have not yet established a permanent or semi-permanent address. Second, for those sampled from the Census, we started with a confirmed address that we know is where the sample member lived in 1990. For those sampled as new graduates, we started with information provided by the sampled colleges and universities. This information varies widely in terms of completeness and timeliness and may only include where the sample member lived while attending college. Third, all cases sampled from the Census were living in the U.S. in 1990, while those sampled as new graduates include non-resident aliens who may have left the country after graduation. Therefore, different decisions may be made for handling nonrespondents for the two groups.

Cases sampled from the 1990 Census. For this group, we should first consider nonrespondents to the baseline (1993) survey. Since this baseline survey was used to identify the eligible sample members, it would be very expensive to go back to nonrespondents from the 1993 survey. This would involve contacting a large number of sample members with unknown eligibility status where no contact attempts have been made since 1993. Instead, we expect that NSF will consider contacting those who responded to the baseline but did not respond to the followup surveys. But it should be pointed out that most of the nonresponse occurred during the baseline survey, with unweighted response rates of 78 percent for the baseline and 90-94 percent for each followup cycle.

Another consideration in contacting previous nonrespondents is the type of nonresponse. We can group nonrespondents into the three broad categories of refusal, non-locatable, and other (including those who were ill or temporarily absent; wrong sample persons; unable to contact despite repeated attempts; or those who were contacted, but for which critical data items were missing). Among the baseline NSCG sample members who were in the 1997 NSCG, there were 2,630 nonrespondents, which was 6 percent of this portion of the sample. Among these nonrespondents, 61 percent were refusals, 18 percent non-locatable, and 21 percent other. While this might seem to indicate that refusal conversion is a greater challenge than tracing, tracing generally becomes more difficult as time elapses, while refusal conversion may become easier. Cases that refused in the early survey cycles may become tracing problems in the 2003 cycle. Therefore, both refusal conversion and tracing are

important considerations for contacting nonrespondents to previous cycles.

The 1993 NSCG obtained a weighted response rate of 80 percent with a mail survey followed up by CATI and a personal interview. In 1995, a decision was made to follow up on only the respondents to the 1993 survey, as locating costs for following nonrespondents would have been prohibitive. Additionally, at the time, it was not expected that the baseline NSCG sample would be contacted after the 1990s decade, so the diminishing cumulative response rate and potential nonresponse bias that was a result of only following respondents was not expected to have a significant impact on the SESTAT system as a whole. It also saved the cost of locating and contacting nonrespondents.

As previously stated both refusal conversion and tracing are important to consider for contacting nonrespondents. No tests were conducted to determine the estimated success rate or cost of additional refusal conversion activities if refusals are added back into the NSCG sample. However, a tracing test of nonrespondents to the 1995 NSCG was conducted by the Census Bureau under NSF's direction, as explained below.

A simple random sample of 25 cases that responded to the 1993 NSCG baseline but did not respond to the 1995 NSCG were included in the tracing test. The sample was selected from a frame of people who were un-locatable or could not be contacted for the 1995 NSCG. The test used only non-invasive searches, so the new listings were not contacted to determine if the correct individual had been found; that is, sample person verification was not conducted. Therefore, all the address and phone number listings found are unconfirmed. Most searches were done using FastData, an address source using United States Postal Service (USPS) National Change of Address database and other information that provides address, phone number, date of birth or age range, persons living in that residence, and length of time at that address. Other searches were done using PowerFinder on CD-Rom and the Internet. A summary of the results for address and telephone number searches is shown in table 1 below.

TABLE 1. Results of the tracing test conducted by the Census Bureau in spring 2001, which included a random sample of 25 cases that responded to the 1993 NSCG but did not respond to the 1995 NSCG survey

	Address Search Results		Telephone Search Results	
2001 Tracing Test Results for Nonrespondents to the 1995 NSCG	Number	Percent	Number	Percent
Total	25	100%	25	100%
Found an unconfirmed address listing and matched on date of birth (these listings are expected to be for the correct sample member but may not be the current address/telephone number).	7	28%	5	20%
Found an unconfirmed listing but could not match on date of birth (these listings may not be for the correct sample member and they may not be the current address/telephone number).	11	44%	5	20%
Telephone number was non-published	NA	NA	8	32%
No listing found	7	28%	7	28%

NA=not available

Please note that results are reported in unweighted numbers and percents because this tracing test used a simple random sample.

As seen above, of the 25 cases in the test, unconfirmed address listings were found for 18 (72 percent). Only 7 of these (28 percent of the total) could be matched on birth date. Unconfirmed phone number listings were found for about 40 percent of the test cases and about one-third (32 percent) of the cases had non-published phone

numbers. While finding address listings for 72 percent of the test cases is encouraging, it should be noted that a number of these are expected to be outdated addresses where the sample member no longer lives, and some will be listings for someone else with the same name as the sample member.

Cases sampled from the NSRCG Baseline and included in the Panel sample. For this group, we should consider nonrespondents separately by whether they were nonrespondents in the baseline survey (1993, 1995, 1997 or 1999 NSRCG baseline survey) or did not respond to the Panel followup survey. We can also classify them by type of nonresponse. Each of the baseline surveys has resulted in a similar distribution of nonrespondents by type. In the 1997 NSRCG baseline, there were 2,573 nonrespondents, 18 percent of the sample. Among these nonrespondents, 37 percent were refusals, 50 percent non-locatable, and 13 percent were other. If NSF decides to include nonrespondents to past NSRCG baseline surveys in the 2003 sample, both tracing and refusal conversion activities are important, with tracing issues predominating. Since no followup survey has been conducted that included nonrespondents to the NSRCG baseline surveys, a tracing test was conducted by the Census Bureau, as described below.

This test included 40 cases that did not respond to the 1995 NSRCG. Since every NSRCG is a baseline survey, these tracing test cases never responded to any survey cycle. The sample was selected from a frame of 1,762 people who were either un-locatable or their household could never be contacted to confirm the graduate lived there in the 1995 cycle. The sample of 40 includes 10 people with foreign addresses provided by the school, 10 with no address provided by the school, and 20 with one or more U.S. addresses at the time of sampling. The sample was drawn using three sampling categories based on the type of address provided on the school sampling list in the 1995 cycle (no address, foreign address, or U.S. address provided). Since different sampling rates were used for the different categories, the tables below include weighted totals and weighted percents. Please note that these weights reflect the sampling for the tracing test only. That is, the weighted total of 1,762 is the number of 1995 NSRCG sample members (nonrespondents) eligible to be included in the tracing test. The weight for the "No address provided" category is 38.4 (384 in frame/10 in sample), for the "Foreign address provided" category the weight is 17.1 (171 in frame/10 in sample), and for the "U.S. address provided" category the weight is 60.35 (1,207 in frame/20 in sample).

The same tracing procedures were followed for this group that were followed for the NSCG tracing test. Both tests used only non-invasive searches, so the new listings were not contacted to determine if the correct individual had been found. A summary of the results for addresses and phone numbers is listed below, with table 2 showing the results of address searches and table 3 showing the results of telephone number searches.

Looking at the weighted percents for address searches in table 2 , we can see that unconfirmed address listings were found for about half (51 percent). However, only 10 percent of the total could be matched on birth date. As expected, results varied by type of address provided on the initial sampling list. For cases with no address or foreign address provided, none of the cases in the test sample could be matched on birth date and only 2 cases in each category had an address listing found.

Table 3 shows that un-confirmed phone number listings were found for 23 percent of the cases (with 3 percent matched on birth date) and 27 percent of the cases had non-published phone numbers. None of the cases in the "no address" and "foreign

TABLE 2. Results of the tracing test (address searches) conducted by the Census Bureau in spring 2001 for a sample of 40 cases that did not respond to the 1995 NSRCG baseline survey

2001 Tracing Test Results for Address Searches For Nonrespondents to the 1995 NSRCG	Unweighted Total	Unweighted counts by sampling category (type of address provided by school in the 1995 cycle)			Weighted Totals*	
		No address provided	Foreign address provided	U.S. address provided	Weighted Number	Weighted Percent
Total	40	10	10	20	1,762	100%
Found an unconfirmed address listing and matched on date of birth (these listings are expected to be for the correct sample member but may not be the current address). Note that 2 of these were listings for a new address and 1 was for an existing address.	3	0	0	3	181	10%
Found an unconfirmed address listing but could not match on date of birth because date of birth was missing (these listings may not be for the correct sample member and they may not be the current address). Of these 14 listings, 9 are for new addresses and 5 are for existing addresses.	14	2	2	10	715	41%
No address listing found	23	8	8	7	866	49%

* Please note that the weighted numbers and percents are weighted for the tracing test sampling only. That is, the weighted total of 1,762 is the number of 1995 NSRCG sample members (nonrespondents) eligible to be included in the tracing test. The weight for the "No address provided" category is 38.4 (384 in frame/10 in sample), for the "Foreign address provided" category it is 17.1 (171 in frame/10 in sample), and for the "U.S. address provided" category it is 60.35 (1,207 in frame/20 in sample).

TABLE 3. Results of the tracing test (telephone number searches) conducted by the Census Bureau in spring 2001 for a sample of 40 cases that did not respond to the 1995 NSRCG baseline survey

2001 Tracing Test Results for Address Searches for Telephone Number Searches For Nonrespondents to the 1995 NSRCG	Unweighted Total	Unweighted counts by sampling category (type of address provided by school in the 1995 cycle)			Weighted Totals*	
		No address provided	Foreign address provided	U.S. address provided	Weighted Number	Weighted Percent
Total	40	10	10	20	1,762	100%
Found an unconfirmed telephone listing and matched on date of birth (these listings are expected to be for the correct sample member but may not be the current phone number).	1	0	0	1	60	3%
Found an unconfirmed telephone number listing but could not match on date of birth because date of birth was missing (these listings may not be for the correct sample member and they may not be the current phone number).	7	1	1	5	357	20%
Telephone number was non-published	9	1	1	7	478	27%
No telephone number listing found	23	8	8	7	866	49%

* Please note that the weighted numbers and percents are weighted for the tracing test sampling only. That is, the weighted total of 1,762 is the number of 1995 NSRCG sample members (nonrespondents) eligible to be included in the tracing test The weight for the "No address provided" category is 38.4 (384 in frame/10 in sample), for the "Foreign address provided" category it is 17.1 (171 in frame/10 in sample), and for the "U.S. address provided" category it is 60.35 (1,207 in frame/20 in sample).

address" categories could be matched by birth date and only 1 un-confirmed telephone number listing was found in each category. It should also be noted that for both addresses and telephone numbers, a number of these listings are expected to be outdated addresses/telephone numbers where the sample member no longer lives, and some will be listings for someone else with the same name as the sample member.

It is interesting to compare the results of this tracing test (for nonrespondents to the NSRCG baseline) with the tracing test conducted for cases that responded to the 1993 NSCG baseline but did not respond to the 1995 NSCG, as discussed in the previous section. The NSCG tracing test involved searching for sample members for the fourth time (first for the 1990 Census, second for the 1993 NSCG, third for the 1995 NSCG, and fourth for the tracing test). The first two contacts resulted in successfully locating the individual and obtaining his/her cooperation. Both of these contacts provided a confirmed address and confirmation of the sample member's identify. The third contact attempt (for the 1995 NSCG) resulted in a nonresponse of either un-locatable or unable to contact. In contrast, the NSRCG tracing test conducted searches for sample members that had never been contacted. For some of these cases, no good address was ever obtained from the sampled college/university. These differences in test samples are highlighted by the different tracing test results, with no address listings found for 28 percent of the NSCG tracing cases and 49 percent of the NSRCG cases. In addition, 28 percent of the NSCG cases had an address listing matched by date of birth while 10 percent of the NSRCG cases could be matched by birth date.

While the NSRCG tracing test gives us information about tracing nonrespondents from the baseline NSRCG survey, we also need to look at the problems associated with tracing nonrespondents to NSRCG Panel followup surveys. In considering this group, we can use the example of cases that responded to the 1995 NSRCG baseline survey, were sampled for the 1997 Panel followup, but did not respond in 1997. These cases were included in the 1999 Panel followup survey, since they responded to the baseline survey. There were a total of 1,191 cases in this group (8 percent of the 1999 followup sample), which can be broken down by the type of 1997 nonresponse as follows: 46 percent refused, 20 percent were not located, 23 percent were other nonresponse, and 11 percent were temporary ineligibles. As would be expected, the response rates in the 1999 survey cycle vary by the type of nonresponse in the 1997 followup survey. Table 4 shows the distribution of cases by 1997 and 1999 survey response category and the 1999 cycle response rates by 1997 response category.

TABLE 4. Cases sampled in the 1995 NSRCG who responded to the baseline survey, by their response categories in the 1997 and 1999 followup surveys

	1997 Cycle Response Category					
1999 Cycle Response Category	Total	Complete	Ineligible	Refused	Not located	Other non-response
Complete	5,554	5,180	46	136	82	110
Ineligible	213	100	72	13	18	10
Refused	893	511	2	327	22	31
Not Located	392	192	9	49	93	49
Other Nonresponse	288	166	7	24	23	68
Total	7,340	6,149	136	549	238	268
1999 Cycle Response Rate	79%	86%	87%	27%	42%	45%

Among the three types of 1997 cycle nonrespondents (refused, not located, and other nonresponse), the refusals are the biggest group and have the lowest response rate (27 percent). For the un-locatable and other nonresponse categories, less than half responded. The cases that were temporarily ineligible in 1997 (mostly living out of the U.S.) had a high response rate in 1999, as most of them continued to be ineligible.

Redesign Option 4

For this option, the NSRCG Panel samples selected during the 1990s would be supplemented by new samples drawn from the original sampling frames. Since there is no viable sampling frame available for the 1993 NSRCG,* these new sample selections would begin with the 1995 NSRCG. There are two main areas of consideration for this. First, there is the issue of college/university cooperation and confidentiality. The second is the operational issues involved in using old sampling lists up to eight years old. Each of these areas is discussed below.

Colleges and universities were asked to provide sampling lists for each NSRCG survey cycle. All materials sent to the colleges implied that the lists were to be used for that one survey cycle, as was intended at the time each list was collected. Although no specific promises were made to colleges about the use of their sampling lists for later survey cycles, most colleges would not expect that their lists would be kept and used years later. The NSRCG confidentiality plan states, "At the close of each study, survey materials are placed in secure storage for a period of 3 years…After this period has lapsed, the materials are disposed of." If this option is chosen, we suggest that NSF consider contacting the sampled colleges to ask their permission to use the old sampling lists they provided. While we expect that most colleges will give permission, some may be concerned that their sampling lists were kept for such a long time period.

The second issue involves the operational steps necessary to restore the old sampling lists and select the new samples. Lists were provided by colleges in either computer files or on paper. For lists provided on paper, the sampling information was keyed for all eligible graduates and was included in the sampling frames along with the computer lists. Therefore, it is possible to draw a new sample from both paper and computer lists once the sampling frame files and documentation have been restored. However, all identifying and locating information from paper lists was keyed after sampling, only for the sampled graduates. This means that if a new sample is drawn, it must be matched back to the paper list by ID number and the information for the new sample must be keyed from these old paper lists. As would be expected, the percent of sampling lists provided by paper varies by cycle, with 26 percent for 1995, 17 percent for 1997, 15 percent for 1999, and 9 percent for 2001. Time and money must be allowed for processing the information from these paper lists if this option is chosen.

In addition to the sampling issues, there are considerations about locating and contacting graduates who have never been contacted in earlier cycles. These issues are similar to those discussed for option 2 and 3 for contacting nonrespondents. Once the new samples are selected, the alumni offices at the sampled schools will be asked to provide updated locating information. We expect that additional substantial tracing will be needed to locate these graduates. In the 1997 NSRCG baseline survey, only 28 percent of the sample were interviewed at the address provided by the schools on the sampling lists or provided by the alumni offices.

Footnotes

* The contractor that drew the 1993 NSRCG sample is no longer involved with the SESTAT studies and is not expected to have maintained the sampling frames for this length of time (since they were not project deliverables). In particular, the lists sent on paper are not expected to be available.

Bibliography

Chu A, Brick J, Kalton G. 1999. Weights for combining surveys across time and space. *Bulletin of the International Statistical Institute, Contributed Papers* 2:103–4.

Collins M. 2001. Web-based data collection in two studies: individuals and institutions. Paper presented at Fed-CASIC conference, February 2001, Washington, DC.

Collins M, Tsapogas J. 2000. An experiment in Web-based data collection. Paper presented at annual meeting of the American Association of Public Opinion Research, May 2000, Portland, OR.

Couper MP. 2001. The promises and perils of Web surveys. Paper presented at conference of the Association for Survey Computing (The Challenge of the Internet); May 2001; London, England.

Deutschmann M, Faulbaum F. 2001. The recruitment of online samples by CATI-screening: problems of non-response. Paper presented at conference of the Association for Survey Computing (The Challenge of the Internet). May 2001, London, England.

Dillman DA. 2000. *Mail and Internet Surveys.* New York: John Wiley and Sons.

Hansen M, Tepping B. 1985. Estimation of variance in NAEP. Unpublished manuscript. Rockville, MD: Westat.

Flatley J. 2001. The internet as a mode of data collection in government social surveys: issues and investigations. Paper presented at conference of the Association for Survey Computing (The Challenge of the Internet), May 2001, London, England.

McCarthy P. 1966. Replication: An approach to the analysis of data from complex surveys. *Vital and Health Statistics* 2(14). Washington, DC: U.S. Department of Health, Education, and Welfare.

Meeks RL, Lanier AT, Fecso RS, Collins MA. 1998. Web-based data collection in National Science Foundation surveys. Proceedings of the Survey Research Methods Section, American Statistical Association.

National Science Foundation, Division of Science Resources Studies (NSF/SRS). 1999. *SESTAT: A Tool for Studying Scientists and Engineers in the United States.* NSF-99-337. Nirmala Kannankutty and R. Keith Wilkinson. Arlington, VA: National Science Foundation.

Poynter R. 2001. A guide to best practice in online quantitative research. Paper presented at conference of the Association for Survey Computing (The Challenge of the Internet), May 2001, London, England.

U.S. Census Bureau. 1998. *1997 National Survey of College Graduates, Methodology Report.*

U.S. Census Bureau. 2001. *1999 Survey of Doctorate Recipients, Methodology Report.*

Westat. 1999. *1997 National Survey of Recent College Graduates, Methodology*

Report. Washington, DC: U.S. Department of Commerce, U.S. Census Bureau.

Westat. 2000. *WESVAR 4.0 User's Guide.* Rockville, MD.

Wolter K. 1985. *Introduction to Variance Estimation.* New York: Springer-Verlag.

Suggested Citation, Acknowledgments

Fecso RS, Baskin R, Chu A, Gray C, Kalton G, Phelps R. 2007. *Design Options for SESTAT for the Current Decade: Statistical Issues.* Working Paper SRS 07-201. Arlington, VA: Division of Science Resources Statistics, National Science Foundation.

The authors thank Norman Bradburn, Lynda Carlson, and Mary Frase for management support and editorial input; Joan Mitchie and Cindy Gray of Westat for review and contract support; and internal reviewers Kelly Kang, Mark Regets, Wendy Scholetzky, and Nirmala Kannankutty for helpful comments on early versions of the report.

Division of Science Resources Statistics

Lynda T. Carlson	Mary J. Frase
Division Director	*Deputy Director*

Division of Science Resources Statistics (SRS)
The National Science Foundation, 4201 Wilson Boulevard, Arlington, Virginia 22230, USA
Tel: (703) 292-8780, FIRS: (800) 877-8339 | TDD: (800) 281-8749