

2023-24 NATIONAL POSTSECONDARY STUDENT
AID STUDY (NPSAS:24) FULL-SCALE
STUDY

INSTITUTION CONTACTING AND LIST COLLECTION

Supporting Statement Part B

OMB # 1850-0666 v. 35

Submitted by
National Center for Education Statistics
U.S. Department of Education

July 2023

Contents

B. Collection of Information Employing Statistical Methods.....	3
1. Respondent Universe.....	3
a. Institution Universe.....	3
b. Student Universe.....	3
2. Statistical Methodology.....	4
a. Institution Sample.....	4
b. Student Sample.....	6
3. Methods for Maximizing Response Rates.....	11
a. NPSAS:24 Institution Contacting.....	11
b. Matching to Administrative Databases.....	13
c. Postsecondary Data Portal (PDP).....	13
1. Tests of Procedures or Methods.....	14
2. Reviewing Statisticians and Individuals Responsible for Designing and Conducting the Study.....	16

Tables and Figures

Table 1. Preliminary number of full-scale institutions in the population and to be sampled, by control and level of institution.....	5
Table 2. NPSAS:24 field test predictor variables for the logistic models predicting baccalaureate status.....	9
Table 3. NPSAS:24 preliminary student population and sample sizes, by control and level of institution and student type.....	10
Table 4. Baccalaureate status determination, by student type.....	14
Table 5. Counts and percentages of surveyed students by baccalaureate status.....	15
Table 6. Comparison of NSC and institution enrollment list counts.....	16
Table 7. Percentage of non-missing important data elements.....	16
Figure 1. Institution contacting.....	12

B. Collection of Information Employing Statistical Methods

This submission requests clearance for the 2023–24 National Postsecondary Student Aid Study (NPSAS:24) institution contacting, enrollment list collection, list sampling, and administrative matching activities. The National Center for Education Statistics (NCES) will submit a separate request for the student data collection, including student record data abstraction and student interviews, in the fall of 2023.

1. Respondent Universe

a. Institution Universe

NPSAS:24 will be nationally representative for both undergraduate and graduate students and will use a two-stage sampling design. The first stage involves the selection of institutions. In the second stage, students are selected from within sampled institutions. Also, the NPSAS:24 sample is designed to serve as the base year for a 2024 cohort of the Baccalaureate and Beyond (B&B) Longitudinal Study and, therefore, will include a nationally representative sample of baccalaureate recipients. Although no funding is available to field the follow-up surveys, this allows NCES to add the collections later should Congress appropriate additional funding. To construct the full-scale institution sampling frame for NPSAS:24, we used institution data collected from various surveys of the Integrated Postsecondary Education Data System (IPEDS). The student sampling frame includes all students who meet eligibility requirements from the participating institutions.

The NPSAS:24 institution (first stage) sampling frame includes all levels (less-than-2-year, 2-year, and 4-year) and control classifications (public, private nonprofit, and private for-profit) of Title IV eligible postsecondary institutions in the 50 states, the District of Columbia, and Puerto Rico. To be eligible for NPSAS:24, an institution must do the following during the 2023–24 academic year:

- 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 administer the institution;
- be located in at least one of the 50 states, the District of Columbia, or Puerto Rico;
- be other than a U.S. service academy;¹ and
- have a signed Title IV participation agreement with the U.S. Department of Education.²

As indicated above, institutions providing only avocational, recreational, or remedial courses or only in-house courses for their own employees will be excluded.

b. Student Universe

The student (second stage) sampling frame is described below. NPSAS-eligible undergraduate and graduate students are those who were enrolled in the NPSAS

¹ The U.S. service academies (the U.S. Air Force Academy, the U.S. Coast Guard Academy, the U.S. Military Academy, the U.S. Merchant Marine Academy, and the U.S. Naval Academy) are not eligible for this financial aid study because of their unique funding/tuition base.

² A Title IV eligible institution is an institution that has a written agreement (program participation agreement) with the U.S. Secretary of Education that allows the institution to participate in any of the Title IV federal student financial assistance programs other than the State Student Incentive Grant and the National Early Intervention Scholarship and Partnership programs.

institution in any term or course of instruction between July 1, 2023 and April 30, 2024³ for the full-scale and who are:

- enrolled in either (1) an academic program; (2) at least one course for credit that could be applied toward fulfilling the requirements for an academic degree; (3) exclusively noncredit remedial coursework that has been determined by their institution to be eligible for Title IV aid; or (4) an occupational or vocational program that requires at least 3 months or 300 clock hours of instruction to receive a degree, certificate, or other formal award; and
- not concurrently enrolled in high school; and
- not enrolled solely in a General Educational Development (GED®)⁴ or other high school

completion program.

2. Statistical Methodology

a. Institution Sample

The NPSAS:24 full-scale institution frame was constructed prior to the start of the field test data collection from the IPEDS 2021-22 Institutional Characteristics (IC) header, 2021-22 IC, 2020-21 Completions, and 2020-21 Full-year Enrollment files.⁵ The field test and full-scale institution samples were both selected during the field test. Prior to the start of full-scale data collection, the institution sample will be freshened using the IPEDS:2022-23 header, 2022-23 IC, 2021-22 Completions, and 2021-22 Full-year Enrollment files. This will ensure that all potentially eligible institutions have a probability of selection. The freshened sample size will be determined based on the number and size of new institutions but is expected to be small. As part of this process, we will use the current IPEDS files to identify full-scale institutions no longer eligible for NPSAS:24 due to closure. It is possible that some for-profit institutions and large chains of for-profit institutions may have been closed or sold after the latest IPEDS data collection. Therefore, we will also conduct web searches for information about closed institutions to identify them. For the small number of institutions on the frame that have missing enrollment information because they are not imputed as part of IPEDS, we will impute the enrollment.

The institution strata for the full-scale study will be the eleven sectors, used since the NPSAS:16 full-scale study, which are based on institution level, control, and highest level of offering:

- Public less-than-2-year;
- Public 2-year;
- Public 4-year, non-doctorate-granting, primarily subbaccalaureate;
- Public 4-year, non-doctorate-granting, primarily baccalaureate;
- Public 4-year, doctorate-granting;
- Private nonprofit less-than-4-year;
- Private nonprofit 4-year, non-doctorate-granting;
- Private nonprofit 4-year, doctorate-granting;
- Private for-profit less-than-2-year;
- Private for-profit 2-year; and

³ The end date is March 31, 2024 for continuous enrollment institutions.

⁴ The GED® credential is a high school equivalency credential earned by passing the GED® test, which is administered by GED Testing Service. For more information on the GED test and credential, see https://ged.com/about_test/test_subjects/.

⁵ A preliminary sampling frame has been created using data from the prior year IPEDS files, and population estimates in the sample size tables are based on this preliminary frame. The frame will be recreated with the most up to date IPEDS data prior to sample selection.

- Private for-profit 4-year.

The full-scale sample of 2,000 institutions was selected first using a variation of probability proportional to size (PPS) sampling called sequential probability minimum replacement (PMR) sampling.⁶ This method selects institutions sequentially with probability proportional to size and with minimum replacement. Selection with minimum replacement means that the actual number of hits for an institution can equal the integer part of the expected number of hits for that institution, or the next largest integer, that is, institutions have a chance of being selected more than once.⁷ Instead of the PMR sampling algorithm selecting some institutions multiple times, prior to the PMR sample selection, we set aside for inclusion in the sample with certainty all institutions with a probability of being selected more than once, that is, adjusting their probability of selection to be one. Then, the probabilities of selection for other institutions were adjusted accordingly prior to PMR selection, so that the total institution sample size target was met. A composite size measure⁸ was used to help achieve self-weighting samples⁹ for student-by-institution strata and to allow flexibility to change sampling rates in selected strata without losing the self-weighting attribute of the sampling method. Institution composite measures of size were determined using undergraduate and graduate student enrollment counts and baccalaureate recipient counts from the 2020-21 IPEDS Full-year Enrollment and Completion files.

The institution sample sizes by the eleven institution strata, prior to sample freshening, are presented in table 1. We expect to obtain an overall 97 percent student eligibility rate and at least an overall 80 percent institution participation (response) rate, which will yield approximately 1,550 institutions providing lists.

Within each institution stratum, additional implicit stratification will be accomplished by sorting the sampling frame by the following classifications, as appropriate:

1. Level of institution;
2. Historically Black Colleges and Universities (HBCU) indicator;
3. Hispanic-serving institutions (HSI) indicator;¹⁰
4. Carnegie classification of postsecondary institutions;¹¹ and
5. The institution measure of size.

The objective of this implicit stratification is to approximate proportional representation of institutions on these measures.

Table 1. Number of full-scale institutions in the population and sampled, by control and level of institution¹

Control and level of institution	Population estimate	Sample size
----------------------------------	---------------------	-------------

⁶ Chromy, J.R. (1979). Sequential Sample Selection Methods. In *Proceedings of the Survey Research Methods Section of the American Statistical Association* (pp. 401-406). Alexandria, VA: American Statistical Association.

⁷ https://support.sas.com/documentation/cdl/en/statug/63347/HTML/default/viewer.htm#statug_surveysselect_a0000000173.htm.

⁸ Folsom, R.E., Potter, F.J., and Williams, S.R. (1987). Notes on a Composite Size Measure for Self-Weighting Samples in Multiple Domains. In *Proceedings of the Section on Survey Research Methods of the American Statistical Association*. Alexandria, VA: American Statistical Association, 792-796.

⁹ Self-weighting samples have equal weights within sampling domains.

¹⁰ A Hispanic-serving institutions indicator is no longer available from IPEDS, so we will create an HSI proxy following the definition of HSI as provided by the U.S. Department of Education (<https://www2.ed.gov/programs/idedhs/definition.html>) and using IPEDS Hispanic enrollment data.

¹¹ We used the 2018 version of Carnegie classification available on the IPEDS files.

Total	5,952	2,000
Public less-than-2-year	227	22
Public 2-year	906	377
Public 4-year, non-doctorate-granting, primarily sub-baccalaureate	179	82
Public 4-year, non-doctorate-granting, primarily baccalaureate	198	98
Public 4-year, doctorate-granting	389	338
Private nonprofit 2-year or less	193	20
Private nonprofit 4-year, non-doctorate-granting	867	325
Private nonprofit 4-year, doctorate-granting	743	268
Private for-profit less-than-2-year	1,377	70
Private for-profit 2-year	538	120
Private for-profit 4-year	335	280

¹ The number of institutions in the population and sample is prior to sample freshening.

NOTE: Details may not sum to totals due to rounding.

SOURCE: Population estimates based on IPEDS 2021-22 data.

When necessary, substitutions for sampled, eligible institutions not providing student enrollment lists may be used so that we have sufficient institution participation and sampled students. To do so, we will recreate the institution sampling frame in the same order as used for sample selection described above. Then, within the institution strata, we will identify institutions on the frame immediately before and after the sampled institution as potential substitutes. Substitutes will not include institutions already selected for either the full-scale or field test sample. Of the two substitute institutions identified, we will use the one that has the closest measure of size to the sampled institution. Any institutions included in the sample with certainty will not have substitutes because they do not have neighboring institutions with a similar measure of size that are not already in the sample.

We plan to identify substitute institutions in the following four strata:

- public less-than-2-year;
- private nonprofit 2-year or less;
- private for-profit less-than-2-year; and
- private for-profit 2-year.

These strata have had historically low institution participation rates. They also have low institution sampling rates, which allows for substitutes to be available for many sampled institutions, and there are not many institutions in these strata that were selected with certainty.

b. Student Sample

Although this submission is not for student data collection, the sample design is included here because part of the design is relevant for list collection, and the sampling of students from the enrollment lists will likely have to begin prior to OMB approval of the student data collection.

Student Enrollment List Collection

To begin NPSAS data collection, sampled institutions are asked to provide a list of their NPSAS-eligible undergraduate and graduate students enrolled in the targeted academic year, covering July 1 through June 30 (methods for contacting the sampled institutions are described below in section B.3, and student list data elements are

described in appendix D). Since NPSAS:04, institutions have been asked to limit listed students to those enrolled through April 30 and, starting in NPSAS:20, continuous enrollment institutions were asked to limit listed students to those enrolled through March 31.¹² These truncated enrollment periods exclude students who first enrolled in May or June (and additionally in April for continuous enrollment institutions), but they allow lists to be collected earlier and, as a result, data collection to be completed in less than 12 months. Any lack of coverage resulting from the truncated enrollment period will be accounted for by the poststratification weight adjustment.

We will request that high school students be included on the enrollment list even though these students are not eligible for NPSAS. While these students will be excluded from sampling, high school student counts are needed later for the weighting poststratification adjustment. We will poststratify the NPSAS students to IPEDS enrollment counts (used as control totals), which include high school students. As dual enrollment becomes more prominent, it is important that we adjust the IPEDS counts downward to account for dual enrollment. Since dual enrollment counts are not currently readily available, using high school student counts from the enrollment lists may be the best source for adjusting IPEDS counts and ensuring accurate control totals.

In addition to collecting typical enrollment lists from institutions, we will also attempt to obtain enrollment lists directly from the National Student Clearinghouse (NSC) for some institutions unable to provide enrollment lists but willing to provide permission for us to obtain their lists from the NSC. This may be done, within the strata that do not have substitute institutions (see above), for those institutions that provide data to NSC. Obtaining enrollment lists directly from NSC has the potential to reduce institution burden and help with refusal conversion. See section B.4 for results of evaluating enrollment lists obtained from the NSC during the field test.

Student Stratification

The student sampling strata will be:

1. potential baccalaureate recipients who are veterans;
2. potential baccalaureate recipients from science, technology, engineering, and mathematics (STEM) programs, White;
3. potential baccalaureate recipients from STEM programs, non-White;
4. potential baccalaureate recipients from teacher education programs, White;
5. potential baccalaureate recipients from teacher education programs, non-White;
6. potential baccalaureate recipients from business programs, White;
7. potential baccalaureate recipients from business programs, non-White;
8. potential baccalaureate recipients from other programs, White;
9. potential baccalaureate recipients from other programs, non-White;
10. other undergraduate students who are veterans;
11. other undergraduate students, Hispanic;
12. other undergraduate students, White;
13. other undergraduate students, Black;
14. other undergraduate students, other;
15. graduate students who are veterans;

¹² Many institutions know their enrolled students prior to April 30 and provide lists in February, March, or April. However, continuous enrollment institutions, including many of the for-profit institutions, typically cannot provide enrollment lists until mid-May, at the earliest, given that the lists include students enrolled through April 30.

16. master's degree students in STEM programs;
17. master's degree students in education and business programs;
18. master's degree students in other programs;
19. doctoral-research/scholarship/other students in STEM programs;
20. doctoral-research/scholarship/other students in education and business programs;
21. doctoral-research/scholarship/other students in other programs;
22. doctoral-professional practice students; and
23. other graduate students.

We are keeping the graduate student strata similar to the sampling strata used since NPSAS:16. If students fall into multiple strata, such as graduate students who are veterans, the ordering of the strata above will be used to prioritize the stratification.

Several student subgroups will be intentionally sampled at rates different than their natural occurrence within the population due to specific analytic objectives. The following groups will be oversampled:

- potential baccalaureate recipients overall and who are:
 - o veterans¹³,
 - o in STEM or teacher education programs,
 - o not White,
 - o in public 4-year, non-doctorate-granting, primarily sub-baccalaureate institutions, or
 - o in private for-profit 4-year institutions;
- other undergraduate students who are:
 - o veterans,
 - o Hispanic,
 - o Black, or
 - o other non-White race; and
- graduate students in STEM programs.

Similarly, we anticipate the following groups will be undersampled:

- undergraduate White students;
- potential baccalaureate recipients who are White, in business programs, in other non-STEM and non-teacher education programs, or in public 4-year, doctorate-granting institutions;
- master's degree students in education and business programs; and
- doctoral-research/scholarship/other students in education and business programs.

Because these groups are so large, sampling in proportion to the population would make it difficult to draw inferences about the experiences of some other types of students.

We will match the student enrollment lists to two supplemental databases prior to sampling (pre-sampling matching), as has been done since NPSAS:16. Because the veterans identified by institutions on the lists are incomplete, in order to identify veterans we will match the student enrollment lists with a list of veterans from the Veterans Benefits Administration (VBA). This veterans' information will be used with

¹³ Oversampled veterans in this document refer to veterans who receive veteran's benefits.

the veteran status from the enrollment lists for full-scale stratification. We will also match the student lists to the National Student Loan Data System (NSLDS) data and use the available financial aid data for student-implicit stratification. Within the student-explicit strata, we will sort the students by federally aided/unaided, and this will allow the sample proportions of aided and unaided students to approximately match the population within institution and student strata.

Identification of baccalaureate recipients

NPSAS:24 may serve as the base year for a 2024 cohort of B&B and will include a nationally representative sample of baccalaureate recipients, hence the stratification described above. This allows NCES to decide later if they will conduct a B&B study should funding become available. We will ask all institutions that award baccalaureate degrees to identify baccalaureate recipients. Instead of waiting until June for institutions to positively identify these students and send in lists, we will request that a baccalaureate indicator be included on the enrollment lists to flag whether students have completed requirements for or received a bachelor's degree between July 1 and the date the enrollment list is provided. In NPSAS:16, we additionally requested that institutions provide a second indicator on the lists to flag students who had not yet received their bachelor's degree but were expected to receive it by June 30.

Because of the difficulty institutions experienced in providing this second indicator in NPSAS:16, we developed a baccalaureate proxy algorithm for the NPSAS:24 field test based on analysis of NPSAS:16 data from student lists and the Central Processing System (CPS). We established the proxy separately for independent and dependent students within baccalaureate-granting institutions. Dependency status was determined based on pre-sampling matching to CPS. We classified as independent any students who, according to the enrollment list, were 24 years or older or a veteran, and did not fill out the Free Application for Federal Student Aid. We flagged all other students for whom there was no CPS data as having an unknown dependency status.

We performed logistic modeling to predict the probability of a student in baccalaureate-granting institutions being a baccalaureate recipient. Prior to running the logistic models, we divided the data into two groups – those to include in and exclude from the models. Dependent students were excluded from the model if they were 20 years old or younger; independent students were excluded from the model if they were doctoral students or other graduate students not enrolled for the master's degree (graduate students at the master's level were retained).

With the students flagged for inclusion, we then ran two logistic models – one for independent students and one for dependent students. The dependent variable in each model is baccalaureate receipt, based on enrollment list data, rather than survey data, because the baccalaureate proxy is meant to replace what an institution would provide on the list. Table 2 below lists the predictor variables used in the models and for which model it was included, independent and/or dependent. For continuous variables, we tried both the continuous and categorical versions in the models and determined that, except for Age for independent students, the categorical versions performed better at predicting baccalaureate status.

Table 2. NPSAS:24 field test predictor variables for the logistic models predicting baccalaureate status

Variable description	Model for independent students	Model for dependent students
Control and level of institution	✓	✓
Class level of student (year of enrollment – 1 st , 2 nd , 3 rd ,...)	✓	✓
Months since high school graduation (categorical – 0 or missing, 1-47, 48-61, 62 or more)	✓	
Indicator of having graduated high school at least 47 months ago		✓
Indicator of enrollment date at least 33 months ago	✓	
Indicator of enrollment date at least 45 months ago		✓
Indicator of having at least 105 credit hours	✓	✓
Age (continuous)	✓	
Age (categorical – 21, 22, 23, missing)		✓
Dependency status (dependent and unknown)		✓

NOTE: Months since high school graduation and since enrollment at NPSAS institution are based on June 30 of the NPSAS year.

We will evaluate the criteria used to include or exclude students from the baccalaureate proxy model using NPSAS:24 field test list and survey data once available, and revise it, as necessary, for use in the full-scale collection. We will include in the evaluation a second indicator, provided by sampled institutions, for students who were expected to complete requirements for or receive a bachelor’s degree between the date the enrollment list was provided and June 30, 2023, similar to what was requested in NPSAS:16. We will calculate both false-positive and false-negative rates for all baccalaureate recipients based on the modeling. Preliminary evaluation results are presented below in section B.4. Once the final evaluation of the field test proxy is complete, we will decide whether the use of a proxy in the full-scale is suitable. If the proxy is determined to be inadequate for the full-scale, we will ask the institutions to provide the second baccalaureate indicator described above.

Sample Sizes and Student Sampling

NPSAS:24 will be designed to sample about 162,000 students. We expect to obtain, minimally, 95 percent eligibility rates and at least a 70 percent response rate overall. This will yield approximately 107,730 student surveys. The graduate student sample is 25,000, matching the sample size targets in both NPSAS:18-AC and NPSAS:20. Table 3 shows the preliminary population and sample sizes, respectively, by control and level of institution and student type.

In setting the NPSAS:24 sample sizes, we need to determine the sample size of potential baccalaureate recipients, who will be part of both NPSAS and the B&B 2024 cohort, if a B&B study is conducted. The B&B cohort would include 30,000 baccalaureate recipients who respond to the NPSAS:24 survey and confirm that they have received their bachelor’s degree in the appropriate time frame. The NPSAS:24 potential baccalaureate recipient preliminary sample size will be approximately 53,040, assuming a 95 percent eligibility rate, a 70 percent survey response rate, a 19.7 percent false-positive rate, and a 3.0 percent false-negative rate among other

undergraduate students, as in NPSAS:16.¹⁴ and this will be updated after the sample optimization.

Table 3. NPSAS:24 preliminary student population and sample sizes, by control and level of institution and student type

Control and level of institution	Potential baccalaureate recipients		Other undergraduate students		Graduate students	
	Population estimate	Sample size	Population estimate	Sample size	Population estimate	Sample size
Total	2,184,067	53,043	19,382,086	83,957	4,044,750	25,000
Public less-than-2-year	0	0	85,839	848	0	0
Public 2-year	0	0	7,047,852	26,556	0	0
Public 4-year, non-doctorate-granting, primarily sub-baccalaureate	26,377	3,694	2,173,207	3,702	2,016	97
Public 4-year, non-doctorate-granting, primarily baccalaureate	162,509	3,694	800,306	3,470	133,743	2,015
Public 4-year, doctorate-granting	1,253,630	13,745	4,825,723	13,905	1,777,833	6,074
Private nonprofit 2-year or less	0	0	79,674	1,084	0	0
Private nonprofit 4-year, non-doctorate-granting	211,898	7,050	972,548	5,333	221,495	3,358
Private nonprofit 4-year, doctorate-granting	425,381	7,852	1,728,474	4,643	1,530,875	5,878
Private for-profit less-than-2-year	0	0	360,802	4,193	0	0
Private for-profit 2-year	0	0	360,545	8,661	0	0
Private for-profit 4-year	104,272	17,008	947,116	11,562	378,788	7,578

SOURCE: Population estimates based on IPEDS 2020-21 data.

Institution-level student sampling rates will be set based on frame data and adjusted to account for the overestimation of enrollment counts in IPEDS data that has been found in prior NPSAS list collections. Based on these adjusted rates, students will be sampled on a flow basis as enrollment lists are received using stratified systematic sampling procedures. As mentioned above, student strata will be sorted by federally aided/unaided students to maintain proportionality between the sample and frame. Within the graduate-student stratum for veterans, the students will be sorted by master's and doctoral degree levels to ensure that the sample will be roughly proportional to the frame. Sample yield will be monitored by institution and student sampling strata, and the sampling rates will be adjusted early, if necessary, to achieve the desired sample yields.

Quality Control Checks for Lists and Sampling

The number of enrollees on each institution's student list will be checked against the latest IPEDS full-year enrollment and completions data for each student level: baccalaureate, undergraduate, and graduate. As has been done in past rounds of NPSAS, only counts within 50 percent of non-imputed IPEDS counts will pass QC and will be allowed to move to student sampling.

¹⁴ $30,000 = (53,043 \text{ potential baccalaureate recipients} * .95 * .70 * .803) + (83,957 \text{ other undergraduate students} * .95 * .70 * .03)$, where $.803 = 1 - .197$.

The false-positive and false-negative rates will be updated for the full-scale based on field test results. The false negative rate will also be updated to account for a small percentage (less than two percent) of students who are sampled as graduate students but are baccalaureate recipients.

Institutions that fail QC will be re-contacted to resolve the discrepancy and to verify that the institution coordinator who prepared the student list clearly understood our request and provided a list of the appropriate students and data. If 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 have either confirmed that the list received is correct or have received a corrected list.

QC is very important for sampling and all statistical activities. All statistical procedures will undergo thorough QC checks, following the Quality Management Plan. We will employ a checklist for all statisticians to use to make sure that all appropriate QC checks are done for student sampling. Some specific student sampling QC checks include, but are not limited to, checking that the:

- students on the sampling frames all have a known, non-zero probability of selection;
- high school students are excluded;
- student strata are populated, as expected, based on institutions stratum;
- email addresses match student names; and
- number of students selected match the target sample sizes.

3. Methods for Maximizing Response Rates

a. NPSAS:24 Institution Contacting

Establishing and maintaining contact with sampled institutions throughout the data collection process is vital to the success of NPSAS:24. Institution participation is required in order to collect enrollment lists and select the student sample. The process by which institutions will be contacted is depicted in figure 1 and described below.

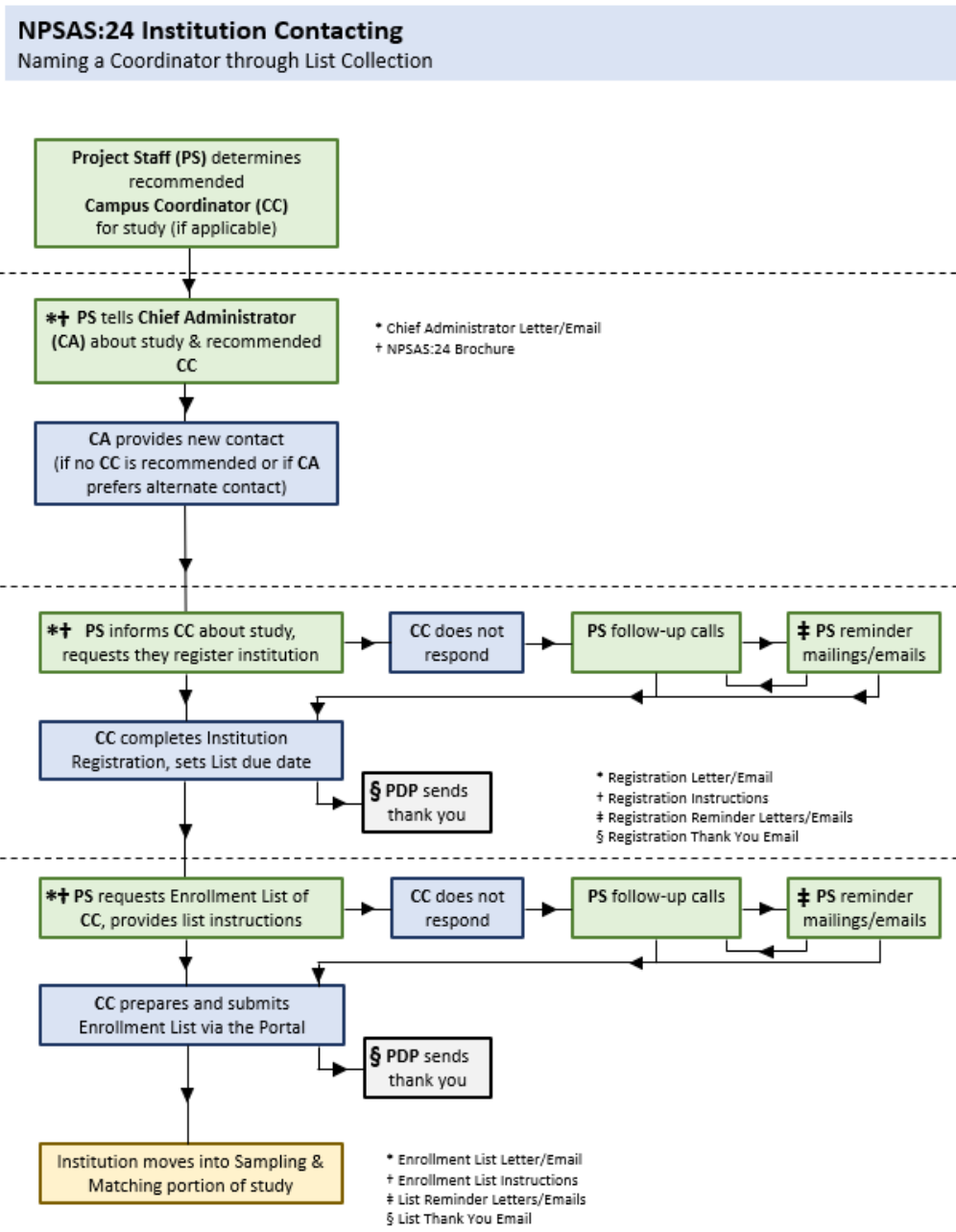
The data collection contractor will be responsible for contacting institutions on behalf of NCES. Each staff member will be assigned a set of institutions that is their responsibility for the duration of data collection. This allows contractor staff to establish rapport with institution staff and provides a reliable point of contact for the institution. 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.

The first step in the process is verification of the chief administrator's contact information. Web searches and verification calls will be conducted to confirm eligibility and confirm contact information obtained from the IPEDS header files before study information is mailed. The Higher Ed Directory (<https://hepinc.com/>) may also be used to verify information. Once the contact information is verified, we will prepare and send an information packet to the chief administrator of each sampled institution. A copy of the letter and brochure can be found in appendix D. The materials provide information about the purpose of the study and the nature of subsequent requests. Two versions of the chief administrator letter will be used, tailored to the institution's situation: (1) one letter for institutions for which we identify and recommend a potential campus coordinator from previous NPSAS participation; (2) another letter for institutions for which we cannot identify and recommend a potential campus coordinator. For institutions without a recommended coordinator, institution contactors will conduct follow-up calls to the chief administrator to secure study

participation and identify a campus coordinator. If the coordinator is not already a Postsecondary Data Portal user, they will be added as a user.

NCES and its contractor will identify relevant multi-campus systems within the sample because these systems can supply enrollment list data at the system level, minimizing burden on individual campuses. Even when it is not possible for a system to supply data from a centralized office, the system can lend support in other ways, such as by prompting institutions under its jurisdiction to participate. NCES and its contractor will undertake additional outreach activities, such as engaging with higher education organizations and networking within the postsecondary community at conferences and professional meetings. These activities are intended to promote the value of NPSAS both to data providers and data users thereby increasing interest and participation in NPSAS:24.

Figure 1. Institution contacting



Once a campus coordinator has been identified for an institution, the contractor will send the coordinator study materials with a request to complete the online Registration Page as the first step. The materials include a letter, the study brochure, and a quick guide to participation in the study (see appendix D). The primary functions of the Registration Page are to confirm the date the institution will be able to provide the student enrollment list and to determine how they will report student records data, by term or by month. Based on the information provided, a customized timeline for collecting the enrollment list will be created for each institution.

After the Registration Page is completed, the campus coordinator will be sent a letter or email requesting an electronic enrollment list of all students enrolled during the academic year. Enrollment lists will be collected from January 2024 to July 2024. As described above, the lists will serve as the frame from which the student samples will be drawn. Follow-up contacts with institutions include telephone prompts, reminder emails and mailers, typically sent prior to a deadline, and touch-base emails typically sent after a period of no outbound contact from study staff (see appendix D). After enrollment lists are received and validated by the contractor for completeness and quality, the campus coordinator will be sent a “thank you” email acknowledging appreciation for their time and effort.

Student Enrollment List Template

In the NPSAS:24 field test, we offered institutions a new option to submit student enrollment lists using a pre-formatted Excel template, similar to the Excel template offered during the student records collection. The enrollment list template option was offered for a few purposes: first, to reduce burden on institutions by making it easier to format their enrollment lists; second, to increase the uniformity of lists by encouraging institutions to submit data in a single format; and third to increase data quality by performing error checks on the list data at the time the list is uploaded. Institutions were given a choice of using the new template option or preparing the student enrollment list file in their own format (consistent with prior rounds of NPSAS). Of the 184 institutions submitting a student enrollment list in the field test, 135 (73 percent) submitted lists using the template option. Of the 42 institutions that received data errors when uploading, 41 percent uploaded a revised file that resolved the errors. Based on the number of institutions that chose the template option and the number that were able to immediately resolve data errors, we will continue to offer both the template option and the “create your own” option in the full-scale data collection.

Alternate Enrollment List Submission Method

As previously described, in addition to collecting typical enrollment lists from institutions, we will also attempt to obtain enrollment lists directly from the National Student Clearinghouse (NSC) for some institutions unable to provide enrollment lists. This may be done in the strata that do not have substitute institutions (see above) for institutions that provide data to NSC. Obtaining enrollment lists directly from NSC has the potential to reduce institution burden and help with refusal conversion. Permission from the institution will be required before obtaining enrollment list data directly from NSC.

b. Matching to Administrative Databases

Information about NPSAS:24 sampled students will be matched with their data from several administrative databases, including NSLDS, CPS including FAFSA, NSC, VBA, and student records obtained directly from postsecondary institutions. Further details

about these matches are provided in the Supporting Statement Part A (sections A.1, A.2, A.10, and A.11) and in appendix C. We continue to explore matches to other potential data sources to be added to the full-scale collection, such as the Supplemental Nutrition Assistance Program (SNAP) recipients' data from the U.S. Department of Agriculture.

c. Postsecondary Data Portal (PDP)

The NPSAS:24 institution data collection will utilize NCES' Postsecondary Data Portal (PDP) website. The flexible design of the website allows it to be used for multiple NCES postsecondary institution sample studies in data collection at the same time, even when those studies collect different types of data. Currently, there are no plans for other postsecondary data collections to be underway using the PDP during the NPSAS:24 full-scale.

The PDP provides to users both general-purpose and study-specific content. General-purpose pages provide overview information about NCES postsecondary studies and use of the website. These pages are identified in appendix D as the "pre-login" pages. Once a user logs in, they see pages with study-specific content. These pages are identified in appendix D as the "after login" content. The NPSAS:24 study-specific content includes FAQs about NPSAS:24, instructions and resources, and pages for providing data (appendix D). Institutions see study-specific PDP content only for the study or studies for which they have been sampled.

The PDP was updated for the NPSAS:24 field test to add a new option for institutions to submit the student enrollment list using a pre-formatted Excel template file. This new feature was designed for institutions that had requested an Excel template in prior rounds of NPSAS and also facilitates real-time error checking during the enrollment list upload process. Enrollment list error checking will provide institution staff with immediate feedback about potential enrollment list data problems, rather than waiting for NPSAS staff to review the enrollment list and follow-up with feedback at a later date. Institutions that prefer to create their own enrollment list files will still be able to do so.

Data Security on the PDP

Because of the risks associated with transmitting confidential data on the internet, the latest technology systems will be incorporated into the web application to ensure strict adherence to NCES confidentiality guidelines. The web server will include a Secure Sockets Layer (SSL) certificate and will be configured to force encrypted data transmission over the Internet. All data-entry modules on this site require the user to log in before accessing confidential data. Logging in requires entering an assigned ID number and two-factor authentication with a code sent via email and a password. Through the PDP, the campus coordinator at the institution will be able to use a "Manage Users" link to add and delete users, as well as reset passwords and assign roles. Each user will have a unique username and will be assigned to one e-mail address. Upon account creation, the new user will be sent a temporary password by the PDP. When logging in for the first time, the new user will be required to create a new password. The system automatically will log out after 20 minutes of inactivity. Files uploaded to the secure website will be stored in a secure project folder that is accessible and visible to authorized project staff only.

1. Tests of Procedures or Methods

The NPSAS:24 field test enrollment list collection was used to evaluate several new

procedures designed to improve data quality and decrease burden on institutions: the bachelor’s degree recipient proxy used for identifying potential B&B cohort members (described above in section B.2), the collection of enrollment lists from National Student Clearinghouse (described above in section B.3), and the new pre-formatted enrollment list template file (described above in section B.3). Tests of procedures or methods related to student data collection will be described in the forthcoming Student Data Collection Package to be submitted in the fall of 2023.

a. Bachelor’s degree recipient proxy

As discussed in section B.2, we developed a baccaureate proxy in the field test to identify students who had not yet received their bachelor’s degree but were expected to receive it by June 30. The institutions provided a baccaureate indicator on the enrollment lists to flag whether students have completed requirements for or received a bachelor’s degree between July 1 and the date the enrollment list is provided. We used this flag and the proxy to identify students to sample as potential baccaureate recipients. For use in evaluating the proxy, we additionally requested that institutions provide a second indicator on the lists to flag students who had not yet received their bachelor’s degree but were expected to receive it by June 30.

We have analyzed preliminary field test data to begin evaluating the baccaureate proxy. Table 4 shows that about 50 percent of students sampled as baccaureates using the proxy and surveyed were confirmed to have received or expected to receive their bachelor’s degree between July 1, 2022 and June 30, 2023. This is a false positive rate of 50 percent (100-50).

Table 4. Baccaureate status determination, by student type¹

Student type	Students surveyed in field test	Confirmed B&B eligibility	
		Number	Unweighted percent
Total	3,050	1,110	36.2
Total undergraduate	2,880	1,100	38.2
Potential Baccaureate	2,040	1,090	53.3
Identified on student lists ²	180	150	83.2
Identified by proxy definition	1,860	940	50.4
Other undergraduate	840	20	1.8
Graduate	170	#	1.8

Rounds to zero.

¹ Data are current as of May 30, 2023.

² Identified on student lists using baccaureate indicator on the enrollment lists for students who have completed requirements for or received a bachelor’s degree between July 1 and the date of the enrollment list.

NOTE: Detail may not sum to totals due to rounding.

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2023–24 National Postsecondary Student Aid Study (NPSAS:24) Field Test.

Table 5 shows that the baccaureate proxy matches the expected baccaureate flag from the enrollment lists for about 61 percent of the surveyed students. The baccaureate proxy correctly identifies about 64 percent of baccaureate recipients, as compared to the survey, and the expected baccaureate flag from the enrollment lists correctly identifies about 77 percent of baccaureate recipients, as compared to the survey.

Table 5. Counts and percentages of surveyed students by baccalaureate status¹

Confirmed baccalaureate recipient in survey ²	Baccalaureate recipient by proxy	Baccalaureate recipient by expected flag on lists ³	Number of surveyed students	Percent of surveyed students
Yes	Yes	Yes	660	21.6
Yes	Yes	No	280	9.1
Yes	No	Yes	60	2.1
Yes	No	No	100	3.4
No	Yes	Yes	190	6.1
No	Yes	No	740	24.1
No	No	Yes	120	4.0
No	No	No	900	29.6

¹ Data are current as of May 30, 2023.

² Baccalaureate recipients not confirmed in the survey include responses of no and missing responses.

³ Baccalaureate recipients not flagged as expected on the list include responses of no, does not apply, and unknown and missing responses.

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2023–24 National Postsecondary Student Aid Study (NPSAS:24) Field Test.

At the end of field test student data collection (scheduled for June 30, 2023), we will conduct a final evaluation of the baccalaureate proxy, including determining if different models would perform better, based on field test data. At that point, we will decide if it is preferable to use a proxy or to ask the institutions to provide the expected baccalaureate indicator.

b. Collection of enrollment lists from National Student Clearinghouse

During the field test, for a subset of institutions, we obtained enrollment list data directly from NSC, in addition to collecting their standard NPSAS enrollment lists from institutions. By comparing institutions' NSC data with their regular NPSAS lists, the viability of collecting lists from NSC on a wider scale in the full-scale study could be assessed. After discussing the plan with staff from the institutions identified for this initiative and obtaining their permission, we were able to obtain NSC enrollment list data for over 110,000 students from 21 institutions.

To evaluate the quality of the NSC lists, student counts were compared between the lists from NSC and the institutions, as shown in Table 6.¹⁵ The NSC counts were higher than institution counts for some institutions and lower for others. Overall, the differences in the counts were minimal. When comparing the NSC and institution list counts to IPEDS counts, as described in section B.2, the results of lists passing or failing the QC checks were the same for all but one institution.

¹⁵ The comparison was done for 18 institutions that had sufficient data on degree program to compare counts by student type.

Table 6. Comparison of NSC and institution enrollment list counts¹

Student type	Median absolute relative percent difference ²	Percent of institutions with same QC results (compared to IPEDS counts)
Baccalaureate recipients	30.0	94.0
Undergraduate students	11.0	94.0
Graduate students	8.0	100.0

¹ The comparison was done for 18 institutions that had sufficient data.

² The median absolute relative percent difference is the median of the absolute value of (institution list count - NSC list count) / institution list count. Baccalaureate counts were included if the institution provided the baccalaureate indicator on the list. Baccalaureate and graduate counts were included only for the 4-year institutions.

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2023–24 National Postsecondary Student Aid Study (NPSAS:24) Field Test.

Additionally, the NSC lists were examined to determine if there were sufficient non-missing data for the most important data elements for student sampling and contacting, as shown in Table 7. The percentage of data received from NSC was generally high, except for race/ethnicity. Also, the NSC data received for the field test subset of institutions were more complete than the institution data for both Social Security number and baccalaureate indicator.

Table 7. Percentage of non-missing important data elements

Data element	Percent of non-missing data	
	From NSC	From institutions
Social Security number	96.1	88.5
Date of birth	99.9	99.9
Race/ethnicity	9.6	86.2
Degree program	92.4	100.0
Received bachelor's since July 1	100.0	65.1
Contact information	100.0	100.0

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2023–24 National Postsecondary Student Aid Study (NPSAS:24) Field Test.

While the NSC lists contained data that were fairly comparable to the lists provided by the institutions, there are some noteworthy disadvantages to the NSC lists, including the low amount of race/ethnicity data; the inclusion of dually enrolled high school students, especially for less-than-4-year institutions; and the lack of data elements to identify students *expected* to receive their bachelor's degree before June 30.

As discussed in sections B.2 and B.3, we plan to obtain enrollment lists from NSC for some institutions refusing to provide enrollment lists but willing to provide permission for us to obtain their lists from NSC. Use of NSC lists will help us obtain student sample size targets, as well as reduce institution burden and help with refusal conversion.

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

NPSAS:24 is being conducted by NCES. The following statisticians at NCES are responsible for the statistical aspects of the study: Dr. Tracy Hunt-White, Dr. David Richards, Dr. Sean Simone, and Dr. Chris Chapman. NCES's prime contractor for NPSAS:24 is RTI International (Contract# 91990022C0017), and subcontractors include Activate Research; ARSIEM Corporation; EurekaFacts; Forum One

Communications; HR Directions; KEN Consulting, Inc.; Leonard Resource Group; Research Support Services; Strategic Communications, Inc.; The Equity Paradigm; and Whitworth Kee Consulting, LLC. Dr. Anthony Jones, Dr. Vincent Castano, Dr. Eric Atchison, Richard Reeves, Dr. Sandy Baum, Dr. Matt Springer, and Dr. Shelly Steward are consultants on the study. The following staff members at RTI are working on the statistical aspects of the study design: Dr. Jennifer Wine, Dr. Josh Pretlow, Peter Siegel, Stephen Black, Ruby Johnson, Jennifer Cooney, Dr. T. Austin Lacy, and Dr. Emilia Peytcheva. Principal professional RTI staff, not listed above, who are assigned to the study include: Kristin Dudley, Jamie Wescott, Ashley Wilson, Austin Caperton, Jeff Franklin, Dr. Jerry Timbrook, and Dr. Erin Velez.