

A. Collection of Information Employing Statistical Methods

This submission requests clearance for the full-scale 2008 National Postsecondary Student Aid Study (NPSAS:08). All procedures, methods, and systems to be used in the full-scale study were tested in a realistic operational environment during the field test conducted during the 2006–07 academic year. Specific plans for full-scale activities are provided below.

1. Respondent Universe

a. Institution Universe

To be eligible for the NPSAS:08 full-scale study, institutions are required during the 2007–08 academic year to:

- offer an educational program designed for persons who have completed secondary education;
- offer at least one academic, occupational, or vocational program of study lasting at least 3 months or 300 clock hours;
- offer courses that are open to more than the employees or members of the company or group (e.g., union) that administers the institution;
- have a signed Title IV participation agreement with the U.S. Department of Education;
- be located in the 50 states, the District of Columbia, or Puerto Rico; and
- be other than a U.S. Service Academy.

Institutions providing only avocational, recreational, or remedial courses or only in-house courses for their own employees are excluded. U.S. Service Academies are excluded because of their unique funding/tuition base.

b. Student Universe

The students eligible for inclusion in the sample for the NPSAS:08 full-scale study are those who were enrolled in a NPSAS-eligible institution in any term or course of instruction at any time from July 1, 2007 through April 30, 2008 and who were

- enrolled in *either* (a) an academic program; (b) at least one course for credit that could be applied toward fulfilling the requirements for an academic degree; *or* (c) an occupational or vocational program that required at least 3 months or 300 clock hours of instruction to receive a degree, certificate, or other formal award;
- not currently enrolled in high school; and
- not enrolled *solely* in a GED or other high school completion program.

2. Statistical Methodology

a. Sample design and proposed augmentations

The details describing the design and allocations of the institutional and student samples are presented in sections 2.b and 2.c. This first section describes two augmentations to the sample design as it was originally proposed.

The first augmentation involves oversampling 5,000 recipients of SMART grants and/or Academic Competitiveness Grants (ACG) (two new sources of student financial aid), to ensure that these students are sufficiently well represented for analysis. RTI will establish sampling rates for SMART grant recipients from a file that is to be provided by ED no later than December 2007. After establishing sampling rates, we will use the ED file to flag SMART grant recipients on lists provided by institutions.

More students are expected to receive ACG than SMART grants, so an oversample of ACG recipients may not be necessary. We will look at sample sizes with and without oversampling and at the effects of oversampling on variance estimates. In consultation with NCES we will decide if an ACG oversample is necessary. If oversampling of ACG recipients is not necessary, then the additional sample of 5,000 students will be only for SMART grant recipients.¹

The second augmentation is contingent upon 1) funding of a pending proposal to the Department Of Education and 2) a planned modification to that proposal based on discussions with the Commissioner of the National Center for Education Statistics as well as attendees of the recent Technical Review Panel meeting (held 8/28-29, 2007). The NPSAS:08 full-scale sample will be augmented to include state-representative samples of undergraduate students in four sectors from six states which will make it possible to produce state-level analyses and comparisons of many of the most pertinent issues in postsecondary financial aid and prices.²

As originally designed, the NPSAS:08 sample yields estimates that are nationally representative but generally not large enough to permit comparison of critical subsets of students within a particular state. Tuition levels for public institutions (attended by about 80 percent of all undergraduates) vary substantially by state, as does the nature of state grant programs (i.e., large versus small, need-based versus merit-based). Therefore, it is possible to analyze the effect of these policies and programs with federal and institutional financial aid policies and programs only at the state level.

The choice of states for the sample augmentation was based on several considerations, including

- **Size of undergraduate enrollments in four sectors:** public 4-year, private not-for-profit 4-year, public 2-year, and private for-profit, degree-granting institutions. We estimate that we will need approximately 1,200 respondents per state in the 4-year and for-profit sectors and 2,000 respondents in the public 2-

¹ The sample design described below assumes that SMART grant recipients will be oversampled and ACG recipients will not be oversampled.

² The field test institutional sample was selected from the complement of institutions selected for the full-scale study to avoid asking an institution to participate in both. After field test data collection, ED requested that RTI augment the full-scale sample to provide state-level representation of students in selected states and sectors. To accomplish this goal, it will be necessary to include a small number of institutions that participated in the field test in the full-scale study.

year sector in order to yield a sufficient number of full-time, dependent, low-income undergraduates—the subset of students that is of particular relevance for the study of postsecondary access. Tuition and grant policies in the states with the largest enrollments have the greatest effect on national patterns and trends. As a practical matter, their representation in a national sample is already so large that the cost of sample augmentation is relatively low.

- **Prior inclusion in the NPSAS:04 12-state sample and high levels of cooperation and participation in that survey.** Participation in NPSAS is not mandatory for institutions, so we depend on institutional cooperation within a state to achieve the response rates and yields required for reliable estimates. Smaller states that were willing and helpful in NPSAS:04 and achieved high yields and response rates are more likely to cooperate again, and with less effort.
- **States with different or recent changes in tuition and state grant policies that provide opportunities for comparative research and analysis.**

Using these criteria, we proposed to augment the samples for the following 6 states: California, Texas, New York, Illinois, Georgia, and Minnesota.

The sample sizes presented in this document reflect the inclusion of the SMART grant oversample and the state-representative samples. The institution sampling strata will be expanded to include strata for the four sectors within each of the six states. For selecting institutions within states and sectors, there are three scenarios. First, for some sectors in the states, there are already enough institutions in the sample, so that no additional sample institutions are necessary. In this case, the institutions already selected will stay in sample. Second, for other sectors in the states, all institutions in the sector in the state will be in sample. Therefore, the institutions already selected will remain in the sample, and the remaining institutions will be added to the sample. Third, for other sectors in the state, additional institutions need to be added to the sample, but not all institutions will be selected. In this case, the originally selected institutions are no longer necessarily in sample, and a new sample will be selected. This is the cleanest method statistically and is also best to keep the unequal weighting effect (UWE) from being too large. In the second and third scenarios, it is anticipated that a total of about 20 field test sample institutions may be included in the full-scale sample.

Also, the student strata will be expanded to include SMART grant recipients and to include in-state and out-of-state students.

b. Institution Sample

The institution samples for the field test and full-scale studies were selected simultaneously, prior to the field test study. The institutional sampling frame for the NPSAS:08 field test was constructed from the 2004-05 Integrated Postsecondary Education Data System (IPEDS) institutional characteristics, header, completions, and fall enrollment files. Three hundred institutions were selected for the field test from the complement of institutions selected for the full-scale study to minimize the possibility that an institution would be burdened with participation in both the field test and full-scale samples, while maintaining the representativeness of the full-scale sample. However, since the decision to augment the full-scale sample to provide state-level representation of students in selected states and sectors was

made after field test data collection was completed, it will be necessary to include in the full scale study about 20 institutions that also participated in the field test (as described above).

The full-scale sample was then freshened in order to add newly eligible institutions to the sample and produce a sample that is representative of institutions eligible in the 2007-08 academic year. To do this, we used the IPEDS:2005-06 header, Institutional Characteristics (IC), Fall Enrollment, and Completions files to create an updated sampling frame of currently NPSAS-eligible institutions. This frame was then compared with the original frame, and 167 new or newly eligible institutions were identified. These 167 institutions make up the freshening sampling frame. Freshening sample sizes were then determined such that the freshened institutions would have similar probabilities of selection to the originally selected institutions within sector (stratum) in order to minimize unequal weights and subsequently variances.

Institutions were selected for the NPSAS:08 full-scale study using stratified random sampling with probabilities proportional to a composite measure of size,³ which is the same methodology that we used for NPSAS:96, NPSAS:2000, and NPSAS:04. Institution measures of size were determined using annual enrollment data from the 2004-05 IPEDS Fall Enrollment Survey and bachelor's degree data from the 2004-05 IPEDS Completions Survey. Using composite measure of size sampling ensures that target sample sizes are achieved within institution and student sampling strata while also achieving approximately equal student weights across institutions.

We expect to obtain an overall eligibility rate of 98 percent and an overall institutional participation (response) rate of 84 percent⁴ (based on the NPSAS:04 full-scale study). Eligibility and response rates are expected to vary by institutional strata. Based on these expected rates, the institution sample sizes (after freshening)⁵ and estimated sample yield, by the nine sectors traditionally used for analyses, are presented in table 7.

³ Folsom, R.E., Potter, F.J., and Williams, S.R. (1987). Notes on a Composite Size Measure for Self-Weighting Samples in Multiple Domains. *Proceedings of the Section on Survey Research Methods of the American Statistical Association*, 792-796.

⁴ The institution response rate of 84 percent assumes that institutional participation will not be mandatory.

⁵ The institution sampling frame was constructed from the IPEDS:2004-05 header, Institutional Characteristics, Fall Enrollment, and Completions files. We freshened the institution sample in order to add newly eligible institutions to the sample and produce a sample that is representative of institutions eligible in the 2007-08 academic year, using the corresponding IPEDS files for 2005-06.

Table 7. NPSAS:08 expected full-scale estimated institution sample sizes and yield

| Institutional sector | Frame count ¹ | Number sampled | Number eligible | List respondents |
|---|--------------------------|----------------|-----------------|------------------|
| Total | 6,777 | 1,962 | 1,940 | 1,621 |
| Public less-than-2-year | 247 | 22 | 19 | 14 |
| Public 2-year | 1,167 | 449 | 449 | 383 |
| Public 4-year non-doctoral | 358 | 199 | 199 | 169 |
| Public 4-year, doctoral | 290 | 290 | 290 | 250 |
| Private not-for-profit less-than-4-year | 326 | 20 | 20 | 18 |
| Private not-for-profit 4-year, non-doctoral | 1,017 | 359 | 346 | 284 |
| Private not-for-profit 4-year doctoral | 591 | 269 | 269 | 209 |
| Private for-profit less-than-2-year | 1,476 | 97 | 91 | 77 |
| Private for-profit 2-year or more | 1,305 | 257 | 257 | 217 |

¹ Institution counts based on IPEDS:2004-05 header file.

NOTE: Detail may not sum to totals because of rounding.

The nine sectors traditionally used for NPSAS analyses were the basis for forming the institutional strata. These are

1. public less-than-2-year
2. public 2-year
3. public 4-year non-doctorate-granting
4. public 4-year doctorate-granting
5. private not-for-profit less-than-4-year
6. private not-for-profit 4-year non-doctorate-granting
7. private not-for-profit 4-year doctorate-granting
8. private for-profit less-than-2-year
9. private for-profit 2-year or more.

Since the NPSAS:08 student sample will be designed to include a new sample cohort for a Baccalaureate and Beyond Longitudinal Study (B&B), these nine sectors will be further broken down to form the same 22 strata used in NPSAS:2000 (the last NPSAS to generate a B&B study) in order to ensure sufficient numbers of sample students within 4-year institutions by various degree types (especially education degrees, an important analysis domain for the B&B longitudinal study). Additionally, 24 strata are necessary for the state sample, as described above. The 46 institutional sampling strata are as follows:

1. public less-than-2-year;
2. public 2-year;
3. public 4-year non-doctorate-granting bachelor's high education;
4. public 4-year non-doctorate-granting bachelor's low education;
5. public 4-year non-doctorate-granting master's high education;
6. public 4-year non-doctorate-granting master's low education;

7. public 4-year doctorate-granting high education;
8. public 4-year doctorate-granting low education;
9. public 4-year first-professional-granting high education;
10. public 4-year first-professional-granting low education;
11. private not-for-profit less-than-2-year;
12. private not-for-profit 2-year;
13. private not-for-profit 4-year non-doctorate-granting bachelor's high education;
14. private not-for-profit 4-year non-doctorate-granting bachelor's low education;
15. private not-for-profit 4-year non-doctorate-granting master's high education;
16. private not-for-profit 4-year non-doctorate-granting master's low education;
17. private not-for-profit 4-year doctorate-granting high education;
18. private not-for-profit 4-year doctorate-granting low education;
19. private not-for-profit 4-year first-professional-granting high education;
20. private not-for-profit 4-year first-professional-granting low education;
21. private for-profit less-than-2-year;
22. private for-profit 2-year or more;
23. California public 2-year;
24. California public 4-year;
25. California private not-for-profit 4-year;
26. California private for-profit degree-granting;
27. Texas public 2-year;
28. Texas public 4-year;
29. Texas private not-for-profit 4-year;
30. Texas private for-profit degree-granting;
31. New York public 2-year;
32. New York public 4-year;
33. New York private not-for-profit 4-year;
34. New York private for-profit degree-granting;
35. Illinois public 2-year;
36. Illinois public 4-year;
37. Illinois private not-for-profit 4-year;
38. Illinois private for-profit degree-granting;
39. Georgia public 2-year;
40. Georgia public 4-year;
41. Georgia private not-for-profit 4-year;
42. Georgia private for-profit degree-granting;
43. Minnesota public 2-year;
44. Minnesota public 4-year;

45. Minnesota private not-for-profit 4-year; and
46. Minnesota private for-profit degree-granting.

Note that “high education” refers to the 20 percent of institutions with the highest proportions of their baccalaureate degrees awarded in education (based on the most recent IPEDS Completions file). The remaining 80 percent of institutions are classified as “low education” (i.e., having a lower proportion of baccalaureate degrees awarded in education).

c. Student Sample

Based on the expected response and eligibility rates, the preliminary expected student sample sizes and sample yield are presented in table 8. This table shows that the full-scale study will be designed to sample a total of 138,066 students, including 29,428 baccalaureate recipients; 86,274 other undergraduate students; and 22,364 graduate and first-professional students. Based on past experience, we expect to obtain, minimally, an overall eligibility rate of 92.0 percent and an overall student interview response rate of 70.0 percent; however, these rates will vary by sector.

Table 8. NPSAS:08 preliminary full-scale student sample sizes and yield

| Institutional sector | Sample students | | | | Eligible students | | | | Study respondents | | | | Responding students per responding institution |
|--|-----------------|----------------|------------------------------|--------------------------------------|-------------------|----------------|------------------------------|--------------------------------------|-------------------|----------------|------------------------------|--------------------------------------|--|
| | Total | Baccalaureates | Other undergraduate students | Graduate/first-professional students | Total | Baccalaureates | Other undergraduate students | Graduate/first-professional students | Total | Baccalaureates | Other undergraduate students | Graduate/first-professional students | |
| Total | 138,066 | 29,428 | 86,274 | 22,364 | 127,073 | 27,827 | 78,026 | 21,220 | 113,178 | 25,567 | 68,110 | 19,501 | 70 |
| Public less-than-2-year | 3,409 | 0 | 3,409 | 0 | 2,719 | 0 | 2,719 | 0 | 2,238 | 0 | 2,238 | 0 | 155 |
| Public 2-year | 31,095 | 0 | 31,095 | 0 | 27,330 | 0 | 27,330 | 0 | 21,719 | 0 | 21,719 | 0 | 57 |
| Public 4-year non-doctoral | 16,592 | 5,722 | 8,710 | 2,153 | 15,739 | 5,430 | 8,266 | 2,043 | 14,139 | 4,878 | 7,425 | 1,835 | 83 |
| Public 4-year doctoral | 37,456 | 12,164 | 14,683 | 10,579 | 35,595 | 11,569 | 13,965 | 10,062 | 32,602 | 10,596 | 12,791 | 9,216 | 130 |
| Private not-for-profit less-than-4-year | 3,077 | 0 | 3,077 | 0 | 2,739 | 0 | 2,739 | 0 | 2,524 | 0 | 2,524 | 0 | 142 |
| Private not-for-profit 4-year non-doctoral | 12,577 | 4,752 | 6,065 | 1,734 | 11,783 | 4,461 | 5,694 | 1,628 | 11,091 | 4,199 | 5,360 | 1,532 | 39 |
| Private not-for-profit 4-year doctoral | 15,784 | 4,080 | 4,236 | 7,486 | 15,005 | 3,874 | 4,022 | 7,108 | 13,860 | 3,579 | 3,715 | 6,566 | 66 |
| Private for-profit less-than-2-year | 7,391 | 0 | 7,391 | 0 | 6,295 | 0 | 6,295 | 0 | 5,839 | 0 | 5,839 | 0 | 76 |
| Private for-profit 2-year or more | 10,679 | 2,710 | 7,608 | 412 | 9,868 | 2,492 | 6,997 | 379 | 9,164 | 2,314 | 6,497 | 352 | 42 |

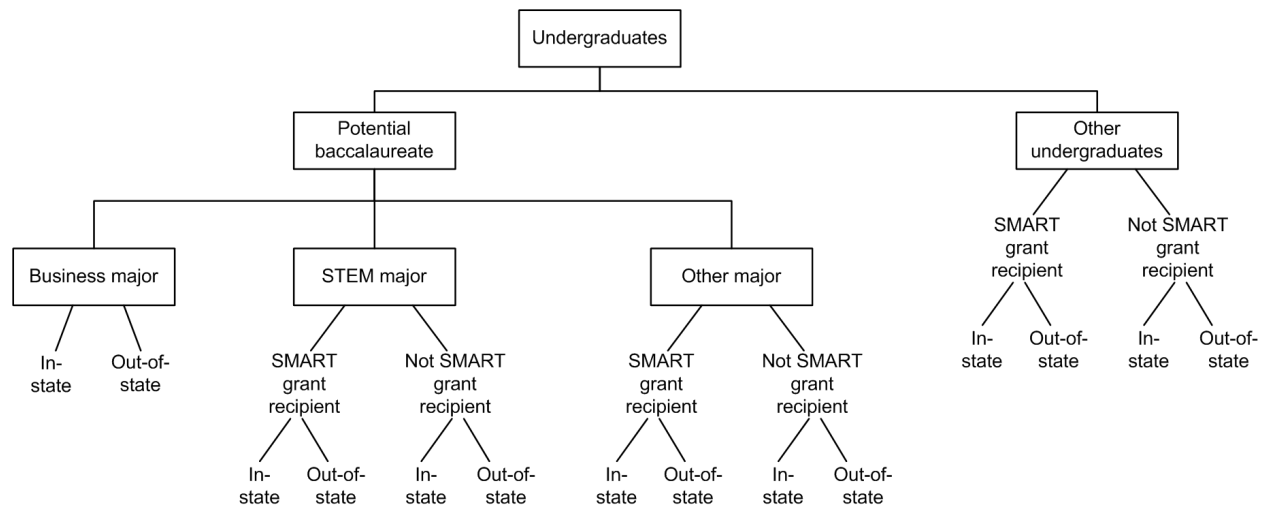
NOTE: NPSAS:08 = 2008 National Postsecondary Student Aid Study.

We plan to employ a variable-based (rather than source-based) definition of study respondent, similar to that used in the NPSAS:08 field test and in NPSAS:04. There are multiple sources of data obtained as part of the NPSAS study, and study respondents must meet minimum data requirements, regardless of source. Using the same variable-based definition from the field test, we expect the overall study response rate to be 89.1 percent, based on NPSAS:04 results. We anticipate, however, that study response rates will vary by institutional sector, as was the case in NPSAS:04. Using the rates we experienced in that study, we expect approximately 113,178 study respondents, including 25,567 baccalaureate recipients; 68,110 other undergraduate students; and 19,501 graduate and first-professional students.

The 18 student sampling strata are listed below and shown graphically in figure 1:

1. in-state potential baccalaureate recipients who are business majors;
2. out-of state potential baccalaureate recipients who are business majors;
3. in-state potential baccalaureate recipients who are science, technology, engineering, or mathematics (STEM) majors and SMART grant recipients;
4. out-of-state potential baccalaureate recipients who are STEM majors and SMART grant recipients;
5. in-state potential baccalaureate recipients who are STEM majors and not SMART grant recipients;
6. out-of-state potential baccalaureate recipients who are STEM majors and not SMART grant recipients;
7. in-state potential baccalaureate recipients in all other majors who are SMART grant recipients;
8. out-of state potential baccalaureate recipients in all other majors who are SMART grant recipients;
9. in-state potential baccalaureate recipients in all other majors who are not SMART grant recipients;
10. out-of state potential baccalaureate recipients in all other majors who are not SMART grant recipients;
11. in-state other undergraduate students who are SMART grant recipients;
12. out-of-state other undergraduate students who are SMART grant recipients;
13. in-state other undergraduate students who are not SMART grant recipients;
14. out-of-state other undergraduate students who are not SMART grant recipients;
15. masters students;
16. doctoral students;
17. other graduate students; and
18. first-professional students.

Figure1. NPSAS:08 undergraduate student sampling strata



As was done in NPSAS:2000 and NPSAS:04, certain student types (potential baccalaureate recipients, other undergraduates, masters students, doctoral students, other graduate students, and first-professional students) will be sampled at different rates to control the sample allocation. Differential sampling rates facilitate obtaining the target sample sizes necessary to meet analytic objectives for defined domain estimates in the full-scale study.

To ensure a large enough sample for the B&B follow-up, the base year sample includes a large percentage of potential baccalaureate recipients (see table 8). The sampling rates for students identified as potential baccalaureates and other undergraduate students on enrollment lists will be adjusted to yield the appropriate sample sizes after accounting for the baccalaureate “false-positives.” This will ensure sufficient numbers of actual baccalaureate recipients. The expected “false positive” rate will be based on the results of the NPSAS:08 field test, comparing B&B status across several sources, and on NPSAS:2000 full scale survey data.⁶

RTI will receive a file of SMART grant recipients from ED and will match that list to each institution’s enrollment list to identify and stratify such students. SMART grant recipients are required to major in a STEM field or in certain foreign languages, so baccalaureate recipients who are STEM or other majors must also be stratified by SMART grant recipient status. However, the strata for baccalaureate recipients who are business majors does not need to be stratified by SMART grant recipient status since they are not eligible to receive the grant because of their major.

Creating Student Sampling Frames. Several alternatives for the types of student enrollment lists that can be provided by the sample institutions are available. Our first preference is to obtain an unduplicated list of all students enrolled in the specified time frame. However, lists by term of enrollment and/or by type of student (e.g., baccalaureate recipient, undergraduate, graduate, and first-professional) will be accepted. The student ID numbers can be used to easily unduplicate electronic files. If an institution has difficulty meeting these requirements, we will be flexible and select the student sample from whatever type of list(s) that the institution can provide, so long as it appears to accurately reflect enrollment during the

⁶ In NPSAS:2000, the “false-positive” rate was 13 percent, but lists were usually sent closer to the end of the spring term than they will be in NPSAS:08, so this rate may be a low estimate for NPSAS:08.

specified terms of instruction. If necessary, we are even prepared to provide institutions with specifications to allow them to select their own sample.

In prior NPSAS studies that spun off a B&B cohort, lists of potential baccalaureate recipients were collected with the student list of all enrolled undergraduates and graduates/first professionals. Unfortunately, these baccalaureate lists often could not be provided until late in the spring or in the summer, after baccalaureate recipients could be positively identified. To help facilitate earlier receipt of lists, we will request that the enrollment lists for 4-year institutions include an indicator of class level for undergraduates (1st year, 2nd year, 3rd year, 4th year, or 5th year). From NPSAS:2000, we estimate that about 55 percent of the 4th- and 5th-year students will be baccalaureate recipients during the NPSAS year, and about 7 percent of 3rd-year students will also be baccalaureate recipients. To increase the likelihood of correctly identifying baccalaureate recipients, we will also request that the enrollment lists for 4-year institutions include an indicator (B&B flag) of students who have received or are expected to receive a baccalaureate degree during the NPSAS year (yes, no, don't know). We will instruct institutions to make this identification before spring graduation so as not to hold up the lists because of this requirement. These two indicators will be used instead of requesting a baccalaureate recipient list, and we plan to oversample 4th and 5th year undergraduates (seniors) and students with a B&B flag of "yes" to ensure obtaining sufficient yield of baccalaureate recipients for the B&B longitudinal study. We expect that most institutions will be able to provide undergraduate year for their students and a B&B flag.

We will also request major field of study and Classification of Instructional Programs (CIP) code on the lists to allow us to undersample business majors and to oversample STEM majors. A similar procedure was used effectively in NPSAS:2000 (the last NPSAS to include a B&B cohort). We expect that most institutions can and will provide the CIP codes. Undersampling business majors is necessary because a disproportionately large proportion of baccalaureate recipients are business majors, and oversampling STEM majors is necessary because there is an emerging longitudinal analytic interest in baccalaureate recipients in these fields.

The following additional data items will be requested for all NPSAS-eligible students enrolled at each sample institution:

- name;
- Date of birth (DOB);
- Social Security number (SSN);
- student ID number (if different from SSN);
- student level (undergraduate, masters, doctoral, other graduate, first-professional);
and
- locating information (local and permanent street address and phone number and school and home e-mail address).

Permanent address will be used to identify and oversample undergraduate in-state students. A similar procedure was used effectively in NPSAS:04. Oversampling of in-state students in the six states with representative samples is necessary because state-level analyses typically only include in-state students, so sufficient sample size is needed. In the other states,

the undergraduate students will be stratified by in-state and out-of-state for operational efficiency, but in-state students will not be oversampled.

As part of initial sampling activities, we will ask participating institutions to provide SSN and DOB for all students on their enrollment list.⁷ We recognize the sensitivity of the requested information, and appreciate the argument that it should be obtained only for sample members. However, collecting this information for all enrolled students is critical to the success of the study for several reasons:

- It is possible that some minors will be included in the study population, so we will need to collect DOB to identify minors and obtain parental consent prior to data collection.
- The NPSAS:08 study includes a special analytic focus on a new federal grant (the National SMART grant) and SSN is needed to identify and oversample recipients of this new grant.
- Having SSN will ensure the accuracy of the sample, because it is used as the unique student identification number by most institutions. We need to ensure that we get the right data records when collecting data from institutions for sampled students. It will also be used to unduplicate the sample for students who attend multiple institutions.
- Making one initial data request of institutions will minimize the burden required for participation (rather than obtaining one set of information for all enrolled students, and then later obtaining a set of information for sampled students).
- An issue related to institutional burden is institutional participation. It is very likely that some institutions will respond to the first request, but not to the second. Refusal to provide SSNs after the sample members are selected will contribute dramatically to student-level nonresponse, because it will increase the rate of unlocatable students (see the following bullet).
- Obtaining SSN early will allow us to initiate locating and file matching procedures early enough to ensure that data collection can be completed within the allotted schedule. The data collection schedule would be significantly and negatively impacted if locating activities could not begin at the earliest stages of institutional contact.
- NPSAS data are critical for informing policy and legislation, and are needed by Congress in a timely fashion. Thus, the data collection schedule is also critical. We must be able to identify the sample, locate students, and finish data collection and data processing quickly. This will not be possible within the allotted time frame if we are unable to initiate locating activities for sampled students once the sample has been selected.

The following section describes our planned procedures to securely obtain, store, and discard sensitive information collected for sampling purposes.

⁷ For institutions unwilling to provide SSN or location data for all students on enrollment lists, we will request SSN or locating data only for sample students immediately after the sample is selected.

Obtaining student enrollment lists. The student sample will be selected from the lists provided by the sampled institutions. To ensure the secure transmission of sensitive information, we will provide the following options to institutions: (1) upload encrypted student enrollment list files to the project’s secure website using a login ID and “strong” password provided by RTI, or (2) provide an appropriately encrypted list file via e-mail (RTI will provide guidelines on encryption and creating “strong” passwords).

In past administrations of this study, hard copy lists were accepted via Fed-Ex or fax. We did not offer this option in the field test and will not offer it in the full-scale study. We expect that a very few institutions will ask to provide a hard copy list (in NPSAS:04 full-scale study, 30 institutions submitted a hard-copy list—mostly via FedEx). In such cases, we will encourage one of the secure electronic methods of transmission. If that is not possible, we will accept a faxed list (but not a Fed-Ex list.) Although fax equipment and software does facilitate rapid transmission of information, this same equipment and software opens up the possibility that information could be misdirected or intercepted by individuals to whom access is not intended or authorized. To safeguard against this, as much as is practical, RTI protocol will only allow for lists to be faxed to a fax machine housed in a locked room and only if schools cannot use one of the other options. To ensure the fax transmission is sent to the appropriate destination, we will require a test run with nonsensitive data prior to submitting the actual list to eliminate errors in transmission from misdialing. RTI will provide schools with a FAX cover page that includes a confidentiality statement to use when transmitting individually identifiable information.⁸ After a sample is selected from an institution, the original electronic or keyed list of all students containing SSNs will be deleted, and faxed lists will be shredded. RTI will ensure that the SSNs for nonselected students are securely discarded (see description below).

Storage of enrollment files.

- Encrypted electronic files sent via e-mail to a secure e-mail folder will be accessible only to a few staff members on the sampling team. These files will then be copied to a project folder that is accessible only to these same staff members. Access to this project folder will be set so that only those who have authorized access will be able to see the included files. The folder will not even be visible to those without access. After being copied, the files will be deleted from the e-mail folder. After selecting the sample of students for each school, the original file containing all students with SSNs will be immediately deleted. While in use, files will be stored on the network that is backed up regularly to avoid the need to recontact the institution to provide the list again should a loss occur. RTI’s information technology service (ITS) will use standard procedures for backing up data, so the backup files will exist for three months.
- Files uploaded to the secure NPSAS website will be copied from the NCES server to the same project folder mentioned above. After being moved, the files will be immediately deleted from the NCES server. After selecting the sample of students for each school, the original file containing all students with SSNs will be immediately deleted. As above, it is necessary for the files to be stored on the project share so that they can be backed up by ITS in case any problems occur that cause us to lose data. ITS will use their standard procedures for backing up data, so the backup files will exist for 3 months.

⁸ These procedures are consistent with those endorsed by HIPAA. See <http://www.hipaadvisory.com/action/faxfacts.htm>

- Paper lists will be kept in one locked file cabinet. Only NPSAS sampling staff will have access to the file cabinet. The paper lists will be shredded immediately after the sample is selected, keyed, and QC'ed. The keying will be done by the same sampling staff who select the sample.

Selection of Sample Students. The unduplicated number of enrollees on each institution's enrollment list will be checked against the latest IPEDS unduplicated enrollment data, which are part of the spring web-based IPEDS data collection. For electronic files, lists will be unduplicated by student ID number. For faxed lists, which are expected to be small, the total number of students listed will be counted. The comparisons will be made for baccalaureates and for each student level: undergraduate, graduate, and first-professional. Based on past experience only counts within 25 percent of nonimputed IPEDs counts will pass edit. There will be one exception based on field test results: if the baccalaureate count is higher than the IPEDs count but within 50 percent, the count will pass edit because we are comparing potential baccalaureate list counts with actual IPEDs counts.

Institutions that fail edit will be recontacted to resolve the discrepancy and to verify that the institution coordinator who prepared the student lists clearly understood our request and provided a list of the appropriate students. When we determine that the initial list provided by the institution was not satisfactory, we will request a replacement list. We will proceed with selecting sample students when we either have confirmed that the list received is correct or have received a corrected list.

Electronic lists will be unduplicated by student ID number prior to sample selection. In addition, all samples, both those selected from electronic files and from paper lists, will be unduplicated by SSN *between* institutions. The duplicate sample member will be deleted from the second institution because the sample is selected on a flow basis. In prior NPSAS studies, we found several instances in which this check avoided multiple selections of the same student. However, we also learned that the ID numbers assigned to noncitizens may not be unique across institutions; thus when duplicate IDs are detected but the IDs are not standard SSNs (do not satisfy the appropriate range check), we will check the student names to verify that they are indeed duplicates before deleting the students.

Student names and SSNs or student IDs will be keyed into Excel for faxed lists, which are expected to be short lists. The keying will be checked thoroughly. These keyed lists will then be unduplicated and sampled similarly to electronic lists. After the sample is selected from a keyed list, the additional information from the original faxed list will be keyed just for the sampled students and checked carefully.

Stratified systematic samples of students will be selected, from both electronic and faxed student lists,⁹ on a flow basis as the lists are received by adapting the procedures we have used successfully for student sampling in prior NPSAS rounds. As the student samples are selected they will be added to a master sample file containing, minimally, for each sample student: a unique study ID number (NPSASID), SSN, the institution's IPEDS ID number (UNITID), institutional stratum, student stratum, and selection probability.¹⁰ Sample yield will be monitored by institutional and student sampling strata, and the sampling rates will be adjusted early, if necessary, to achieve the desired full-scale sample yields.

⁹ Based on NPSAS:04 and the field test, we expect that 2 percent or less of the enrollment lists received will be paper list.

¹⁰ The selection probability is based on the unduplicated list.

Quality Control Checks for Sampling. All statistical procedures will undergo thorough quality control checks. We have technical operating procedures (TOPs) in place for sampling and general programming. These TOPs describe how to properly implement statistical procedures and QC checks. We will use a checklist for all statisticians to use to make sure that all appropriate QC checks are done.

Some specific sampling QC checks will include, but are not limited to, checking that

- the students on the sampling frames all have a known, non-zero probability of selection; and
- the number of students selected match the target sample sizes.

3. Methods for Maximizing Response Rates

Response rates in NPSAS:08 are a function of success in two basic activities: identifying and locating the sample members involved, then contacting them and gaining their cooperation. Two classes of respondents are involved: institutions, and students (undergraduate, graduate, and first-professionals) who were enrolled in those institutions.

a. Institution Contacting

The success of NPSAS:08 is closely tied to the active participation of selected institutions. Because institution contacting is the first stage of the study, upon which all other stages depend, obtaining the cooperation of as many institutions as possible is critical. The consent and cooperation of an institution's chief administrator is essential and helps to encourage the timely completion of the institutional tasks. Most chief administrators are aware of NPSAS and recognize the study's importance to postsecondary education. For those administrators who may believe that the study is overly burdensome, the first contact provides an opportunity to have a senior staff member address their concerns. At institutions newly selected for participation in NPSAS:08, the chief administrator contact provides an invaluable opportunity to establish rapport.

Proven Procedures. NPSAS:08 procedures will be developed from those used successfully in NPSAS:04. Initial institution contact information will be obtained from the IPEDS-IC file and used to telephone each institution (to verify data of record—e.g., the institution's name, address, and telephone number and the name and address of the chief administrator). Verification calls will begin in September 2007 and last approximately 1 week. Materials will be mailed to chief administrators in late September 2007, with follow-up calls continuing through early November. This schedule follows the model implemented in 2004 that established contact with the coordinator prior to the holiday season. The descriptive materials sent to chief administrators will be clear, concise, and informative about the purpose of the study and the nature of subsequent requests. The package of materials sent to chief administrators will contain

- an introductory letter from the NCES Commissioner on U.S. Department of Education letterhead;
- a pamphlet describing NPSAS:08, including a study summary, outline of the data collection procedures, the project schedule, and details regarding the protection of respondent privacy and study confidentiality procedures; and

- a form confirming the institution's willingness to participate in the study, identifying an Institution Coordinator, and requesting contact information for the chief administrator and the institution coordinator.

Follow-up calls to secure field test participation and name a study coordinator occur after allowing adequate time for materials to reach the chief administrators. Identified coordinators will receive a package containing duplicates of materials sent to the chief administrators plus materials clearly explaining the coordinator's critical role in gaining access, consideration, and participation from staff within their institution. Also provided will be checklists clearly describing the steps of the data collection process and anticipated levels of effort.

Experienced staff from RTI's Call Center Services (CCS) carry out these contacts and are assigned a set of institutions that is their responsibility throughout the process. This allows RTI staff members to establish rapport with the institution staff and provides a reliable point of contact at RTI. Staff members are thoroughly trained in basic financial aid concepts and in the purposes and requirements of the study, which helps them establish credibility with the institution staff.

Endorsements. In previous NPSAS studies, the specific endorsement of relevant associations was extremely useful in persuading institutions to cooperate. Endorsements from 26 professional associations have been secured for NPSAS:08. These associations are listed in appendix F. In addition to providing general study endorsement, the National Association of Student Financial Aid Administrators (NASFAA) promotes the study at its national and regional meetings and through the association's publications.

Minimizing Burden. As in prior NPSAS studies, different options for providing enrollment lists and for extracting/recording the data requested for sampled students are offered. The coordinator is invited to select the methodology of greatest convenience to the institution. The optional strategies for obtaining the data are discussed later in this section. With regard to student record abstractions, "preloading" a customized list of financial aid awards into the computer assisted data entry (CADE) for each institution reduces the amount of data entry required for the institution and more closely tailors CADE to award names likely to be found in students' financial aid records. During institution contacting, the names of up to four of the most commonly awarded institution grants and scholarships are identified to assist in this process. Data on institution attributes such as institution level and control, highest level of offering, and other attributes are verified and updated as well.

b. Institutional Data Collection Training

Institution Coordinator Training. The purpose of an effective plan for training institution coordinators is two-fold: to make certain that survey procedures are understood and followed, and to motivate the coordinators. The project relies on these procedures to ensure institutional data are recorded accurately and completely. Because institution coordinators are a critical element in this process, communicating instructions about their survey tasks clearly is essential.

Institution coordinators will be trained during the course of telephone contacts by call center staff. Written materials will be provided to coordinators explaining each phase of the project (enrollment list acquisition, student sampling, institution data abstraction, etc.) as well as their role in each.

Training of institution coordinators is geared toward the method of data collection selected by the institution. All institution coordinators will be informed about the purposes of NPSAS, provided with descriptions of their survey tasks, and assured of our commitment to maintaining the confidentiality of institution, student, and parent data. The CADE system is a Web application; and the CADE website, accessible only with an ID and password, provides institution coordinators with instructions for all phases of study participation. Copies of all written materials, as well as answers to frequently asked questions, are available on the website.

In addition to the training activities described above, RTI established an exhibit booth at NASFAA's national conference in July of 2007. Attending this conference allows project management to meet staff from institutions who have previously participated, and to answer any questions regarding the study, CADE, or institution burden. Because the date of this conference coincided with the completion of field test institution data collection activities, we were also able to solicit feedback from financial aid administrators of field test institutions.

Field Data Collector Training. RTI will develop the training plan and training materials for the field-CADE data collectors and make arrangements for the training. One training session will be held, conducted by staff members who will be responsible for management of the institutional records data collection and who are experienced in conducting data collection from educational institutions. The training is designed to ensure that the data collectors are fully prepared to identify problems that may be encountered in working with schools and school records and to apply solutions that will result in the collection of consistently high quality data by all field staff. The training will include

- a thorough explanation of the background, purpose, and design of the survey;
- an overview of the NPSAS institutional records data collection activity and its importance to the success of the study;
- a description of the role of the NPSAS data collector and his/her responsibility for obtaining complete and accurate data;
- an explanation of the role of the institution coordinator and how the data collector will interact with him/her;
- a full explanation of confidentiality and privacy regulations that apply to the data collector, including signing of nondisclosure affidavits;
- procedures for obtaining financial aid data from sample schools that must be visited;
- use of the CADE module and field case management system to collect, manage, and transmit data;
- completion and review of sample exercises simulating the various situations that will be encountered collecting student financial aid data from the various types of institutions included in the sample; and
- communication and reporting procedures.

The NPSAS Field Data Collector Manual will fully address each of the training topics and will describe all field data collection procedures in detail. The manual will be designed to serve as both a training manual and a reference manual for use during actual data collection. Training will emphasize active participation of the trainees and provide extensive opportunities

for them to deal with procedures and the Information Management System (IMS). A major goal is preparing trainees to interact appropriately with the variety of school staff and different types of financial aid administration and record-keeping systems they will encounter at the NPSAS:08 sample institutions.

c. Collection of Student Data from Institutional Records

The highest priority goal for NPSAS:08 reflects its student aid focus. Institutions and federal financial aid databases are the best source for these data. Historically, institutional records have been a major source of student financial aid, enrollment, and locating data for NPSAS. As part of the institution contacting, institution coordinators will be asked to select a method of data collection—self-CADE (CADE completed by the institution via data entry through a secured website), field-CADE (CADE with the assistance of field data collectors), or data-CADE (submission of an electronic data file via a secured website). We have assumed, based on our previous NPSAS experience, that 21 percent of eligible institutions will submit data-CADE—with 13 percent requiring a field data collector and the remaining 66 percent performing the abstraction themselves.

Prior to data collection, student records are matched to the U.S. Department of Education Central Processing System (CPS)—which contains data on federal financial aid applications—for locating purposes and to reduce the burden on the institutions for the student record abstractions. The vast majority of the federal aid applicants (about 95 percent) will match successfully to the CPS prior to CADE data collection, so we will ask the institution to provide the student’s last name and Social Security number for the small number of federal aid applicants who did not match to the CPS prior to CADE. We will collect these two pieces of information in CADE and then submit the new names and Social Security numbers to CPS for file matching after CADE data collection has ended. Any new data obtained for the additional students will be delivered on the Electronic Code Book (ECB) with the data obtained prior to CADE. Under either scenario, we will have reduced the level of effort at the institution and thereby reduced the CADE cycle time.

Self-CADE via the Internet. Goals for NPSAS:08 CADE include reducing the data collection burden on NPSAS institutions (thereby reducing project costs by reducing the need for field data collectors), expediting data delivery, improving data quality, and ultimately ensuring the long-term success of NPSAS. NPSAS:2000 demonstrated the viability of a web-based approach to CADE data collection, and NPSAS:04 saw increased use of data-CADE, particularly by institutional systems. We plan to use a self-CADE instrument nearly identical to that used in NPSAS:04.

We had success with the self-CADE instrument in NPSAS:04 and believe more institutions are becoming accustomed to web applications, which will result in significant data collection schedule efficiencies. Under self-CADE, the NPSAS schedule will further benefit from the fact that multiple offices within the institution can enter data into CADE simultaneously, as successfully demonstrated in NPSAS:2000 and NPSAS:04.

Because the open Internet is not conducive to transmitting confidential data, any internet-based data collection effort necessarily raises the question of security. However, we intend to incorporate the latest technology systems into our web-CADE¹¹ application to ensure strict

¹¹ To be used for both self-CADE and data-CADE submissions.

adherence to NCES confidentiality guidelines. Our web server will include a Secure Sockets Layer (SSL) Certificate, resulting in encrypted data transmission over the Internet. The SSL technology is most commonly deployed and recognizable in electronic commerce applications that alert users when they are entering a secure server environment, thereby protecting credit card numbers and other private information. Also, all of the data entry modules on this site are password protected, requiring the user to log in to the site before accessing confidential data. The system automatically logs the user out after 20 minutes of inactivity. This safeguard prevents an unauthorized user from browsing through the site. Additionally, we will stay attuned to technological advances to ensure the NPSAS:08 data are completely secure.

Data-CADE. Our CADE experience in NPSAS:2000 and NPSAS:04 confirmed that some coordinators prefer submitting files containing the institution data, rather than performing data entry into CADE. Allowing the institutions to submit CADE data in the form of a data file (via upload to the project's secure website) provides a more convenient mechanism by which institutions can provide data electronically (without performing data entry). Detailed specifications will be provided to the institutions that request this method. We will contact the institution to discuss thoroughly the content of the file and to clarify the exact specification requirements. To mitigate the costs of RTI programmers processing files in various formats, we will request that institutions providing CADE data files use the .CSV format.

Security for the CADE data files will be the same as that described above for self-CADE. File transmission via the website will be protected by industry-standard SSL encryption technology.

Field-CADE. Field data collectors will conduct data abstractions at institutions not choosing self-administered CADE. The data collectors will arrange their visit to the institution with the coordinator and, once there, will abstract data from student records and key the data into CADE software using an RTI-provided laptop computer. The field-CADE data collection system will be identical to the self-CADE instrument but will run in local mode on the laptop, enabling the field data collector to enter the data without needing access to a data line at the institution.

Field data collectors will use a CADE procedures checklist to help them conduct discussions with the coordinator and perform all necessary tasks. The data collector will be provided with electronic files containing CADE preload information for all sampled students. When records abstraction is completed, the data collector will transmit a completed CADE file to RTI.

Data security will be of primary importance during field-CADE data collection. The following steps will be taken to ensure the protection of confidential information in the field.

Field laptops will be encrypted using a whole-disk encryption software package, Pointsec. Pointsec encrypts the entire disk sector by sector, including the system files, temp files, and even deleted files. Boot protection authenticates users before the computer is booted, this prevents the operating system from being subverted by unauthorized persons.

- Field laptops will be configured so that during the startup a warning screen will appear, stating that the computer is the property of RTI and that criminal penalties apply to any unauthorized persons accessing the data on the laptop. The user must acknowledge this warning screen before startup will complete. Each laptop will have affixed a printed version of the same warning with a toll-free number to call if the

laptop is found. Laptops will be configured to require a login and password at startup, and the case management system software will require an additional login and password before displaying the first menu. Field staff are instructed never to write down the passwords anywhere.

- To reduce the risk of intrusion should a laptop be obtained by an unauthorized person, communications software on field laptops will be configured to connect to RTI's network for data transfer (described in the paragraph below). The SQL server database used for data transfer will contain only case assignment and status data, including name and locating information; survey response data will be retrieved from the laptops and stored in a restricted project share. Completed cases' data files will be removed from the laptop during transmission after the data have been verified as being received at RTI.
- Data being sent to and from field laptops are stored in a domain of the RTI network that is behind the RTI firewall but allows access, with appropriate credentials, to users accessing RTI resources while physically outside the private domain (the innermost security login level accessible only by internal RTI staff). The particular file share in which the ingoing and outgoing data are housed is protected by NT security, which allows access to the data only by RTI system administrators, field system programmers, and the controlled programs that are invoked when field interviewers' laptops connect via direct dialup to RTI's modems and communicate with the Integrated Field Management System (IFMS).

CADE Quality Control. As part of our quality control procedures, we will emphasize to CADE data abstractors the importance of collecting information for items. Items will not only have edit-checks applied to them during the CADE abstraction, they will also be analyzed by CADE when abstraction for a student is complete for a given section of the instrument. This CADE feature indicates which key items are missing or out of range and will provide both field data collectors and institution staff with an indication of the overall quality of their abstraction efforts.

As data are collected at institutions, either by field data collectors or institution staff, they will ultimately reside on the Integrated Management System (IMS). In the case of self-CADE institutions, the data will already be resident on the RTI web server and will be copied directly into a special CADE subdirectory of the IMS. Web-based CADE will also allow improved quality control over the CADE process, as RTI central staff will be able to monitor data quality for participating schools closely and on a regular basis. When CADE institutions call for technical or substantive support, we will be able to query the institution's data and communicate much more effectively regarding any problems.

In the case of field-CADE institutions, the CASES files will be transmitted electronically from their modem-equipped laptop computers to the same location. From this subdirectory, automated quality control software, running nightly, will read the data files that arrived that day and produce quality control reports. These reports will summarize the completeness of the institution data and make comparisons to all other participating institutions, as well as to similar (i.e., same Level and Control) institutions.

d. Student Locating

Student interviews and student institutional record abstraction will occur simultaneously so that schedule requirements are met. To achieve the desired response rate, we will use a tracing approach that consists of up to four steps designed to yield the maximum number of locates with the least expense. The steps of our tracing plan include the following elements:

- **Tracing prior to the start of data collection.** Our advance tracing operation will involve batch database searches and interactive database searches.
- **Lead letter mailings to sample members.** A personalized letter (signed by an NCES official), study leaflet, and information sheet will be mailed to all sample members to initiate data collection. This letter will include a toll-free 800 number, study website address, and study ID and password, and will request that sample members call to schedule an appointment to complete the interview by telephone, or complete the self-administered interview. One week after the lead letter mailing, a thank you/reminder postcard will be sent to sample members.
- **Intermediate tracing (during CATI but before intensive tracing).** Cases are processed in batches through Accurint for address and telephone updates. All new information is loaded into our CATI system for attempts to contact the sample members. Cases for which no new information is returned are forwarded to Call Center Services (CCS) tracing services.
- **Intensive tracing.** The goal of intensive tracing is to obtain a telephone number where a CATI interviewer can reach the sample member in a cost-effective manner. Tracing procedures may include (1) checking Directory Assistance for telephone listings at various addresses; (2) using criss-cross directories to obtain the names and telephone numbers of neighbors and calling them; (3) calling persons with the same unusual surname in small towns or rural areas to see if they are related to or know the sample member; and (4) contacting the current or last known residential sources such as neighbors, landlords, and current residents at the last known address. Other more intensive tracing activities could include (1) database checks for sample members, parents, and other contact persons, (2) credit database and insurance database searches, (3) drivers' license searches through the appropriate state departments of motor vehicles, and (4) calls to alumni offices and associations.

e. Student Data Collection: Self-Administered Web and CATI

Training Procedures. Training programs for those involved in survey data collection are critical quality control elements. Training for the help desk operators who answer questions for the self-administered web-based student interview and CATI telephone interviewers will be conducted by a training team with extensive experience. We will establish thorough selection criteria for help desk operators and telephone interviewers to ensure that only highly capable persons—those with exceptional computer, problem-solving, and communication skills—are selected to serve on the project and will contribute to the quality of the NPSAS data.

Contractor staff with extensive experience in training interviewers will prepare the NPSAS:08 Student Survey Telephone Interviewer Manual, which will provide detailed coverage of the background and purpose of NPSAS, sample design, questionnaire, and procedures for the CATI interview. This manual will be used in training and as a reference during interviewing.

(Interview-specific information will be available to interviewers in the Call Center in the form of question-by-question specifications providing explanations of the purpose of each question and any definitions or other details needed to aid the interviewers in obtaining accurate data.) Along with manual preparation, training staff will prepare training exercises, mock interviews (specially constructed to highlight the potential of definitional and response problems), and other training aids.

A comprehensive training guide will also be prepared for use by trainers to standardize training and to ensure that all topics are covered thoroughly. Among the topics to be covered at the telephone interviewer training will be

- the background purposes and design of the survey;
- confidentiality concerns and procedures (interviewers will take an oath and sign an affidavit agreeing to uphold the procedures);
- importance of locating/contacting sample members and procedures for using the IMS/CATI locating and tracing module;
- special practice with online coding systems used to standardize sample member responses to certain items (e.g., institution names);
- review, discussion, and practice of techniques for explaining the study, answering questions asked by sample members, explaining the respondent's role, and obtaining cooperation;
- extensive practice in applying tracing and locating procedures;
- demonstration interviews by the trainers;
- round-robin (interactive mock interviews for each section of each questionnaire, followed by review of the question-by-question specifications for each section);
- completion of classroom exercises;
- practice interviews with trainees using the web/CATI instrument to interview each other while being observed by trainers, followed by discussion of the practice results; and
- explanation of quality control procedures, administrative procedures, and performance standards.

Telephone survey unit supervisors will be given project-specific training in advance of interviewer training and will assist in monitoring interviewer performance during the training.

Student Interviews (web/CATI). Student interviews will be conducted using a single web-based survey instrument for both self-administered and CATI data collection. The data collection activities will be accomplished through the Case Management System (CMS), which is equipped with the following capabilities:

- online access to locating information and histories of locating efforts for each case;
- state-of-the-art questionnaire administration module with full "front-end cleaning" capabilities (i.e., editing as information is obtained from respondents);
- sample management module for tracking case progress and status; and

- automated scheduling module which delivers cases to interviewers and incorporates the following features:
 - *Automatic delivery of appointment and call-back cases at specified times.* This reduces the need for tracking appointments and helps ensure the interviewer is punctual. The scheduler automatically calculates the delivery time of the case in reference to the appropriate time zone.
 - *Sorting of nonappointment cases according to parameters and priorities set by project staff.* For instance, priorities may be set to give first preference to cases within certain sub-samples or geographic areas; cases may be sorted to establish priorities between cases of differing status. Furthermore, the historic pattern of calling outcomes may be used to set priorities (e.g., cases with more than a certain number of unsuccessful attempts during a given time of day may be passed over until the next time period). These parameters ensure that cases are delivered to interviewers in a consistent manner according to specified project priorities.
 - *Restriction on allowable interviewers.* Groups of cases (or individual cases) may be designated for delivery to specific interviewers or groups of interviewers. This feature is most commonly used in filtering refusal cases, locating problems, or foreign language cases to specific interviewers with specialized skills.
 - *Complete records of calls and tracking of all previous outcomes.* The scheduler tracks all outcomes for each case, labeling each with type, date, and time. These are easily accessed by the interviewer upon entering the individual case, along with interviewer notes, thereby eliminating the need for a paper record of calls of any kind.
 - *Flagging of problem cases for supervisor action or supervisor review.* For example, refusal cases may be routed to supervisors for decisions about whether and when a refusal letter should be mailed, or whether another interviewer should be assigned.
 - *Complete reporting capabilities.* These include default reports on the aggregate status of cases and custom report generation capabilities.

The integration of these capabilities reduces the number of discrete stages required in data collection and data preparation activities and increases capabilities for immediate error reconciliation, which results in better data quality and reduced cost. Overall, the scheduler provides a highly efficient case assignment and delivery function by reducing supervisory and clerical time, improving execution on the part of interviewers and supervisors by automatically monitoring appointments and call-backs, and reducing variation in implementing survey priorities and objectives.

In addition to the management aspect of data collection, the survey instrument is another component designed to maximize efficiency and yield high-quality data. Below are some of the basic questionnaire administration features of the web-based instrument:

- Based on responses to previous questions, the respondent or interviewer is automatically routed to the next appropriate question, according to predesignated skip patterns.
- The web-based interview automatically inserts “text substitutions” or “text fills” where alternate wording is appropriate depending on the characteristics of the respondent or his/her responses to previous questions.
- The web-based interview can incorporate or preload data about the individual respondent from outside sources (e.g., previous interviews, sample frame files). Such data are often used to drive skip patterns or define text substitutions. In some cases, the information is presented to the respondent for verification or to reconcile inconsistencies.
- With the web/CATI instrument, numerous question-specific probes may be incorporated to explore unusual responses for reconciliation with the respondent, to probe “don’t know” responses as a way of reducing item nonresponse, or to clarify inconsistencies across questions.
- Coding of multilevel variables. An innovative improvement to previous NPSAS data collections, the web-based instrument uses an assisted coding mechanism to code text strings provided by respondents. Drawing from a database of potential codes, the assisted coder derives a list of options from which the interviewer or respondent can choose an appropriate code (or codes if it is a multilevel variable with general, specific, and/or detail components) corresponding to the text string.
- Iterations. When identical sets of questions will be repeated for an unidentified number of entities, such as children, jobs, or schools, the system allows respondents to cycle through these questions multiple times.

In addition to the functional capabilities of the CMS and web instrument described above, our efforts to achieve the desired response rate will include using established procedures proven effective in other large-scale studies we have completed. These include:

- providing multiple response modes, including self-administered and interviewer-administered options;
- offering incentives to encourage response (see incentive structure described below);
- prompting calls initiated prior to the start of data collection to remind sample members about the study and the importance of their participation;
- assigning experienced CATI data collectors who have proven their ability to contact and obtain cooperation from a high proportion of sample members;
- training the interviewers thoroughly on study objectives, study population characteristics, and approaches that will help gain cooperation from sample members;
- providing the interviewing staff with a comprehensive set of questions and answers that will provide encouraging responses to questions that sample members may ask;
- maintaining a high level of monitoring and direct supervision so that interviewers who are experiencing low cooperation rates are identified quickly and corrective action is taken;

- making every reasonable effort to obtain an interview at the initial contact, but allowing respondent flexibility in scheduling appointments to be interviewed;
- providing hesitant respondents with a toll-free number to use to telephone RTI and discuss the study with the project director or other senior project staff; and
- thoroughly reviewing all refusal cases and making special conversion efforts whenever feasible (see next section).

Refusal Aversion and Conversion. Recognizing and avoiding refusals is important to maximize the response rate. We will emphasize this and other topics related to obtaining cooperation during data collector training. Supervisors will monitor interviewers intensely during the early days of data collection and provide retraining as necessary. In addition, the supervisors will review daily interviewer production reports produced by the CATI system to identify and retrain any data collectors who are producing unacceptable numbers of refusals or other problems.

After encountering a refusal, the data collector enters comments into the CMS record. These comments include all pertinent data regarding the refusal situation, including any unusual circumstances and any reasons given by the sample member for refusing. Supervisors will review these comments to determine what action to take with each refusal. No refusal or partial interview will be coded as final without supervisory review and approval. In completing the review, the supervisor will consider all available information about the case and will initiate appropriate action.

If a follow-up is clearly inappropriate (e.g., there are extenuating circumstances, such as illness or the sample member firmly requested that no further contact be made), the case will be coded as final and will not be recontacted. If the case appears to be a “soft” refusal, follow-up will be assigned to an interviewer other than the one who received the initial refusal. The case will be assigned to a member of a special refusal conversion team made up of interviewers who have proven especially adept at converting refusals.

Refusal conversion efforts will be delayed for at least one week to give the respondent some time after the initial refusal. Attempts at refusal conversion will not be made with individuals who become verbally aggressive or who threaten to take legal or other action. Refusal conversion efforts will not be conducted to a degree that would constitute harassment. We will respect a sample member’s right to decide not to participate and will not impinge this right by carrying conversion efforts beyond the bounds of propriety.

Incentives to Convert Refusals, Difficult and Unable-to-Locate Respondents. As described in the justification section (section A), we plan to offer incentive payments to nonresponding members of the sample population. We believe there will be three groups of nonrespondents: persons refusing to participate during early response or production interviewing, persons who have proven difficult to interview (i.e., those who repeatedly break appointments with an interviewer), and those who cannot be located or contacted by telephone. Our approach to maximizing the response of these persons—and thereby limiting potential nonresponse bias—involves an incentive payment to reimburse the respondent for time and expenses. The NPSAS:08 field test was used to conduct an experiment to determine whether a \$10 prepaid nonresponse incentive followed by \$20 upon survey completion yielded higher response rates

than the promise of a \$30 incentive. Additional detail about the experiments and their results is provided in section B.4.

Additional Quality Control. In addition to the quality control features inherent in the web-based interview (described in section 3), we will use data collector monitoring as a major quality control measure. Supervisory staff from RTI's Call Center Services (CCS) will monitor the performance of the NPSAS:08 data collectors throughout the data collection period to ensure they are following all data collection procedures and meeting all interviewing standards. In addition, members of the project management staff will monitor a substantial number of interviews. In all cases, students will be informed that the interview may be monitored by supervisory staff.

"Silent" monitoring equipment is used so that neither the data collector nor respondent is aware when an interview is being monitored. This equipment will allow the monitor to listen to the interview and simultaneously see the data entry on a computer screen. The monitoring system allows ready access to any of the work stations in use at any time. The monitoring equipment also enables any of the project managers and client staff at RTI or NCES to dial in and monitor interviews from any location. In the past, we have used this capability to allow the analysts to monitor interviews in progress; as a result, they have been able to provide valuable feedback on specific substantive issues and have gained exposure to qualitative information that has helped their interpretation of the quantitative analyses.

Our standard practice is to monitor 10 percent of the interviewing done by each data collector to ensure that all procedures are implemented as intended and that the procedures are effective, and to observe the utility of the questionnaire items. Any observations that might be useful in subsequent evaluation will be recorded and all such observations will be forwarded to project management staff. Staff monitors will be required to have extensive training and experience in telephone interviewing as well as supervisory experience.

4. Tests of Procedures and Methods in the NPSAS:08 Field Test

Our September 2006 submission to OMB described four tests of procedures and methods that we planned to conduct in the NPSAS:08 field test. These tests pertained to four areas of data collection believed to affect overall study response: (1) presentation of notification materials; (2) reminder prompting; (3) early response incentive offers; and (4) nonresponse conversion incentives.

Based on discussions with OMB, we decided not to conduct the third test, pertaining to use of early response incentives. Because our experience in the BPS:04/06 field test showed that a \$30 early response incentive was effective, we determined that it was not necessary to compare the relative effectiveness of \$10 and \$30 early response incentives.

In this section, we describe the results of the remaining three tests of procedures and methods, and their effect on our plans for full-scale data collection. The tests we conducted allowed us to evaluate the effect of sending notification materials by Priority (vs. regular) mail, of making outbound prompting calls to sample members, and of offering prepaid nonresponse conversion incentives. The tests were also designed to allow us to evaluate the effectiveness of combining two or more of these strategies.

a. Notification Materials

In survey research, the method of mail delivery has been found to be an important factor affecting study response. Our past experience in conducting studies for NCES has also suggested that the look of study materials is important. This is especially true in the NPSAS study, where data collection begins so soon after the student sample is selected that there is not enough time for an “advanced notification” mailing. This scheduling limitation makes it essential for the first contact with students to attract attention.

In the NPSAS:08 field test, we conducted an experiment to test the effectiveness of U.S. Postal Service Priority Mail versus regular mail for the initial mailing to sample members (which includes the introduction to the study and the invitation to participate). Prior to the start of data collection, the field test sample was randomly assigned to two groups: one group received the initial study materials via Priority Mail and the other group received the same materials via regular mail, as had been done in the past. The initial mailing contained important information about the study, including the study brochure and information needed to log into the study website to complete the interview. A toll-free telephone number was provided so sample members could contact the study’s Help Desk for assistance, and could also complete a telephone interview if desired. Finally, the sample member was informed of the details of the incentive offer and the expiration date of the early response period.¹² Results were measured by comparing the response rates at the end of the early response period for these two groups to determine whether response was greater for those who received the Priority Mail.

At the end of the early response period, 38.7 percent of those who received Priority Mail had completed the survey, versus only 33.3 percent of those who received regular mail. This difference of 5.4 percentage points was statistically significant at the <.01 probability level. As a result, we decided that all sample members in the full-scale study would receive Priority Mail instead of regular mail. This method of mailing is relatively inexpensive and easy to manage. We plan to send a Priority Mailing to all of a sample member’s known addresses in order to maximize the likelihood that the individual will receive the material.

b. Prompting

The second experiment conducted in the field test involved the use of prompting calls, or outbound calls made by project staff to sample members. Research has shown that additional contacts with sample members increase the likelihood of participation (Moore and Dillman, 1980). Prompting calls are likely effective because they provide another reminder about a study, give interview staff an additional opportunity to provide the information needed to participate, and provide an early indication of the quality of locating information for a case. Our past experience suggests that prompting calls may have a differential effect on various types of students; e.g., base year respondents vs. nonrespondents in a longitudinal follow-up.

We tested the effectiveness of prompting calls during the early response period in the NPSAS:08 field test by randomly assigning sample members to two groups. One group received telephone prompting calls reminding them to log in to the study website and complete an interview and the other group received no prompting calls. For those in the treatment group, prompting calls occurred approximately 2 weeks into the 3-week early response period.

¹² The “early response period” is defined as the first 3 weeks after the data collection notification is sent to a sample member. Sample members are notified of the study and asked to participate by completing a web-based self-administered survey. Help Desk staff are available during this time period to assist sample members and complete a telephone interview if desired.

Response rates at the end of the early response period for the two groups were compared to determine whether the prompting calls are significantly associated with higher response rates during the early response period. At the end of the early response period, 22.2 percent of sample members who received only prompting calls (and not Priority Mail) had completed the survey, versus 17.5 percent of those who did not receive prompting calls. This difference of 4.7 percentage points was statistically significant at .01.

We also compared response rates among sample members who received *both* Priority Mailing and prompting calls and those who did not receive both. At the end of the early response period, 40.8 percent of those who received both had responded, in comparison to 33.9 percent of those who received neither. This difference of 6.9 percentage points was statistically significant at the .01 level. Thus, the two treatments, when combined, had a significant effect on response rates.

However, the number of cases in the field test sample was not sufficient to support a comparison of the effects of prompting across subgroups.¹³ Because the cost of prompting all sample members in a study as large as NPSAS could be prohibitive, we plan to examine response patterns for the last full-scale administration (NPSAS:04) to identify subgroups of students for whom prompting is expected to be most effective (e.g., those who responded via the web, but not during the early response period, sample members who were classified as refusals, or those with high call counts) In the full-scale study, we will make use of prompting calls for these targeted subgroups only.

c. Nonresponse Conversion Incentives

Another strategy commonly used to obtain sufficient response to survey data collections is the nonresponse conversion incentive. The model used recently for the NCES studies has typically required that a case meet one of the following conditions:

- refusal to participate,
- hard to locate (e.g., have a mailing address, but not a good telephone number), or
- high call count (e.g., >10 or 15).

Once a case has been identified as eligible for the nonresponse conversion incentive, an additional incentive offer is made in an attempt to obtain a completed interview. In previously conducted NCES studies, the nonresponse conversion incentive amount has been high (or at least as high), as the early response incentive amount. The NPSAS:04 full-scale study initially offered a nonresponse conversion incentive of \$20, which did not elicit sufficient response. When the incentive was increased to \$30 toward the end of data collection, there was a dramatic increase in response rates.

Evidence suggests that nonresponse conversion incentives are an effective tool for increasing response rates among sample members who do not respond to early attempts to obtain a completed interview. However, little research has been done to identify the most effective combinations of incentive amount offers over the course of data collection—from early response period to production interviewing¹⁴ through nonresponse conversion.

¹³ Results did suggest, however, that prompting may have had a greater impact for students in less-than-4-year institutions.

¹⁴ Production interviewing is the phase of data collection between the early response period and the nonresponse conversion period, during which outbound computer-assisted telephone interviewing (CATI) occurs.

RTI's experience, particularly in NPSAS:04, has demonstrated that using an early response incentive (e.g., \$10) and a high nonresponse conversion incentive (e.g., \$30) increased response rates. What we had not previously tested was whether prepayment increases response rates more than promised incentive offers. There is much evidence to suggest that prepaid incentives increase response rates more than promised incentives (Dillman, 2000; U.S. Department of Education, 2004; Groves, et. al. 2004) However, prepaid incentives are operationally very difficult to administer, especially with a large sample such as this, that includes many cases that require tracing. To accommodate an evaluation of the impact of prepaid incentives, then, we limited our analysis to the difficult cases at the end of data collection—those determined to be eligible for the nonresponse conversion incentive. This allowed us to assess the impact of prepayment on a reduced scale to determine whether it would be effective to implement for targeted groups in the full-scale sample.

In the NPSAS:08 field test, we compared the effectiveness of a \$10 prepaid nonresponsive conversion incentive, with a promise of an additional \$20 upon completion, to a \$30 incentive upon completion. The response rate among those who received the \$10 prepaid incentive was 34.3 percent, in comparison to a response rate of 34.0 percent among those who did not receive the \$10 prepaid incentive. This difference of 0.3 percentage points was not statistically significant. As a result, we do not plan to use prepaid incentives in the full-scale study. We will investigate the possible utility of prepaid incentives further in other studies; e.g., longitudinal studies where better contact information is available.

5. Reviewing Statisticians and Individuals Responsible for Designing and Conducting the Study

Names of individuals consulted on statistical aspects of study design along with their affiliation and telephone numbers are provided below.

| Name | Affiliation | Telephone |
|---------------------------|-------------|----------------|
| Dr. Lutz Berkner | MPR | (510) 849-4942 |
| Dr. Susan Choy | MPR | (510) 849-4942 |
| Dr. E. Gareth Hoachlander | MPR | (510) 849-4942 |
| Dr. John Riccobono | RTI | (919) 541-7006 |
| Dr. James Chromy | RTI | (919) 541-7019 |
| Dr. Karol Krotki | RTI | (202) 728-2485 |
| Dr. Roy Whitmore | RTI | (919) 541-5809 |
| Mr. Peter Siegel | RTI | (919) 541-6348 |

In addition to these statisticians and survey design experts, the following statisticians at NCES have also reviewed and approved the statistical aspects of the study: Dr. Dennis Carroll, Dr. Thomas Weko, Dr. James Griffith, and Dr. Paula Knepper.

6. Other Contractors' Staff Responsible for Conducting the Study

The study is being conducted by the Postsecondary Longitudinal Studies Branch of the National Center for Education Statistics (NCES), U.S. Department of Education. NCES's prime contractor is the RTI International (RTI). RTI is being assisted through subcontracted activities by MPR Associates and the National Association of Student Financial Aid Administrators

(NASFAA). Principal professional staff of the contractors, not listed above, who are assigned to the study are provided below:

| Name | Affiliation | Telephone |
|-------------------------|-------------|----------------|
| Ms. Mary Ann O'Connor | NASFAA | (202) 785-0453 |
| Ms. Vicky Dinger | MPR | (510) 849-4942 |
| Mr. Tim Gabel | RTI | (919) 541-7415 |
| Dr. Laura Horn | MPR | (510) 849-4942 |
| Mr. Jeff Franklin | RTI | (919) 541-2614 |
| Ms. Christine Rasmussen | RTI | (919) 541-6775 |
| Ms. Melissa Cominole | RTI | (919) 990-8456 |
| Ms. Kristin Dudley | RTI | (919) 541-6855 |
| Mr. Brian Kuhr | RTI | (312) 456-5263 |