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

STUDENT DATA COLLECTION AND STUDENT RECORDS

Supporting Statement Part B

OMB # 1850-0666 v. 40

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

September 2023
revised July 2024
revised August 2024
revised October 2024

Contents

Tables and Figures

B. Collection of Information Employing Statistical Methods

This submission requests clearance for the 2023–24 National Postsecondary Student Aid Study (NPSAS:24) full-scale student data collection materials and procedures, which include the institution student record data abstraction and student survey. It carries over respondent burden, procedures, and materials related to the NPSAS:24 institution contacting, enrollment list collection, list sampling, and administrative matching activities previously approved in September 2023 (OMB#1859-0666 v. 35). Specific plans are provided below.

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 institution in any term or course of instruction between July 1, 2023 and April 30, 2024[3] 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

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

- 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. Prior to the start of full-scale data collection, the institution sample was freshened using the IPEDS:2022-23 header, 2022-23 IC, 2021-22 Completions, and 2021-22 Full-year Enrollment files. This ensures that all potentially eligible institutions have a probability of selection. As part of this process, we also used the current IPEDS files to identify full-scale institutions no longer eligible for NPSAS:24 due to closure. Because it is possible that some for-profit institutions and large chains of for-profit institutions will have closed or been sold after the latest IPEDS data collection, we will 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
- Private for-profit 4-year.

The full-scale sample of 2,020 institutions was selected first using a variation of probability proportional to size (PPS) sampling called sequential probability minimum replacement (PMR) sampling.[5] 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.[6] 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 measure[7] was used to help achieve self-weighting samples[8] 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

[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] 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.
[6] https://support.sas.com/documentation/cdl/en/statug/63347/HTML/default/viewer.htm#statug_surveyselect_a0000000173.htm.
[7] 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.
[8] Self-weighting samples have equal weights within sampling domains.

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.

Additionally, four institutions were added to the sample during the freshening process. The freshened sample size was determined based on the enrollment of the 66 new institutions such that the probabilities of selection are similar to the originally sampled institutions in the same strata. The total institution sample size is 2,025.

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 1,570 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;[9]
4. Carnegie classification of postsecondary institutions;[10] 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 institution[1]

Control and level of institution	Population estimate	Sample size
Total	**5,920**	**2,025**
Public less-than-2-year	224	22
Public 2-year	896	378
Public 4-year, non-doctorate-granting, primarily sub-baccalaureate	186	83
Public 4-year, non-doctorate-granting, primarily baccalaureate	214	116
Public 4-year, doctorate-granting	391	340
Private nonprofit 2-year or less	178	20
Private nonprofit 4-year, non-doctorate-granting	854	325
Private nonprofit 4-year, doctorate-granting	751	268
Private for-profit less-than-2-year	1,369	71
Private for-profit 2-year	530	122
Private for-profit 4-year	327	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 2022-23 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

[9] 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/idueshsi/definition.html) and using IPEDS Hispanic or Latino enrollment data.
[10] We used the 2018 version of Carnegie classification available on the IPEDS files.

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

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 D1). 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.[11] 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;

[11] 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.

3. potential baccalaureate recipients from STEM programs, all other race/ethnicity categories (Black or African American, Hispanic or Latino, Asian, American Indian or Alaska Native, Native Hawaiian or Pacific Islander);
4. potential baccalaureate recipients from teacher education programs, White;
5. potential baccalaureate recipients from teacher education programs, all other race/ethnicity categories (Black or African American, Hispanic or Latino, Asian, American Indian or Alaska Native, Native Hawaiian or Pacific Islander);
6. potential baccalaureate recipients from business programs, White;
7. potential baccalaureate recipients from business programs, all other race/ethnicity categories (Black or African American, Hispanic or Latino, Asian, American Indian or Alaska Native, Native Hawaiian or Pacific Islander);
8. potential baccalaureate recipients from other programs, White;
9. potential baccalaureate recipients from other programs, all other race/ethnicity categories (Black or African American, Hispanic or Latino, Asian, American Indian or Alaska Native, Native Hawaiian or Pacific Islander);
10. other undergraduate students who are veterans;
11. other undergraduate students, Hispanic or Latino;
12. other undergraduate students, White;
13. other undergraduate students, Black or African American;
14. other undergraduate students, other;
15. graduate students who are veterans;
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:
    - veterans,[12]
    - in STEM or teacher education programs,
    - all race/ethnicity categories other than White (Black or African American, Hispanic or Latino, Asian, American Indian or Alaska Native, Native Hawaiian or Pacific Islander),
    - in public 4-year, non-doctorate-granting, primarily sub-baccalaureate institutions, or
    - in private for-profit 4-year institutions;
- other undergraduate students who are:
    - veterans,
    - Hispanic or Latino,
    - Black or African American, or
    - all race/ethnicity categories other than White (Black or African American, Hispanic or Latino, Asian, American Indian or Alaska Native, Native Hawaiian or Pacific Islander);
- graduate students in STEM programs.

[12] Oversampled veterans in this document refer to veterans who receive veteran's benefits.

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 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 July 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 those to 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.

We evaluated the criteria used to include or exclude students from the baccalaureate proxy model using NPSAS:24 field test list and survey data, and included 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 calculated both false-positive and false-negative rates for all baccalaureate recipients based on the modeling. Evaluation results are presented below in section B.4. Based on results of the evaluation, we will adjust and test a new model using field test results. During full-scale sampling, we will continue to ask institutions to provide the second baccalaureate indicator described above. If an institution indicates not knowing a student's baccalaureate status (selects "don't know"), we will run the proxy to estimate the student's status. Having the proxy as a supplement will help to prevent underestimation of the number of baccalaureates for sampling.

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 – 1st, 2nd, 3rd,…)	✓	✓
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.

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.[13] and this will be updated after the sample optimization.

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.

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

	Potential baccalaureate recipients		Other undergraduate students		Graduate students	
Control and level of institution	**Population estimate**	**Sample size**	**Population estimate**	**Sample size**	**Population estimate**	**Sample size**
Total	**2,134,236**	**53,043**	**18,928,868**	**83,957**	**4,081,808**	**25,000**
Public less-than-2-year	0	0	84,920	848	0	0
Public 2-year	0	0	6,705,311	26,556	0	0
Public 4-year, non-doctorate-granting, primarily sub-baccalaureate	26,686	3,694	2,167,527	3,702	2,708	97
Public 4-year, non-doctorate-granting, primarily baccalaureate	157,581	3,694	766,915	3,470	132,748	2,015
Public 4-year, doctorate-granting	1,217,550	13,745	4,787,445	13,905	1,817,471	6,074
Private nonprofit 2-year or less	0	0	73,432	1,084	0	0
Private nonprofit 4-year, non-doctorate-granting	205,571	7,050	970,369	5,333	212,716	3,358
Private nonprofit 4-year, doctorate-granting	423,620	7,852	1,747,074	4,643	1,549,338	5,878
Private for-profit less-than-2-year	0	0	374,444	4,193	0	0
Private for-profit 2-year	0	0	355,346	8,661	0	0
Private for-profit 4-year	103,228	17,008	896,085	11,562	366,827	7,578

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

[13] 30,000 = (53,043 potential baccalaureate recipients * .95 * .70 * .803) + (83,957 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

Achieving high response rates in NPSAS:24 will depend on successfully recruiting institutions, obtaining institution data and other administrative data, and identifying and locating sample members and being able to gain their cooperation. The following sections outline methods for maximizing institution and student response to in the NPSAS full-scale data collection.

a. Collection of Data from Institutions

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 D1. 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.

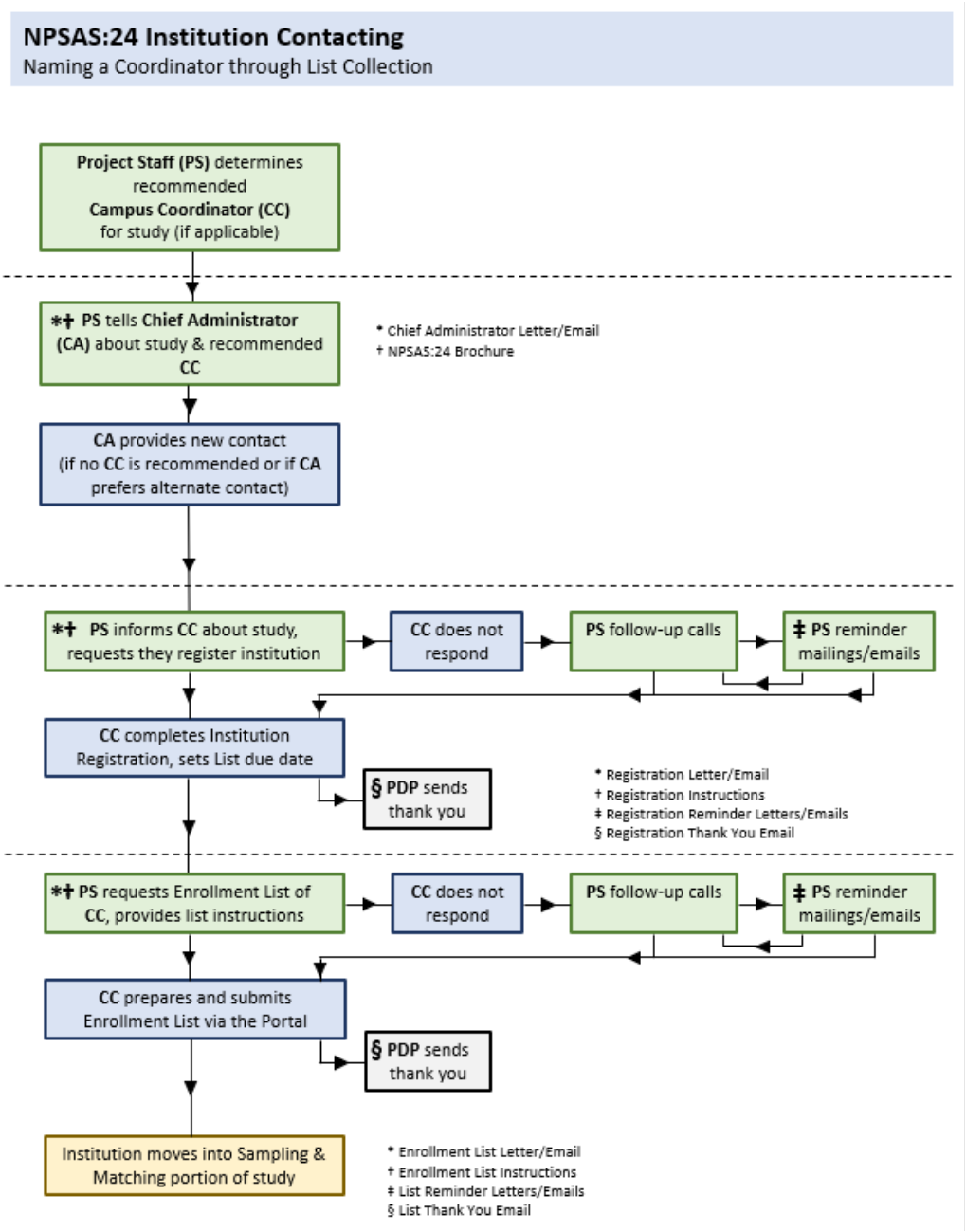
NCES and its contractor will identify relevant multi-campus systems within the sample because these systems can supply enrollment list data at the system level, minimizing burden on individual campuses. Even when it

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.

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 D1). 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 D1). 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.

Figure 1. Institution contacting



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.

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 D1 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 D1). 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.

Student Records

After students are sampled from an institution's enrollment list, the institution coordinator will receive a mailing containing a letter requesting student records data for those sampled students. Institutional contactors will follow up after the mailing to ensure receipt of the package and to answer any questions. Follow-up contacts include telephone prompts, reminder emails that are typically sent 2 weeks prior to a deadline, and touch-base emails typically sent when 3–4 weeks have passed with no outbound contact from study staff. Contact materials are included in Appendix D2. Staff will also be available by telephone and email to help when institution staff have questions or encounter problems.

As with the enrollment list collection, the student record collection will utilize the PDP. The content of the PDP specific to student records collection is included in Appendix F (the student records instrument content) and Appendix D2 (student records communication materials). The following options will be offered to institutions for collecting student records:

- Web-based data entry interface. The web-based data entry interface allows the coordinator to enter data by student, by year.
- Excel workbook. An Excel workbook will be created for each institution and will be preloaded with each sampled student's ID, name, date of birth, and last four digits of SSN (if available). To facilitate simultaneous data entry by different offices within the institution, the workbook contains a separate worksheet for each of the following topic areas: Student Information, Financial Aid, Enrollment, and Budget. The user will download the Excel worksheet from the PDP, enter the data, and then upload the data. Validation checks will occur both within Excel as data are entered

and when the data are uploaded. Data will be imported into the web application so that institution staff can check their submission for quality control purposes.

- CSV (comma separated values) file. Institutions with the means to export data from their internal database systems to a flat file may use this method of supplying student records. Institutions that select this method will be provided with detailed import specifications, and all data uploading will occur through the PDP. Like the Excel workbook option, data will be imported into the web application so that institution staff can check their submission before finalizing.
- Alternate method. Institutions will be offered an alternate submission format allowing staff to upload data in any format or file type that is convenient, rather than making their data conform to our template or CSV specifications, as a refusal aversion strategy (described below).

Refusal Aversion Strategies with Institutions

If institution staff report a lack of time or resources needed to provide student records data, the following additional accommodations will be offered:

- reimbursement to help offset labor or staffing costs;
- a reduced set of the most critical data elements (see data elements marked with an asterisk in Appendix D2); and/or
- an alternate submission format allowing staff to upload data in any format or file type that is convenient, rather than making their data conform to our template or CSV specifications.

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.

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. Collection of Student Survey Data

Tracing of Sample Members

To yield the maximum number of located cases with the least expense, we designed an integrated tracing approach, with the following elements.

- Advance tracing activities, which will occur prior to the start of data collection, include initial batch database searches, such as to the National Change of Address (NCOA) databases, for cases with sufficient contact information to be matched. To handle cases for which contact information is invalid or unavailable, additional advance tracing through proprietary interactive databases will expand on leads found.
- Hard copy mailings, emails, and text messages will be used to maintain ongoing contact with sample members, prior to and throughout data collection. The student contacting materials, which will be developed with a design appealing to students in 2024, are provided in Appendix J. The initial mailing to sample members will include a letter announcing the start of data collection, requesting that the sample member complete the web survey, and including a toll-free number, the study website address, a Study ID and password, and a study brochure. We will send a similar email message mirroring information provided in the mailing.
- Sample members will have a variety of means to provide updated contact information and contact preferences. Students can use an Update Contact Information page on the secure NPSAS:24 website to provide their contact information, including cell phone number, as well as provide contacting preferences with respect to phone calls, mailings, emails, and text messages. Help Desk calls and emails providing information about a sample member's text message preferences will be monitored and the sample member's data record updated as soon as the information becomes known.
- The telephone locating and surveying stage includes calling all available telephone numbers and following up on leads provided by parents and other contacts.
- The pre-intensive batch tracing stage consists of the LexisNexis SSN and Premium Phone batch searches that will be conducted between the telephone locating and surveying stage and the intensive tracing stage.
- Once all known telephone numbers are exhausted, a case will move into the intensive tracing stage during which tracers will conduct interactive database searches using all known contact information for a sample member. With interactive tracing, a tracer assesses each case on an individual basis to determine which resources are most appropriate and the order in which each should be used. Sources that may be used, as appropriate, include credit database searches, such as Experian, various public websites, and other integrated database services.
- Other locating activities will take place as needed, including a LexisNexis email search conducted for nonrespondents toward the end of data collection.

Training for Data Collection Staff

Telephone data collection will include supervisors and interviewers. Training programs for these staff members are critical to maximizing response rates and collecting accurate and reliable data.

Team supervisors, who are responsible for all supervisory tasks, will attend their own project-specific training, in addition to the interviewer training. They will receive an overview of the study, background and objectives, and the data collection instrument through a question-by-question review. Supervisors will also receive training in the following areas: providing direct supervision during data collection; handling refusals; monitoring interviews and maintaining records of monitoring results; problem resolution; case review; specific project procedures and protocols; reviewing reports generated from the ongoing Computer Assisted Telephone Interviewing (CATI); and monitoring data collection progress.

Training for interviewers is designed to help staff become familiar with and practice using the CATI case management system and the survey instrument, as well as to learn project procedures and requirements. Particular attention will be paid to quality control initiatives, including refusal avoidance and methods to ensure that quality data are collected. Interviewers will receive project-specific training on telephone interviewing and answering questions from web participants regarding the study or related to specific items within the survey. At the conclusion of training, all NPSAS data collection staff must meet certification requirements by successfully completing a certification interview. This evaluation consists of a full-length interview with project staff observing and evaluating interviewers, as well as an oral evaluation of interviewers' knowledge of the study's Frequently Asked Questions.

Case Management System

Surveys will be conducted using a single web-based survey instrument for both web (including mobile devices) and CATI data collection. Control of data collection activities will be accomplished through a CATI case management system, which is equipped with the numerous capabilities, including: on-line access to locating information and histories of locating efforts for each case; a 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. The automated scheduling module 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 non-appointment 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.
- 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 an 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.

NPSAS:24 data collection comprises a total of 9 sample waves which correspond to time in data collection. Sample members in the earliest waves of data collection will receive multiple contacts (e.g., emails, texts) over the course of the data collection period. Repetition of this nature can cause "wearout" (Pechmann & Stewart, 1988), which reduces sensitivity and attention to communications.[14] Over time, sample members may become desensitized to the content of study communication materials making them less likely to notice the offer of a boost incentive or abbreviated survey. Consequently, we plan to pause reminders for nonresponding sample members in Waves 1 - 3 for a period of four weeks (weeks of 7/8/2024 to 7/29/2024), as noted below in table 4. After this pause, we will re-engage with these sample members with either the incentive boost, if approved (described below), or the abbreviated survey. We expect that the pause in data collection results in an increased salience of the boost/abbreviated survey offer when it arrives, thereby increasing the propensity

[14] Pechmann, C. and Stewart, D.W., 1988. Advertising repetition: A critical review of wearin and wearout. *Current issues and research in advertising*, 11(1-2), pp.285-329.

of response to the new request (Groves, Singer and Corning, 2000).[15] We will offer 4-week breaks in communications according to the wave of data collection, as noted in red in the table. In addition, all waves will receive a break of almost 2 weeks ahead of the United States Presidential Election on November 5, 2024.

Table 4. Weeks in NPSAS:24 data collection, by data collection wave

Week	Date beginning[1]	Wave								
		1	2	3	4	5	6	7	8	9
1	26-Feb	1								
2	4-Mar	2								
3	11-Mar	3								
4	18-Mar	4								
5	25-Mar	5	1							
6	1-Apr	6	2							
7	8-Apr	7	3							
8	15-Apr	8	4	1						
9	22-Apr	9	5	2						
10	29-Apr	10	6	3						
11	6-May	11	7	4	1					
12	13-May	12	8	5	2					
13	20-May	13	9	6	3					
14	27-May	14	10	7	4	1				
15	3-Jun	15	11	8	5	2				
16	10-Jun	16	12	9	6	3				
17	17-Jun	17	13	10	7	4	1			
18	24-Jun	18	14	11	8	5	2			
19	1-Jul	19	15	12	9	6	3			
20	8-Jul	20	16	13	10	7	4	1		
21	15-Jul	21	17	14	11	8	5	2		
22	22-Jul	22	18	15	12	9	6	3		
23	29-Jul	23	19	16	13	10	7	4	1	
24	5-Aug	24	20	17	14	11	8	5	2	
25	12-Aug	25	21	18	15	12	9	6	3	
26	19-Aug	26	22	19	16	13	10	7	4	1
27	26-Aug	27	23	20	17	14	11	8	5	2
28	2-Sep	28	24	21	18	15	12	9	6	3
29	9-Sep	29	25	22	19	16	13	10	7	4
30	16-Sep	30	26	23	20	17	14	11	8	5
31	23-Sep	31	27	24	21	18	15	12	9	6
32	30-Sep	32	28	25	22	19	16	13	10	7
33	7-Oct	33	29	26	23	20	17	14	11	8

[15] Groves RM, Singer E, Corning AD. Leverage-salience theory of survey participation: Description and an illustration. *Public Opinion Quarterly.* 2000;64:299-308.

34	**14-Oct**	34	30	27	24	21	18	15	12	9
35	**21-Oct**	35	31	28	25	22	19	16	13	10
36	**28-Oct**	36	32	29	26	23	20	17	14	11
37	**4-Nov**	37	33	30	27	24	21	18	15	12
38	**11-Nov**	38	34	31	28	25	22	19	16	13
39	**18-Nov**	39	35	32	29	26	23	20	17	14
40	**25-Nov**	**Expected end of data collection**								

[1]All dates shown are in 2024.

Survey Instrument Design

The survey will employ a web-based instrument and deployment system, which has been in use since NPSAS:08. The system provides multimode functionality that can be used for self-administration, including on mobile devices, CATI, or data entry.

In addition to the functional capabilities of the case management system and web instruments 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 mobile-friendly self-administered and interviewer-administered options.
- Offering incentives to encourage response.
- Assigning experienced CATI interviewers 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.
- 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 during the initial contact, but allowing respondent flexibility in scheduling appointments to be interviewed.
- Thoroughly reviewing all refusal cases and making special conversion efforts whenever feasible (see next section).
- Implementing and assuring participants of confidentiality procedures, including restricting the ability for the respondent to view survey responses from prior log in sessions (i.e. no ability to use navigation buttons to go to "Previous" survey questions from another log in session) and the survey automatically logging out of a session after 10 minutes of inactivity.
- For the NSAS:24 full-scale collection, item-by-item toggling between English and Spanish languages at the discretion of the web respondent, or telephone interviewer when warranted.

Refusal Aversion and Conversion of Student Sample Members

Recognizing and avoiding refusals is important to maximize the response rate. We will emphasize this, and other topics related to obtaining cooperation during interviewer training. Supervisors will monitor interviewers intensely during the early days of outbound calling 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.

Refusal conversion efforts will be delayed for at least 1 week to give the respondent 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.

Offer of Incentives

Sample members will be offered $30 for a completed survey, paid by the respondent's choice of check or PayPal. In addition, we anticipate that a boost incentive will be needed for key student subgroups to further encourage participation and reduce the potential for nonresponse bias.

Researchers have commonly used incentive boosts as a nonresponse conversion strategy for sample members who have implicitly or explicitly refused to complete the survey (e.g., Groves and Heeringa 2006; Singer and Ye 2013).[16] [17] These boosts are especially common in large federal surveys during their nonresponse follow-up phase and have been implemented successfully in other postsecondary education surveys (e.g., HSLS:09 second follow-up; BPS:12/17; NPSAS:20). In NPSAS:20, a $10 incentive boost increased the overall response rate by about 3.2 percentage points above the projected response rate.

To determine whether a boost incentive is needed, we are monitoring response rates overall, within data collection wave, and within key student subgroups, particularly those that will be oversampled because of historic underrepresentation. At least at two points in time (June and September), we will run logistic regression models to predict likelihood of response for nonrespondents, prioritizing members of the oversampled groups (see section B.2), with their sampling characteristics and any potential interactions as predictors, identifying those most likely to be nonrespondents. Results of the June analysis and recommendation on which cases are being offered a $10 boost are provided below, together with the results and recommendation from the September analysis. .

Below we provide an example of the data collection flow for the oversampled cases for the earliest data collection waves:

1. **Feb-March:** All oversampled members are invited to complete the NPSAS:24 survey and offered a $30 promised incentive.
2. **June (Sept for later waves):** Predictive models identify cases likely to be missing from the respondent pool. Those cases receive a $10 incentive boost once approved.[18]
3. **August:** Data collection reminders are paused for 2-3 weeks.
4. **September:** Nonresponding oversample members are reinvited to complete the NPSAS:24 survey ($40 promised incentive).
5. **October:** All nonresponding cases are offered an abbreviated survey.
6. **November:** Final push and end of data collection.

For the rest of the sample, the data collection plan will be as follows:

1. All sample members are invited to complete the NPSAS:24 survey and offered a $30 promised incentive.
2. Data collection reminders are paused for 2-3 weeks for cases who have been in data collection for about 8-10 weeks.
3. Cases who have been in data collection for about 5 months (3-4 for later sample releases) are offered the abbreviated survey.
4. Begin end of data collection notifications for cases who have been in data collection for 6 months; such cases might be recontacted for a final push at the beginning of November, if needed.
5. End of data collection for all cases in November.

[16] Groves R.M., and Heeringa, S.G. (2006). Responsive Design for Household Surveys: Tools for Actively Controlling Survey Errors and Costs. Journal of the Royal Statistical Society Series A-Statistics in Society, 169(3): 439-457.
[17] Singer, E. and Ye, C. (2013). The Use and Effects of Incentives in Surveys. Annals. Annals of the American Academy of Political and Social Science, 645(1): 112-141.
[18] It is possible all nonresponding oversampled cases are eligible for the $10 boost.

Identification of Subgroups for Boost Incentive in June 2024

To identify subgroups for the incentive boost, we estimated a binary logistic regression model predicting the probability of a NPSAS:24 survey response for respondents, partial respondents, and nonrespondents in Waves 1-5 of data collection. This model included nine sample member characteristics as substantive predictors: gender (coded as male, female, other), age (25 or younger, 26 to 39, 40 and older), race (White, races other than White), ethnicity (Hispanic/Latino, not Hispanic/Latino), veteran status (veteran, not a veteran), control of sample member's institution (public, private non-profit, private for-profit), level of sample member's institution (less than 2 year, 2 to less than 4 year, 4 year or higher non-doctoral, 4 year or higher doctoral), whether sample member's course of study is STEM (STEM, not STEM), and undergraduate status (undergraduate, not an undergraduate). These variables were obtained from enrollment lists. In cases where data from the enrollment lists were missing, we replaced missing values with sample members' substantive answers to the NPSAS:24 survey, where available.

The model also included three variables controlling for design features of the survey: the sample member's data collection wave, whether the sample member was assigned for CATI calling, and time of day that reminder emails were sent to the sample member. The overall model fit was good – the pseudo R-square for the final model was 0.4.

We then used this model to estimate predicted probabilities of NPSAS:24 survey response for each category of each of our nine sample member characteristics, holding all other variables at their means. Table 2 below displays these predicted probabilities, along with response rates for each subgroup as of July 1st, 2024.

We identified three subgroups that had lower response rates and/or propensities – *control* of private for-profit (low response rate of 38.4 percent), *institution level* of less-than-2 year (low response propensity of 0.58 and low response rate of 34.1 percent), and *institution level* of 2-years but less-than-4 year institutions (low response propensity of 0.66 and low response rate of 37.5 percent). Those three groups have historically responded at lower rates across NPSAS surveys and may benefit from a design change in line with the leverage saliency theory, stipulating that "one-fits all" incentive amount is not a good solution to nonresponse error (Groves, Singer and Corning, 2000).[19]

Differential incentives have been proven successful in bringing in groups of focal importance who were otherwise underrepresented (e.g., Groves, Singer and Corning, 2000; Groves and Heeringa, 2006; Peytcheva, Kirchner and Cooney, 2018).[20,21] Such strategy was successfully employed in NPSAS:20 when additional $10 were offered to nonrespondents in three key analyses groups during the last 8 waves of data collection, resulting in an average response rate increase of 17.53 percent across waves relative to the projected response rate under the original design. We therefore recommend offering a $10 promised incentive boost for cases belonging to any of the three groups mentioned above in the earliest waves, to encourage participation and reduce the potential for nonresponse bias. This would result in an incentive boost for approximately 4,500 nonresponding sample members from Waves 1 – 3. For the rest of the sample member characteristics, response rates and propensities were generally similar across subgroups, but we will continue to monitor them during data collection and we will re-evaluate the remaining waves in September 2024.

Identification of Subgroups for Boost Incentive in September 2024

To identify subgroups for the September incentive boost with Waves 4 through 9, we again estimated a binary logistic regression model predicting the probability of a NPSAS:24 survey response for respondents, partial

---

[19] Ibid.

[20] Groves RM, Heeringa SG. Responsive design for household surveys: tools for actively controlling survey errors and costs. *Journal of the Royal Statistical Society Series a-Statistics in Society.* 2006;169(3):439-457. doi: DOI 10.1111/j.1467-985X.2006.00423.x.

[21] Peytcheva, E, Kirchner, A., and Cooney, J. 2018. Experimental Comparison of Two Data Collection Protocols for Previous Wave Nonrespondents. Paper presented at the Methodology of Longitudinal Surveys II conference, Essex, U.K.

respondents, and nonrespondents in Waves 1-7 of data collection. This model included nine sample member characteristics as substantive predictors: gender (coded as male, female, other), age (25 or younger, 26 to 39, 40 and older), race (White, races other than White), ethnicity (Hispanic/Latino, not Hispanic/Latino), veteran status (veteran, not a veteran), control of sample member's institution (public, private non-profit, private for-profit), level of sample member's institution (less than 2 year, 2 to less than 4 year, 4 year or higher non-doctoral, 4 year or higher doctoral), whether sample member's course of study is STEM (STEM, not STEM), and undergraduate status (undergraduate, not an undergraduate). These variables were obtained from enrollment lists. In cases where data from the enrollment lists were missing, we replaced missing values with sample members' substantive answers to the NPSAS:24 survey, where available.

The model also included three variables controlling for design features of the survey: the sample member's data collection wave, whether the sample member was assigned for CATI calling, and time of day that reminder emails were sent to the sample member. The pseudo R-square for the final model was 0.22.

We used this model to estimate predicted probabilities of NPSAS:24 survey response for each category of each of our nine sample member characteristics, holding all other variables at their means. Table 5 below displays these predicted probabilities, along with response rates for each subgroup as of August 11, 2024[22].

We identified four subgroups that had lower response rates and/or propensities –*institution level* of less-than-2 years (low response propensity of 0.43 and low response rate of 36.6 percent), *institution level* of 2-years but less-than-4 year institutions (low response propensity of 0.45 and low response rate of 38.6 percent), *control* of private for-profit (low response rate of 38.0 percent), and *undergraduates* (low response propensity of 0.46 and low response rate of 43.2 percent).

As the *institution level* of 2-year but less-than-4 year institutions, *control* of private for-profit, and *undergraduates* groups have large sample sizes, we conducted further analyses to investigate differences in nonresponse across our other substantive predictors. We began these additional analyses by first estimating a binary logistic regression model predicting the probability of a NPSAS:24 survey response for respondents, partial respondents, and nonrespondents in Waves 1-7 of data collection, subset to sample members whose *institution level* is 2-years but less-than-4[23]. We used the same substantive predictors and controls as the previous model, excluding the institution level, the variable driving the subset. The pseudo R-square for the final model was 0.17. We then used this model to estimate predicted probabilities of NPSAS:24 survey response for each category of each of our seven sample member characteristics, holding all other variables at their means. Table 6 below displays these predicted probabilities, along with response rates for each subgroup as of August 11, 2024. This new analysis identified one additional subgroup of interest for sample members whose institutional level is 2-years but less-than-4 year institutions: sample members whose *gender* is male (low response rate of 35.6 percent; 9 percentage points lower than females).

Similarly, as the *control* of private for-profit and *undergraduates* groups have large sample sizes, we again conducted further analyses to investigate differences in nonresponse across our other substantive predictors. We estimated a binary logistic regression model predicting the probability of a NPSAS:24 survey response for respondents, partial respondents, and nonrespondents in Waves 1-7 of data collection, subset to sample members who are *undergraduates* and whose *control* is private for-profit. We use the same substantive predictors and controls as the first model, excluding the variables driving the subset: undergraduate status and control. The pseudo R-square for the final model was 0.15. We then used this model to estimate predicted probabilities of NPSAS:24 survey response for each category of each of our seven sample member characteristics, holding all other variables at their means. Table 7 below displays these predicted probabilities, along with response rates for each subgroup as of August 11, 2024.

[22] Notably, this date was prior to implementing two data collection changes approved in OMB# 1850-0666 v.38: 1) sample members in Waves 1-3 of data collection receiving a boost offer if identified as eligible for boost, and 2) the start of a contacting materials experiment.
[23] As all sample members in the *institutional level* of two years but less than four years are undergraduates, it was not necessary to explicitly include undergraduate status as a subset.

This final analysis then identified four additional subgroups of interest within undergraduates in the private for-profit control: 1) sample members whose *gender* is male (low response propensity of 0.34 and low response rate of 33.9 percent), 2) sample members who are *veterans* (low response propensity of 0.35 and low response rate of 32.5 percent), 3) sample members whose *age group* is 40 years of age or older (low response propensity of 0.36 and low response rate of 33.2 percent), and 4) sample members who identify as *Hispanic* (low response propensity of 0.36 and low response rate of 37.6 percent).

Given these findings, we recommend offering a $10 promised incentive boost during the remaining waves of data collection for cases belonging to the following groups identified in our analyses: 1) *institution level* is less-than-2 years, 2) *institution level* is 2-years but less-than-4 years and *gender* is male, 3) *undergraduates* whose *control* is private for-profit and *gender* is male, 4) *undergraduates* whose *control* is private for-profit and are *veterans*, 5) *undergraduates* whose *control* is private for-profit and *age group* is 40 years of age or older, and 6) *undergraduates* whose *control* is private for-profit and who identify as *Hispanic*. The purpose of this boost is to encourage participation and reduce the potential for nonresponse bias. This would result in an incentive boost for approximately 22,000[24] nonresponding sample members from Waves 4 – 9. For the rest of the sample member characteristics, response rates and propensities were generally similar across subgroups.

Table 5. Response Rates and Mean Predicted Propensities for Selected Subgroups

Sample member characteristic	Predicted Probability of Survey Response	Standard Error	Response Rate as of 8/11/2024
Gender			
Male	0.55	0.00	44.1%
Female	0.45	0.00	50.1%
Other	0.64	0.02	64.3%
Age			
25 or younger	0.50	0.00	44.4%
26 - 39	0.46	0.00	47.9%
40 or older	0.53	0.01	47.4%
Race			
Races other than White	0.50	0.00	46.6%
White	0.48	0.00	47.9%
Ethnicity			
Not Hispanic	0.49	0.00	48.0%
Hispanic	0.49	0.00	45.4%
Veteran Status			
Not a Veteran	0.49	0.00	45.7%
Veteran	0.47	0.01	44.6%
Control of Institution			
Public Institution	0.48	0.00	44.8%
Private non-profit Institution	0.53	0.00	50.4%
Private for-profit Institution	0.49	0.01	38.0%
Institution Level			
Less-than-2-year	0.43	0.02	36.6%
2-year but less-than-4 year	0.45	0.00	38.6%
4-year or higher non-doctorate granting	0.50	0.00	47.8%
4-year or higher doctorate granting	0.52	0.01	50.3%
STEM Status			
Not in a STEM Program	0.49	0.00	44.9%
In a STEM Program	0.51	0.01	50.8%

[24] As Wave 9 of data collection has not been fielded, this estimate reflects our assumption that 30 percent of the sample members in Wave 9 will respond to the NPSAS:24 survey before the boost is offered.

Sample member characteristic	Predicted Probability of Survey Response	Standard Error	Response Rate as of 8/11/2024
Undergraduate Status			
Not an Undergraduate	0.58	0.00	53.4%
Undergraduate	0.46	0.00	43.2%

Table 6. Response Rates and Mean Predicted Propensities for Selected Subgroups for Sample Members whose Institutional Level is Two Years but Less than Four Years

Sample member characteristic	Predicted Probability of Survey Response	Standard Error	Response Rate as of 8/11/2024
Gender			
Male	0.44	0.01	35.6%
Female	0.37	0.01	44.5%
Other	0.59	0.07	56.3%
Age			
25 or younger	0.41	0.01	38.0%
26 - 39	0.38	0.01	41.4%
40 or older	0.38	0.01	35.9%
Race			
Races other than White	0.42	0.01	41.3%
White	0.36	0.01	38.2%
Ethnicity			
Not Hispanic	0.41	0.01	40.5%
Hispanic	0.38	0.01	39.7%
Veteran Status			
Not a Veteran	0.40	0.00	38.7%
Veteran	0.40	0.02	36.5%
Control of Institution			
Public Institution	0.39	0.00	38.4%
Private non-profit Institution	0.47	0.03	50.9%
Private for-profit Institution	0.42	0.01	39.0%
STEM Status			
Not in a STEM Program	0.40	0.00	38.6%
In a STEM Program	0.62	0.16	71.4%*

*Note that the response rate for this group is high due to a very small sample size in this cell.

Table 7. Response Rates and Mean Predicted Propensities for Selected Subgroups for Undergraduates whose Control is Private For-Profit

Sample member characteristic	Predicted Probability of Survey Response	Standard Error	Response Rate as of 8/11/2024
Gender			
Male	0.34	0.01	33.9%
Female	0.39	0.01	39.7%
Other	0.62	0.16	61.5%
Age			
25 or younger	0.37	0.01	37.1%
26 - 39	0.37	0.01	38.0%
40 or older	0.36	0.02	33.2%
Race			
Races other than White	0.38	0.01	38.1%
White	0.35	0.01	39.1%
Ethnicity			
Not Hispanic	0.38	0.01	39.8%
Hispanic	0.36	0.01	37.6%
Veteran Status			
Not a Veteran	0.37	0.01	37.4%
Veteran	0.35	0.02	32.5%
Institution Level			
Less-than-2-year	0.35	0.02	36.7%
2-year but less-than-4 year	0.38	0.01	39.0%
4-year or higher non-doctorate granting	0.38	0.01	35.6%
4-year or higher doctorate granting	0.12	0.13	16.7%
STEM Status			
Not in a STEM Program	0.37	0.01	36.9%
In a STEM Program	0.41	0.05	38.9%

Identification of Subgroups for Boost Incentive in October 2024

Data from the federal Free Application for Student Aid (FAFSA) is a key data source for NPSAS, providing information submitted by and about students selected for participation in the study, including demographics and information needed to determine aid eligibility. Many important variables in NPSAS are derived from data available only from the FAFSA or the survey. FAFSA data, available through 2023-24 from the Central Processing System (CPS), are also important for obtaining locating information during data collection. In NPSAS:20, about 61 percent of students matched to CPS overall, and the data collection contractor received new address information or confirmed existing information from this CPS match.

To date for NPSAS:24, about 70 percent of the sample submitted FAFSA applications (matched in CPS). Of those who did not submit, about 58 percent have yet to complete the student survey. FAFSA non-filers are a group that is systematically different from FAFSA filers. Anecdotally, a common question asked of NCES by sample members is whether or not their participation in the NPSAS survey is needed since they did not apply for student aid. Our expectation is that this sentiment is common among non-filers. For non-filers, the NPSAS interview becomes the only source for the critical data (e.g., income, dependents) that are missing without a CPS match. Together with the targeted contacting materials described above, we believe a $10 boost offered to non-filers will increase the likelihood of their participation. About 14,900 non-filers have not yet received a separate boost offer.

4. Tests of Procedures or Methods

The NPSAS:24 field test was used to evaluate several new procedures designed to improve data quality and decrease burden on institutions: the collection of enrollment lists from National Student Clearinghouse (described above in section B.2), the bachelor's degree recipient proxy used for identifying potential B&B cohort members (described above in section B.2), and the new pre-formatted enrollment list template file (described above in section B.3). In addition, the student data collection tested three incentive approaches to determine the effects on student participation. Results of these tests are described below.

a. Collection of enrollment lists from National Student Clearinghouse

For a subset of field test institutions, we obtained enrollment list data directly from NSC, in addition to collecting standard NPSAS enrollment lists directly from the 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 8.[25] 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.

Table 8. Comparison of National Student Clearinghouse (NSC) and institution enrollment list counts[1]

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.
NOTE: QC = Quality Control; IPEDS = Integrated Postsecondary Education Data System.
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 9. 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.

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

Table 9. 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

NOTE: NSC = National Student Clearinghouse.
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, 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.

b. Bachelor's degree recipient proxy

As discussed in section B.2, we developed a baccalaureate 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 baccalaureate 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 baccalaureate 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 baccalaureate proxy. Table 10 shows that about 51 percent of students who were both sampled as baccalaureates 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 49 percent (100-51).

Table 10. Baccalaureate status determination, by student type[1]

Student type	Students surveyed in field test	Confirmed B&B eligibility	
		Number	Unweighted percent
Total	**3,470**	**1,260**	**36.4**
Total undergraduate	**3,270**	**1,260**	**38.4**
Potential Baccalaureate	2,320	1,240	53.5
Identified on student lists[2]	200	170	84.7
Identified by proxy definition	2,110	1,070	50.5
Other undergraduate	950	20	1.7
Graduate	200	#	2.0

Rounds to zero.
[2] Identified on student lists using baccalaureate 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: B&B = Baccalaureate and Beyond Longitudinal Study. 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 11 shows that the baccalaureate proxy matches the expected baccalaureate flag from the enrollment lists for about 61 percent of the surveyed students. The baccalaureate proxy correctly identifies about 67 percent of baccalaureate recipients, as compared to the survey, and the expected baccalaureate flag from the enrollment lists correctly identifies about 80 percent of baccalaureate recipients, as compared to the survey.

At the end of field test student data collection, we began a final evaluation of the baccalaureate proxy, including determining if different models would perform better, based on field test data. Given the equivocal results, we plan to use the proxy to supplement, not replace, institutions reporting an unknown degree status.

Table 11. Counts and percentages of surveyed students by baccalaureate status[1]

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	760	23.3
Yes	Yes	No	310	9.5
Yes	No	Yes	10	0.3
Yes	No	No	10	0.3
No	Yes	Yes	210	6.4
No	Yes	No	840	25.7
No	No	Yes	130	3.9
No	No	No	1,000	30.7

[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.

c. Incentive Experiment

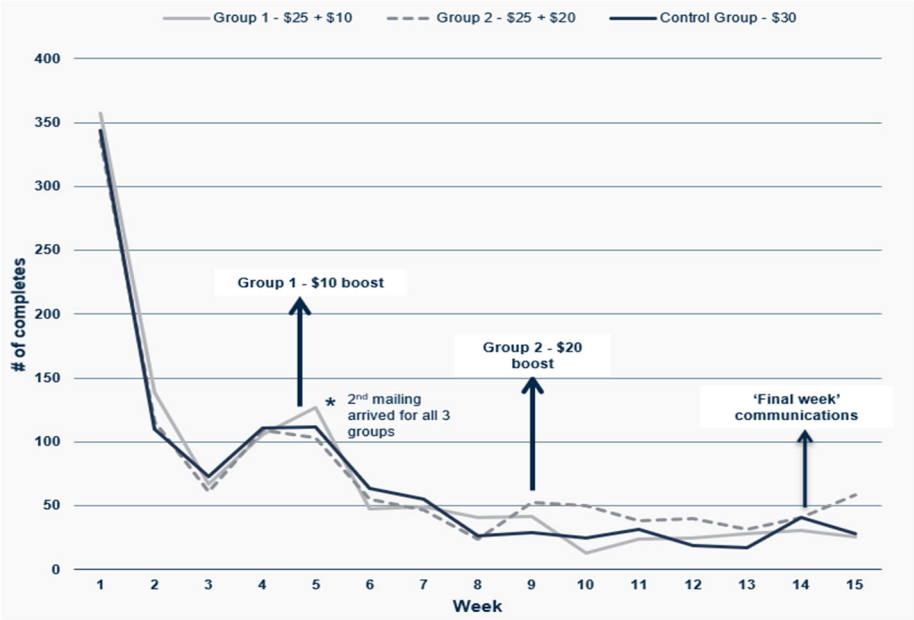
During the field test, we conducted an experiment to determine the optimal amount and delivery timing of incentives, given fixed overall data collection costs (e.g., the costs of incentives, nonresponse follow-up). The field test sample of 6,000 students was divided at random into three experimental groups. Group 1 received an initial incentive offer of $25, followed by an offer of an additional $10 to those who remained nonrespondents after four weeks of data collection (i.e., a boosted incentive). Group 2 received an initial offer of $25 as well, but with an additional $20 boost offered to nonrespondents at the eight-week point. Group 3, a control group, received an incentive offer of $30 throughout the entire period of data collection. We found no significant advantage of the incentive boost or timing of the boost over the control condition ($30 promised incentive). We also found no differences across key demographics across the experimental groups; nor cost per complete advantages (see Appendix I for detailed results).

As indicated in section B.3., sample members will be offered $30 for a completed survey, in line with the control condition from the field test incentive experiment. While the two experimental groups in the NPSAS:24 field test, which initially offered sample members $25, achieved numerically higher response rates (58.5 percent for the +$10 boost and 58.9 percent for the +$20 boost relative to the 57.0 percent for the $30 control condition), a $25 incentive without a boost was not tested. Figure 2 provides some limited insight into the performance of the $25 incentive before a boost is introduced. While response rates were fairly

comparable to that of the $30 control group, the $25 groups did not show the same sustained level of participation except after the offer of the boost.

The field test sample size, originally specified as 3,500, was increased to 6,000 sample members to accommodate the incentive experiment with 3 equally sized groups of 2,000 but, due to cost constraints, could not be increased further to detect smaller differences in groups. The design yielded sufficient power to detect differences in response rates across conditions of no less than 4 percentage points at alpha=0.05 and 80% power; thus, we failed to reject the null hypothesis of no difference across conditions with observed differences of 1.5 and 1.9 percentage points. Consequently, we recommend using the control condition incentive amount for the NPSAS:24 data collection but monitoring data collection throughout for groups that are underrepresented in the respondent pool and may benefit from an incentive boost. As noted in part A.9, our experience on NPSAS:20

Figure 2. NPSAS:24 field test completes, by experimental group



showed that a small boost of $10 can yield an overall increase in response rates, over projected response rates, of 17.5 percent. By data collection wave, the boost resulted in increases over expected rates ranging from 8.05 percent to 25.19 percentage points.[26]

d. Modifications to Student Follow-up Emails

A set of follow-up emails was provided in appendix J, Student Data Collection Materials, of the NPSAS:24 student data collection forms clearance package (OMB#1859-0666 v. 37). While effective in increasing the likelihood of participation, particularly on or after the date sent, a recent experiment as part of a National Science Foundation survey, found that shorter (140 word) emails to follow up with nonrespondents were more likely to elicit a response than longer (212 word) emails.[27] After the first 4 weeks of data collection, response rates were observed to be statistically significantly higher for those who had received the shorter email (18 percent; $p \leq 0.05$) than the longer email (16 percent) after the first reminder. According to the

[26] Memorandum summarizing results submitted by NCES to OMB on April 20, 2021.
[27] Bowman, M., Bryant, A., Griffiths, R., Hare, A., Huey, L., McCall, J., Scanlon, J., and Wakar, B. (2023). *Assessment of Stakeholder Experiences with NSF's Merit Review Process: Findings form the 2021 Merit Review Survey.* Alexandria, CA: National Science Foundation.

authors, this 2 percent difference was maintained through to the end of data collection although at the $p \leq 0.075$ level.

Given this finding, we have created a set of shortened NPSAS:24 nonrespondent reminder emails, removing all but the most essential text from reminder emails 1, 2, 3, 4, and 8. Both the original and shorten version are in ~~Appendix~~appendix J (e.g., see Reminder email 1_original, Reminder email 1_shorten)... Nonrespondents within the last or last two (depending on timing of approval) data collection waves will be split at random into two groups with one receiving the original, longer follow-up emails (the Control group) and the other receiving the new, shorter emails (the Experimental group). Participation rates will be compared immediately before the next follow-up email is scheduled to be sent for the wave.

The proposed experimental design will allow us to test the null hypothesis that there is no statistically significant difference in participation rates between the Control and the Experimental groups (effect of shorter email).

In addition, appendix J was revised to (1) add holiday-themed contacts, including a greeting card, emails, and texts for Thanksgiving and the winter holidays; (2) update contact materials for the end of data collection; and (3) add contact materials to address relevance of participation to federal aid application non-filers.

5. 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.