

# 2019-20 NATIONAL POSTSECONDARY STUDENT AID STUDY (NPSAS:20)

INSTITUTION CONTACTING AND LIST COLLECTION

Supporting Statement Part B

OMB # 1850-0666 v.23

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

**April 2019  
revised June 2019**

# Contents

<b>B. Collection of Information Employing Statistical Methods.....</b>	<b>1</b>
1. Respondent Universe.....	1
a. Institution Universe.....	1
b. Student Universe.....	1
2. Statistical Methodology.....	2
a. Institution Sample.....	2
b. Student Sample.....	5
3. Methods for Maximizing Response Rates.....	13
a. NPSAS:20 Institutional Contacting.....	13
b. Matching to Administrative Databases.....	15
c. Postsecondary Data Portal (PDP).....	15
4. Tests of Procedures or Methods.....	16
5. Reviewing Statisticians and Individuals Responsible for Designing and Conducting the Study.....	16

## Tables and Figures

Table 1.....Preliminary number of institutions to be sampled, by state .....	4
Table 2.....Preliminary number of institutions to be sampled, by control and level of institution.....	5
Table 3.....Potential first-time beginners' false positive rates, by source and control and level of institution: 2011-12.....	8
Table 4.....First-time beginner status determination, by student type: 2011-12 .....	9
Table 5...First-time beginner false-positive rates, by control and level of institution: 2011- 12.....	9
Table 6.....Preliminary graduate student sample sizes, by control and level of institution .....	9
Table 7.. Preliminary undergraduate student interview sample sizes, by control and level of institution.....	10
Table 8.....Preliminary undergraduate student sample sizes, by state .....	11
Table 9.. Preliminary first-time beginning student (FTB) sample sizes, by control and level of institution.....	12
Figure 1..... Institution contacting .....	14

## **B. Collection of Information Employing Statistical Methods**

This submission requests clearance for the 2019–20 National Postsecondary Student Aid Study (NPSAS:20) institution contacting, enrollment list collection, list sampling, and administrative matching activities. The National Center for Education Statistics (NCES) will submit a separate request for the student data collection, including student record data abstraction and student interviews, in the summer of 2019.

### **1. Respondent Universe**

#### **a. Institution Universe**

NPSAS:20 will be nationally representative for both undergraduate and graduate students and state-representative for undergraduate students overall and in public 2-year and 4-year institutions. NPSAS:20 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, NPSAS:20 will serve as the base year for the 2020 cohort of the Beginning Postsecondary Students (BPS) Longitudinal Study and will include a nationally-representative sample of first-time beginning students (FTBs). To construct the full-scale institution sampling frame for NPSAS:20, we will use 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:20 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:20, an institution must do the following during the 2019–20 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<sup>1</sup>; and
- have a signed Title IV participation agreement with the U.S. Department of Education.<sup>2</sup>

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

---

<sup>1</sup> 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.

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

institution in any term or course of instruction between July 1, 2019 and April 30, 2020 and who are:

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

## 2. Statistical Methodology

### a. Institution Sample

The NPSAS:20 institution sampling frame will be constructed from the Integrated Postsecondary Education Data System (IPEDS) 2018-19 Institutional Characteristics Header, 2018-19 Institutional Characteristics, 2017-2018 12-Month Enrollment, and 2017 Fall Enrollment files.<sup>4</sup> Freshening the institution sample will not be needed because we will be using the most up-to-date institution frame available. It is possible that some for-profit institutions and large chains of for-profit institutions may have been closed or sold after the latest IPEDS data collection. We will take this into account in the sample design by using all available resources, such as conducting web searches for articles about closed institutions, to identify these closed for-profit institutions. When using IPEDS to create the sampling frame, we will identify and exclude institutions that are still in IPEDS but are no longer eligible for NPSAS:20 due to closure. 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 data using the latest IPEDS imputation procedures to guarantee complete data for the frame.<sup>5</sup>

The institution strata will be the following three sectors within each state and territory, for a total of 156 (52 x 3) sampling strata:

- public 2-year;
- public 4-year;<sup>6</sup> and
- all other institutions, including:
  - public less-than-2 year;
  - private nonprofit (all levels); and
  - private for-profit (all levels).

The estimated institution sample sizes by the 156 institution strata are presented in table 1.<sup>7</sup> The sample sizes presented in table 1 will allow us to have state-

---

<sup>3</sup> 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 <http://www.gedtesting.com/ged-testing-service>.

<sup>4</sup> A preliminary sampling frame has been created using data from the prior year IPEDS files, and population estimates in the sample size tables in the appendix are based on this preliminary frame. The frame will be recreated with the most up-to-date IPEDS data prior to sample selection.

<sup>5</sup> See <https://nces.ed.gov/pubs2018/2018195.pdf> for further detail on imputation in IPEDS.

<sup>6</sup> The public 4-year institution stratum includes all eligible institutions that IPEDS classifies as public 4-year institutions, including those that are non-doctorate-granting, primarily sub-baccalaureate institutions.

<sup>7</sup> All sample sizes are preliminary and will be updated in fall 2019 once more recent IPEDS and NPSAS:18-AC data are available to evaluate sample sizes for a national- and state-representative study. The final numbers may affect how the sample distributes over states and strata, but are not expected to affect the overall estimated numbers of institutions and students to be sampled, nor the estimated response burden.

representative<sup>8</sup> undergraduate student samples for public 2-year and public 4-year institutions as well as overall. The sample will be nationally representative for both undergraduate and graduate students. Table 2 shows the approximate distribution of the sample sizes across control and level of institution:

- Public less-than-2-year;
- Public 2-year;
- Public 4-year, non-doctorate-granting, primarily sub-baccalaureate;
- 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.

We will select a total of 3,106 institutions which will include a census of all public 2-year and all public 4-year institutions and a sample of 1,381 institutions from the “all other institutions” stratum. Based on NPSAS:18-AC, we expect about a 99 percent eligibility rate, an 85 percent rate for provision of student enrollment lists, and a 93 percent rate for provision of student records among institutions providing lists. This will yield approximately 2,614 enrollment lists, and student records from 2,431 institutions. Within the “all other institutions” stratum, our goal is to sample at least 30 institutions per state so that institutions in this stratum will be sufficiently represented within the state and national samples. We propose using the following criteria from NPSAS:18-AC to determine institution sample sizes within the “all other institutions” stratum:

1. In states with 30 or fewer institutions in the “all other institutions” strata, we will take a census of these institutions.
2. In states with more than 30 institutions in the “all other institutions” strata and where selecting only 30 institutions would result in a very high sampling fraction, we will take a census of institutions. We have arbitrarily chosen 36 institutions as the cutoff to avoid high sampling fractions. This cutoff will result in taking a census of institutions in states that have between 31 and 36 institutions in the “all other institutions” strata.<sup>9</sup>
3. In states with more than 36 institutions in the “all other institutions” strata, we will sample 30 of these institutions.

Within the “all other institutions” stratum, we propose selecting institutions using a variation of probability proportional to size (PPS) sampling called sequential probability minimum replacement (PMR) sampling.<sup>10</sup> 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.<sup>11</sup> Instead of the PMR sampling algorithm selecting some institutions multiple

<sup>8</sup> From this point forward, the word “state” will refer to the 50 states, the District of Columbia, and Puerto Rico.

<sup>9</sup> Based on the latest IPEDS data, there are only three states (Mississippi, Nebraska, and Nevada) that have between 31 and 36 institution in the “other” stratum and will be affected by this cutoff.

<sup>10</sup> 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.

<sup>11</sup> [https://support.sas.com/documentation/cdl/en/statug/63347/HTML/default/viewer.htm#statug\\_surveysselect\\_a0000000173.htm](https://support.sas.com/documentation/cdl/en/statug/63347/HTML/default/viewer.htm#statug_surveysselect_a0000000173.htm).

times, prior to the PMR sample selection, we will set aside for inclusion with certainty in the sample 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 are adjusted accordingly prior to PMR selection, so that the total institution sample size target is met. A composite size measure<sup>12</sup> will be used to help achieve self-weighting samples<sup>13</sup> for student-by-institution strata and to allow flexibility to change sampling rates in selected strata without losing the self-weighting attribute of the sampling method. Institution composite measures of size will be determined using undergraduate and graduate student enrollment counts and FTB counts from the most recent IPEDS 12-Month Enrollment and Fall Enrollment files, respectively.

Within the “all other institutions” stratum, additional implicit stratification will be accomplished by sorting the sampling frame by the following classifications, as appropriate:

1. Control and level of institution;
2. Historically Black Colleges and Universities (HBCUs) indicator;
3. Hispanic-serving institutions (HSIs) indicator;<sup>14</sup>
4. Carnegie classifications of postsecondary institutions;<sup>15</sup> and
5. The institution measure of size.

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

---

<sup>12</sup> 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.

<sup>13</sup> Self-weighting samples have equal weights within sampling domains.

<sup>14</sup> 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/ideshsi/definition.html>) and using IPEDS Hispanic enrollment data.

<sup>15</sup> We will decide what, if any, collapsing is needed of the categories for the purposes of implicit stratification.

**Table 1. Preliminary number of institutions to be sampled, by state**

State	Number of institutions							
	Public 2-year		Public 4-year		Other sectors		All Institutions	
	Population estimate	Sample size	Population estimate	Sample size	Population estimate	Sample size	Population estimate	Total sample size
<b>Total</b>	<b>974</b>	<b>974</b>	<b>751</b>	<b>751</b>	<b>4,836</b>	<b>1,381</b>	<b>6,561</b>	<b>3,106</b>
Alabama	26	26	14	14	47	30	87	70
Alaska	0	0	4	4	5	5	9	9
Arizona	20	20	10	10	85	30	115	60
Arkansas	22	22	11	11	47	30	80	63
California	105	105	49	49	504	30	658	184
Colorado	13	13	17	17	78	30	108	60
Connecticut	13	13	10	10	51	30	74	53
Delaware	0	0	3	3	13	13	16	16
District of Columbia	0	0	2	2	20	20	22	22
Florida	32	32	42	42	280	30	354	104
Georgia	24	24	29	29	106	30	159	83
Hawaii	6	6	4	4	14	14	24	24
Idaho	4	4	4	4	28	28	36	36
Illinois	48	48	12	12	200	30	260	90
Indiana	1	1	14	14	96	30	111	45
Iowa	16	16	3	3	68	30	87	49
Kansas	25	25	8	8	46	30	79	63
Kentucky	16	16	8	8	72	30	96	54
Louisiana	15	15	17	17	89	30	121	62
Maine	7	7	8	8	22	22	37	37
Maryland	16	16	13	13	55	30	84	59
Massachusetts	16	16	15	15	140	30	171	61
Michigan	25	25	21	21	122	30	168	76
Minnesota	31	31	12	12	63	30	106	73
Mississippi	15	15	8	8	33	33	56	56
Missouri	17	17	13	13	140	30	170	60
Montana	10	10	7	7	14	14	31	31
Nebraska	9	9	7	7	32	32	48	48
Nevada	1	1	6	6	33	33	40	40
New Hampshire	7	7	6	6	27	27	40	40
New Jersey	19	19	13	13	119	30	151	62
New Mexico	19	19	9	9	19	19	47	47
New York	38	38	43	43	356	30	437	111
North Carolina	59	59	16	16	100	30	175	105
North Dakota	5	5	9	9	15	15	29	29
Ohio	33	33	35	35	233	30	301	98
Oklahoma	24	24	17	17	85	30	126	71
Oregon	17	17	9	9	60	30	86	56
Pennsylvania	17	17	45	45	294	30	356	92
Puerto Rico	5	5	5	5	121	30	131	40
Rhode Island	1	1	2	2	18	18	21	21
South Carolina	20	20	13	13	70	30	103	63
South Dakota	5	5	7	7	17	17	29	29
Tennessee	39	39	10	10	119	30	168	79
Texas	61	61	47	47	296	30	404	138
Utah	4	4	7	7	60	30	71	41
Virginia	24	24	16	16	118	30	158	70
Washington	8	8	35	35	67	30	110	73
West Virginia	11	11	13	13	49	30	73	54
Wisconsin	17	17	17	17	69	30	103	64
Wyoming	7	7	1	1	2	2	10	10

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

SOURCE: Population estimates based on IPEDS 2017-2018 data.

**Table 2. Preliminary number of institutions to be sampled, by control and level of institution**

Control and level of institution	Population estimate	Sample size
<b>Total</b>	<b>6,561</b>	<b>3,106</b>
Public less-than-2-year	235	35
Public 2-year	974	974
Public 4-year, non-doctorate-granting, primarily sub-baccalaureate	151	151
Public 4-year, non-doctorate-granting, primarily baccalaureate	225	225
Public 4-year, doctorate-granting	375	375
Private nonprofit, less-than-4-year	228	35
Private nonprofit, 4-year, non-doctorate-granting	954	390
Private nonprofit, 4-year, doctorate-granting	701	336
Private for-profit, less-than-2-year	1,454	185
Private for-profit, 2-year	788	205
Private for-profit, 4-year	476	195

NOTE: Details may not sum to totals due to rounding.  
 SOURCE: Population estimates based on IPEDS 2017-2018 data.

### **b. Student Sample**

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

#### **Student Enrollment List Collection**

To begin NPSAS data collection, sampled institutions are asked to provide a list of all their NPSAS-eligible undergraduate and graduate students enrolled in the targeted academic year, covering July 1 through June 30 (methods for contacting the sampled institutions are described below in section B.3, and student list data elements are described in appendix D). Since NPSAS:04, institutions have been asked to limit listed students to only those enrolled through April 30. This truncated enrollment period excludes students who first enrolled in May or June, but it allows lists to be collected earlier and, in turn, 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.

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. This results in students from these institutions having less time in data collection and potentially lower interview response rates than other students. For institutions with continuous enrollment, we will change the endpoint of enrollment from April 30 to March 31 to receive their enrollment lists earlier, allowing more time for student data collection. We conducted research using NPSAS:16 data and concluded that we will not significantly harm representation of the target population by excluding students who enroll in continuous enrollment institutions in April for the first time during the academic year. Again, any lack of coverage resulting from the truncated enrollment period will be accounted for by the poststratification weight adjustment.

#### **Student Stratification**

The student sampling strata will be:



1. undergraduate students who are potential FTBs;<sup>16</sup>
2. other undergraduate students;
3. graduate students who are veterans;
4. master's degree students in science, technology, engineering, and mathematics (STEM) programs;
5. master's degree students in education and business programs;
6. master's degree students in other programs;
7. doctoral-research/scholarship/other students in STEM programs;
8. doctoral-research/scholarship/other students in education and business programs;
9. doctoral-research/scholarship/other students in other programs;
10. doctoral-professional practice students; and
11. other graduate students.

To be comparable to NPSAS:16 and NPSAS:18-AC, we are keeping the graduate strata similar to the sampling strata used in those studies.

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:

- undergraduate students who are potential FTBs;
- graduate students who are veterans;
- master's degree students in STEM programs;
- doctoral-research/scholarship/other students in STEM programs; and
- master's degree students enrolled in for-profit institutions.

Similarly, we anticipate the following groups will be undersampled:

- master's degree students in education and business programs; and
- doctoral-research/scholarship/other students in education and business programs.

Because these two groups are so large, sampling in proportion to the population would make it difficult to draw inferences about the experiences of other master's degree and doctoral students, respectively.

As was done for NPSAS:16 and NPSAS:18-AC, we will match the student enrollment lists to two supplemental databases prior to sampling (pre-sampling matching). To identify veterans, we will match the student enrollment lists with a list of veterans from the Veterans Benefits Administration (VBA) because the veterans identified by institutions on the lists are incomplete. This veterans information will be used with the veteran status from the enrollment lists to explicitly stratify graduate students and implicitly stratify undergraduate students. As in NPSAS:18-AC, the undergraduate students who are veterans will not be oversampled within each state because that would require too large of a total sample size. The implicit stratification will allow the sample proportions of veterans to approximately match the population within institution and student strata, which will ensure that we have enough undergraduate veterans in the sample for analytic purposes. We will also match the student lists to the National Student Loan Data System (NSLDS) data and use the financial aid data for student-implicit stratification. Within the student-explicit strata for graduate students

---

<sup>16</sup> If a decision is made to oversample any subgroup of FTBs for BPS, then that subgroup would be a separate stratum. The final sampling plan will be provided in the student data collection submission in the summer of 2019.

and veterans implicit strata for undergraduate students, 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 FTBs**

As mentioned in section 1a, NPSAS:20 will serve as the base year for the 2020 cohort of BPS and will include a nationally-representative sample of FTBs, hence the stratification described above. Accurately qualifying sample members as FTBs is a continuing challenge. Correctly classifying FTBs is important because unacceptably high rates of misclassification (i.e., false positives and false negatives) can and have resulted in: (1) excessive cohort loss with too few eligible sample members to sustain the longitudinal study, (2) excessive cost to “replenish” the sample with little value added, and (3) inefficient sample design (excessive oversampling of “potential” FTBs) to compensate for anticipated misclassification error.

In NPSAS:04, the FTB false-positive and false negative rates were 53.5 and 25.3 percent, respectively, because institutions have difficulty identifying FTBs on the student enrollment lists. In NPSAS:12 (the next NPSAS after NPSAS:04 and the last NPSAS prior to NPSAS:20 to spin off a BPS cohort), we greatly improved the identification of FTBs from what was provided on the student enrollment lists, and we will take several steps early in the NPSAS:20 listing and sampling processes to similarly improve the rate at which FTBs are correctly classified for sampling. First, in addition to an FTB indicator, we will request that enrollment lists provided by institutions (or institution systems) include degree program, class level, date of birth, dual enrollment in high school indicator, and high school completion date. Students identified by the institution as FTBs, but also identified as in their third year or higher and/or not an undergraduate student, will not be classified as FTBs for sampling. Additionally, students who are dually enrolled at the postsecondary institution and in high school based on the enrollment in high school (or completion program) indicator and the high school graduation date will not be eligible for sampling. If the FTB indicator is not provided for a student on the list but the student is 18 years old or younger and does not appear to be dually enrolled, the student will be classified as an FTB for sampling. Otherwise, if the FTB indicator is not provided for a student on the list and the student is over the age of 18, then the student will be sampled as an “other undergraduate,” but will be part of the BPS cohort if identified during the interview as an FTB.

Second, prior to sampling, we will match all students listed as potential FTBs to NSLDS records to determine if any have a federal financial aid history pre-dating the NPSAS year (earlier than July 1, 2019). Since NSLDS maintains current records of all Title IV grant and loan funding, any students with data showing disbursements from prior years can be reliably excluded from the sampling frame of FTBs. Given that about 68 percent of FTBs receive some form of Title IV aid in their first year, this matching process will not be able to exclude all listed FTBs with prior enrollment but will significantly improve the accuracy of the listing prior to sampling, yielding fewer false positives. All potential FTBs will be sent to NSLDS because 11 percent of students 18 and younger sampled as FTBs and interviewed in NPSAS:12 were not FTBs (false positives). In NPSAS:12, matching to NSLDS identified about 20 percent of the cases sent for matching as false positives (see table 3). NPSAS:12 showed that it is feasible to send all potential FTBs to NSLDS for matching. NSLDS has a free process to match the FTBs, and lists were usually returned to us in one day.

Third, simultaneously with NSLDS matching, we will match all potential FTBs to the Central Processing System (CPS) to identify students who, on their Free Application for Federal Student Aid (FAFSA), indicated that they had attended college previously. In NPSAS:12 we identified about 17 percent of the cases sent for CPS matching as false positive (see table 3). CPS has an automated, free process for matching that we have used in NPSAS:12 for this purpose, as well as for other purposes in the past for NPSAS sample students. This matching can handle large numbers of cases, and the matching usually takes one day.

Fourth, after NSLDS and CPS matching, we will match a subset of the remaining potential FTBs to the National Student Clearinghouse (NSC) for further narrowing of FTBs based on the presence of evidence of earlier enrollment. In NPSAS:12, matching to NSC identified about seven percent of the remaining potential FTBs, after NSLDS and CPS matching, as false positives. NSC worked with us to set up a process that can handle a large number of potential FTBs and return FTB lists to us within two or three days. There is a “charge per case matched” for NSC matching, so we plan a targeted approach to the matching. We plan to target potential FTBs over the age of 18 in the public 2-year and for-profit sectors because these sectors had high false-positive rates in NPSAS:12 and have large NPSAS:20 sample sizes.

Fifth, in setting our FTB selection rates, we will take into account the false-positive rates, based on the NPSAS:12 interview, as shown in table 4 (overall, with details) and table 5 (by control and level of institution). In NPSAS:12, of the 36,620 interview respondents sampled as potential FTBs, the interview confirmed that 28,550 were FTBs, for an unweighted false positive rate of 22 percent (100 percent minus 78 percent). Conversely, of the 48,380 interview respondents sampled as other undergraduate or graduate students, about 1,590 were FTBs, for a false negative rate of 4.6 percent unweighted. With the help of the presampling matching, the NPSAS:12 overall false-positive rate of 22 percent was much less than the 53.5 percent false positive rate in NPSAS:04, when pre-sampling matching was not conducted. The false-negative rate is small, but we will also account for it when setting the FTB selection rates.

**Table 3. Potential first-time beginners' false positive rates, by source and control and level of institution: 2011–12**

Control and level of institution	Total			Source								
	Sent for matching	False positives	Percent false positive	NSLDS			CPS			NSC		
				Sent for matching	False positives	Percent false positive	Sent for matching	False positives	Percent false positive	Sent for matching	False positives	Percent false positive
<b>Total</b>	<b>2,103,620</b>	<b>571,130</b>	<b>27.1</b>	<b>2,103,620</b>	<b>417,910</b>	<b>19.9</b>	<b>2,103,620</b>	<b>364,350</b>	<b>17.3</b>	<b>719,450</b>	<b>48,220</b>	<b>6.7</b>
Public												
Less-than-2-year	3,690	2,030	54.9	3,690	1,720	46.5	3,690	1,520	41.2	†	†	†
2-year	816,150	276,500	33.9	816,150	188,630	23.1	816,150	153,150	18.8	584,950	45,300	7.7
4-year, non-doctorate-granting	194,600	26,500	13.6	194,600	17,180	8.8	194,600	18,010	9.3	†	†	†
4-year, doctorate-granting	517,380	53,870	10.4	517,380	28,000	5.4	517,380	42,840	8.3	†	†	†
Private nonprofit												
Less-than-4-year	2,570	1,020	39.6	2,570	750	29.0	2,570	640	24.8	†	†	†
4-year, non-doctorate-granting	106,800	18,860	17.7	106,800	13,880	13.0	106,800	15,830	14.8	†	†	†
4-year, doctorate-granting	152,450	13,940	9.1	152,450	8,680	5.7	152,450	11,850	7.8	†	†	†
Private for-profit												
Less-than-2-year	16,800	9,820	58.4	16,800	8,800	52.4	16,800	4,940	29.4	7,110	130	1.8
2-year	69,070	42,980	62.2	69,070	37,920	54.9	69,070	29,730	43.0	26,770	680	2.5
4-year	224,110	125,610	56.0	224,110	112,370	50.1	224,110	85,850	38.3	100,620	2,120	2.1

† Not applicable.

NOTE: NSLDS = National Student Loan Data System; NSC = National Student Clearinghouse; and CPS = Central Processing System. Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2011–12 National Postsecondary Student Aid Study (NPSAS:12).

**Table 4. First-time beginner status determination, by student type: 2011–12**

Student type	Students interviewed	Confirmed FTB eligibility	
		Number	Unweighted percent
<b>Total</b>	<b>85,000</b>	<b>30,140</b>	<b>35.5</b>
Total undergraduate	71,000	30,140	42.4
Potential FTB	36,620	28,550	78.0
FTB in certificate program	10,900	7,670	70.3
Other FTB	25,720	20,880	81.2
Other undergraduate	34,380	1,580	4.6
Graduate	14,000	10	#

# Rounds to zero.

NOTE: *Students interviewed* includes all eligible sample members who completed the interview. FTB = first-time beginner. Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2011–12 National Postsecondary Student Aid Study (NPSAS:12).

**Table 5. First-time beginner false-positive rates, by control and level of institution: 2011–12**

Control and level of institution	FTB false positive rate
<b>Total</b>	<b>22.0</b>
Public less-than-2-year	41.6
Public 2-year	23.5
Public 4-year, non-doctorate-granting	11.9
Public 4-year, doctorate-granting	8.8
Private nonprofit, less-than-4-year	24.0
Private nonprofit, 4-year, non-doctorate-granting	11.2
Private not-for-profit, 4-year, doctorate-granting	10.0
Private for-profit, less-than-2-year	31.2
Private for-profit, 2-year	31.2
Private for-profit, 4-year	27.3

NOTE: There were 10 categories of control and level of institution defined for NPSAS:12, instead of 11 in NPSAS:16 and NPSAS:20.

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2011–12 National Postsecondary Student Aid Study (NPSAS:12).

## Sample Sizes and Student Sampling

NPSAS:20 will be designed to sample a total of 400,000 students; 150,000 of which will be asked to complete an interview, and 250,000 of which will not be asked to complete an interview. Student records and administrative data will be collected for all sample students. Based on past rounds of NPSAS, we expect about a 95 percent eligibility rate (among all sampled students, with eligibility determined based on the interview, student records, and administrative data), a 70 percent interview response rate, and a 90 percent student records completion rate.<sup>17</sup> This will yield approximately 100,000 interviews and 342,000 student records, and the average student sample size per institution will be approximately 165 students. We expect to sample 25,000 graduate students, and the remaining sample will be of undergraduate students. All sampled graduate students will be asked to complete an interview, in addition to us collecting student records and administrative data. The preliminary graduate student sample sizes by institution strata are presented in table 6.

**Table 6. Preliminary graduate student sample sizes, by control and level of institution**

Control and level of institution	Population estimate	Sample size
<b>Total</b>	<b>3,875,095</b>	<b>25,000</b>
Public 4-year non-doctorate-granting primarily sub-baccalaureate	1,521	78
Public 4-year non-doctorate-granting primarily baccalaureate	166,519	1,510
Public 4-year doctorate-granting	1,627,687	7,040
Private not-for-profit 4-year non-doctorate-granting	247,948	2,623
Private not-for-profit 4-year doctorate-granting	1,411,067	6,490
Private for-profit 4-year	420,353	7,260

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

SOURCE: Population estimates based on IPEDS 2016-2017 data.

The undergraduate student sample will be both nationally representative and state-representative for public 2-year and public 4-year institutions, as well as overall.

<sup>17</sup> The expected student records completion rate and yield are preliminary pending the completion of NPSAS:18-AC student records collection.

125,000 of the 150,000 students asked to complete an interview will be undergraduates. The preliminary distribution of these 125,000 undergraduate sample students by control and level of institution is shown in table 7. The remaining 250,000 undergraduate students will be sampled only for collecting student records and administrative data. All 375,000 undergraduate sample students will be included in the state-representative sample. For the state-representative sample, we propose initially dividing the undergraduate sample evenly between states (resulting in 7,212 students per state) and proportionally within states to obtain the preliminary sample sizes for each stratum, as shown in table 8. This will allow sufficient precision in each state, but the final sample sizes per state will be determined taking the NPSAS:18-AC data into account.

As part of setting the NPSAS:20 sample sizes, we need to determine the sample size of FTBs, which will be part of both NPSAS and the BPS 2020 cohort. The BPS:20/22 sample size is planned to be about 37,000, including 30,000 FTBs who respond to the NPSAS:20 interview and confirm that they are FTBs, and 7,000 potential FTBs who do not respond to the interview. The NPSAS:20 potential FTB sample size will be approximately 55,400, assuming 95 and 70 percent NPSAS:20 eligibility and interview response rates, respectively, a 22 percent false-positive rate, and a 4.6 false-negative rate, as in NPSAS:12.<sup>18</sup> The preliminary distribution of potential FTBs by control and level of institution is shown in table 9.

During the NPSAS:18-AC Technical Review Panel (TRP) meeting, panel members expressed an interest in being able to create their own groupings of institutions for analysis (i.e. institutions within specific university systems). We expect to sample at least 200 undergraduates, on average, per public 2-year and 4-year institutions. The minimum sample size will vary by institution depending on the strata and enrollment size of the institution. Therefore, the sample size will be sufficient to allow researchers to aggregate institutions for analysis of undergraduate students in public 2-year and public 4-year institutions.

Institution-level student sampling rates will be set based on frame data and adjusted, based on NPSAS:18-AC data, to account for IPEDS data overestimating the enrollment counts for the student lists. Based on these adjusted rates, students will be sampled on a flow basis as student lists are received. Stratified systematic sampling procedures will be used. Within the graduate-student strata for veterans, the students will be sorted by master's and doctoral to ensure that the sample will be roughly proportional to the frame. As mentioned above, undergraduate student strata will be sorted (implicitly stratified) by veteran status and all strata will be sorted by federally aided/unaided students to maintain proportionality between the sample and 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.

After undergraduate students are initially sampled, they will be randomly divided into two groups, within student stratum within institution. The proportion of students in each group within strata within an institution will be determined such that the overall sample sizes of 125,000 undergraduate students for the interview, including 55,400 FTBs, will be achieved. One group will include students who will receive the interview, and the other group will not receive the interview. Both groups will have student records and administrative data.

---

<sup>18</sup>  $55,400 = 30,000 / .95 / .70 / .78 - (53,125 \text{ non-FTBs} * .046)$ , where  $.78 = 1 - .22$ .

**Table 7. Preliminary undergraduate student interview sample sizes, by control and level of institution**

<b>Control and level of institution</b>	<b>Population estimate</b>	<b>Sample size</b>
<b>Total</b>	<b>23,030,788</b>	<b>125,000</b>
Public less-than-2-year	74,141	1,600
Public 2-year	8,724,915	48,600
Public 4-year, non-doctorate-granting, primarily sub-baccalaureate	1,748,376	3,975
Public 4-year, non-doctorate-granting, primarily baccalaureate	1,207,840	6,654
Public 4-year, doctorate-granting	5,828,957	17,706
Private nonprofit, less-than-4-year	73,373	1,600
Private nonprofit, 4-year, non-doctorate-granting	1,408,181	6,946
Private nonprofit, 4-year, doctorate-granting	2,024,218	5,694
Private for-profit, less-than-2-year	350,055	4,518
Private for-profit, 2-year	430,138	10,588
Private for-profit, 4-year	1,160,594	17,119

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

SOURCE: Population estimates based on IPEDS 2016-2017 data.

**Table 8. Preliminary undergraduate student sample sizes, by state**

State	Number of students							
	Public 2-year		Public 4-year		Other sectors		All institutions	
	Population estimate	Sample size	Population estimate	Sample size	Population estimate	Sample size	Population estimate	Total sample size
<b>Total</b>	<b>8,724,915</b>	<b>120,079</b>	<b>8,785,173</b>	<b>154,452</b>	<b>5,520,700</b>	<b>100,493</b>	<b>23,030,788</b>	<b>375,000</b>
Alabama	118,972	2,576	150,217	3,252	63,901	1,384	333,090	7,212
Alaska	0	0	40,288	6,740	2,822	472	43,110	7,212
Arizona	294,365	2,886	164,298	1,611	276,935	2,715	735,598	7,212
Arkansas	67,914	2,688	92,477	3,660	21,856	865	182,247	7,212
California	1,832,697	3,930	1,041,511	2,233	488,875	1,048	3,363,083	7,212
Colorado	93,753	1,630	217,354	3,779	103,694	1,803	414,801	7,212
Connecticut	71,492	2,378	60,184	2,002	85,144	2,832	216,820	7,212
Delaware	0	0	44,705	5,275	16,419	1,937	61,124	7,212
District of Columbia	0	0	5,242	637	54,063	6,575	59,305	7,212
Florida	76,289	412	928,602	5,011	331,492	1,789	1,336,383	7,212
Georgia	163,372	1,991	314,093	3,828	114,261	1,393	591,726	7,212
Hawaii	35,522	3,347	27,163	2,559	13,856	1,306	76,541	7,212
Idaho	37,681	1,593	53,548	2,264	79,382	3,356	170,611	7,212
Illinois	553,121	4,181	153,534	1,161	247,370	1,870	954,025	7,212
Indiana	164,851	2,411	227,092	3,321	101,183	1,480	493,126	7,212
Iowa	134,204	3,142	71,199	1,667	102,617	2,403	308,020	7,212
Kansas	128,435	3,576	86,688	2,414	43,881	1,222	259,004	7,212
Kentucky	108,182	2,848	116,762	3,074	49,017	1,290	273,961	7,212
Louisiana	92,856	2,393	138,895	3,580	48,072	1,239	279,823	7,212
Maine	23,525	2,042	32,031	2,781	27,523	2,389	83,079	7,212
Maryland	172,695	3,171	173,555	3,187	46,507	854	392,757	7,212
Massachusetts	128,297	2,006	116,393	1,820	216,612	3,387	461,302	7,212
Michigan	222,662	2,524	309,533	3,509	104,037	1,179	636,232	7,212
Minnesota	171,168	3,064	132,289	2,368	99,435	1,780	402,892	7,212
Mississippi	98,883	3,666	74,606	2,766	21,021	779	194,510	7,212
Missouri	127,432	2,234	143,740	2,520	140,274	2,459	411,446	7,212
Montana	12,015	1,521	38,989	4,935	5,979	757	56,983	7,212
Nebraska	62,982	3,119	52,224	2,586	30,413	1,506	145,619	7,212
Nevada	15,893	809	109,822	5,589	15,998	814	141,713	7,212
New Hampshire	21,733	1,001	26,369	1,214	108,523	4,997	156,625	7,212
New Jersey	217,050	3,241	172,033	2,569	93,860	1,402	482,943	7,212
New Mexico	103,714	4,509	54,055	2,350	8,108	353	165,877	7,212
New York	433,328	2,324	397,424	2,132	513,935	2,756	1,344,687	7,212
North Carolina	317,005	3,625	203,834	2,331	109,786	1,256	630,625	7,212
North Dakota	9,403	1,178	41,302	5,174	6,863	860	57,568	7,212
Ohio	257,646	2,491	312,564	3,022	175,616	1,698	745,826	7,212
Oklahoma	92,340	2,613	116,046	3,284	46,478	1,315	254,864	7,212
Oregon	160,820	3,875	105,549	2,543	32,953	794	299,322	7,212
Pennsylvania	188,102	1,804	254,964	2,446	308,825	2,962	751,891	7,212
Puerto Rico	3,124	108	18,862	651	186,939	6,453	208,925	7,212
Rhode Island	20,162	1,679	24,573	2,046	41,878	3,487	86,613	7,212
South Carolina	118,283	3,126	99,317	2,624	55,319	1,462	272,919	7,212
South Dakota	8,478	1,000	40,682	4,798	11,988	1,414	61,148	7,212
Tennessee	132,194	2,653	127,126	2,552	100,008	2,007	359,328	7,212
Texas	1,083,113	3,956	668,026	2,440	223,374	816	1,974,513	7,212
Utah	54,097	1,021	166,950	3,150	161,236	3,042	382,283	7,212
Vermont	8,626	1,317	20,109	3,070	18,498	2,824	47,233	7,212
Virginia	244,220	2,943	193,123	2,328	161,045	1,941	598,388	7,212
Washington	56,772	876	359,839	5,550	50,956	786	467,567	7,212
West Virginia	23,505	950	62,221	2,514	92,777	3,748	178,503	7,212
Wisconsin	132,721	2,497	192,297	3,618	58,293	1,097	383,311	7,212
Wyoming	29,221	5,153	10,874	1,918	803	142	40,898	7,212

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

SOURCE: Population estimates based on IPEDS 2016-2017 data.



**Table 9. Preliminary first-time beginning student (FTB) sample sizes, by control and level of institution**

Control and level of institution	Sample size
<b>Total</b>	<b>55,393</b>
Public less-than-2-year	952
Public 2-year	19,683
Public 4-year, non-doctorate-granting, primarily sub-baccalaureate	1,741
Public 4-year, non-doctorate-granting, primarily baccalaureate	2,595
Public 4-year, doctorate-granting	5,131
Private nonprofit, less-than-4-year	1,244
Private nonprofit, 4-year, non-doctorate-granting	3,519
Private nonprofit, 4-year, doctorate-granting	2,258
Private for-profit, less-than-2-year	3,100
Private for-profit, 2-year	4,979
Private for-profit, 4-year	10,191

NOTE: Population estimates will be added once there is a final sampling frame. Details may not sum to totals due to rounding.

## Calibration Sample

NPSAS:20 will include a calibration sample to inform the design regarding incentive structure and nonresponse follow-up strategies in the absence of a field test.<sup>19</sup> The data collected from the calibration sample will be retained and included in final data files with the main sample data.

We will select the student calibration sample in December 2019 from fall enrollment lists, which are provided by institutions selected from among the 3,106 sample institutions. The calibration institution sample size will be approximately 86 institutions to yield 60 participating institutions, assuming a 70 percent calibration sample participation rate and about 100 students sampled per institution from the fall lists, on average. The 86 institutions will be selected purposively from among the full NPSAS:20 sample across the control and level of institution and will include both small and large institutions, as well as systems and individual institutions. A purposive subsample will allow us to target institutions that we have a good relationship with and have a good idea that they would be willing to provide both fall and spring lists.

The calibration sample of students will be selected using the same sampling design as the main sample, with the exception that all students selected from the institutions in the calibration sample will be included in the interview. The interview is needed for the calibration sample experiments, and this sampling approach allows for a smaller calibration institution sample size. Students will be sampled from the fall enrollment lists using the same sampling rates as will be used for the spring lists for these institutions, so that there are not unequal probabilities of selection, and thus unequal weights, within student strata in an institution. The institutions that provide a fall list will be asked to provide another list in the spring that includes either all eligible students or eligible students not included on the fall list. If the former, then the fall and spring lists for an institution will be deduplicated by SSN and, if there is no SSN, by name and date of birth. This will ensure that students will have one chance of selection per institution. Students will be sampled from the spring lists using the same sampling rates as used for the fall lists for these institutions. Ideally, all pre-sampling matching will occur for both the fall and spring enrollment lists to identify FTBs, veterans, and aided/unaided students. However, depending on how quickly institutions can provide fall lists, we may not have time to send data from all institutions to all sources for pre-sampling matching in time to select the calibration student samples.

<sup>19</sup> Details of the calibration sample and its use for determining incentives and follow-up strategies will be provided in the NPSAS:20 Student Data Collection Package in the summer of 2019.

A potential issue with requesting fall and spring enrollment lists is that some of the calibration sample institutions may send a fall list and then decide later to not send a spring list. Those institutions could be treated as nonresponding and their student data using the fall lists could be excluded, that is using an institution nonresponse weighting adjustment. Alternatively, the student data could be included, with the student poststratification weight adjustment accounting for the student undercoverage. The latter approach is preferred, as long as the fall list has sufficient students such that the poststratification can reduce any undercoverage bias. While enrollment lists would be collected twice for the calibration sample institutions, we would only collect student records data once, after students are sampled from the spring lists.

### **Quality Control Checks for Lists and Sampling**

The number of enrollees on each institution's student list will be checked against the latest IPEDS 12-month enrollment. The comparisons will be made for each student level: undergraduate and graduate. Based on experience with past rounds of NPSAS, we recommend that, in order for an institution's student list to pass quality control (QC) and be moved on to student sampling, the student counts must be within 50 percent of non-imputed most recent IPEDS counts.

Institutions that fail QC will be recontacted to resolve the discrepancy and to verify that the institution's campus coordinator who prepared the student list clearly understood our request and provided a list of the appropriate students and data items. 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.

All statistical procedures will undergo thorough quality-control checks. The data collection contractor has a Quality Management Plan (QMP) in place for sampling and all statistical activities. All statisticians will employ a checklist to ensure that all appropriate QC checks are done for student sampling.

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

- institutions and students on the sampling frames all have a known, non-zero probability of selection;
- distribution of implicit stratification for institutions is reasonable; and
- number of institutions and students selected match the target sample sizes.

### **3. Methods for Maximizing Response Rates**

#### **a. NPSAS:20 Institutional Contacting**

Establishing and maintaining contact with sampled institutions throughout the data collection process is vital to the success of NPSAS:20. Institutional participation is required in order to collect enrollment lists and draw the student sample. The process by which institutions will be contacted is depicted in figure 1 and described below.

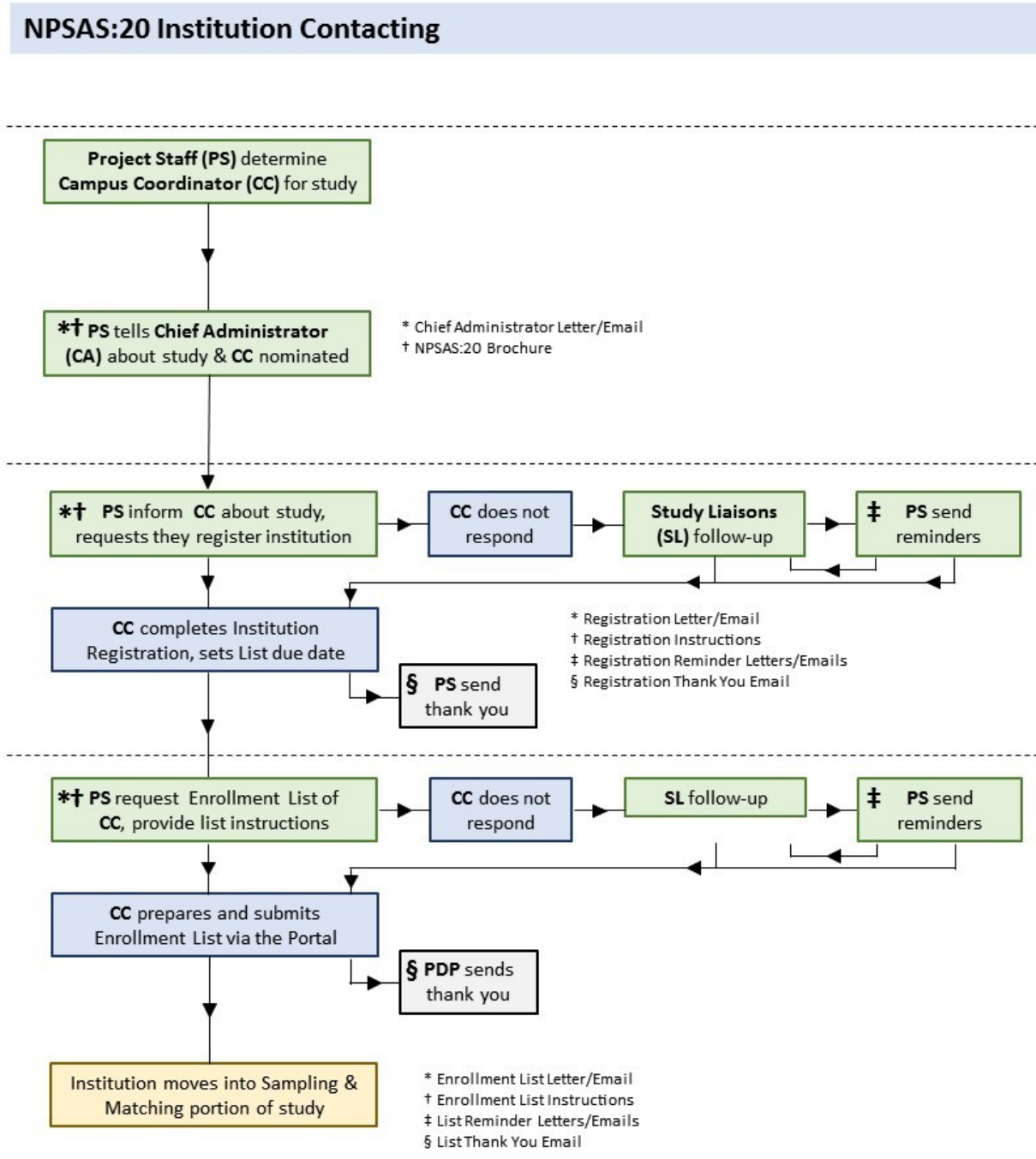
The data collection contractor will be responsible for contacting institutions on behalf of NCES. Each staff member will be assigned a set of institutions that is their responsibility throughout the data collection process. This allows the contractor's staff members to establish rapport with the institution 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 using the Higher Ed Directory (<https://hepinc.com/>). Web searches and verification calls will be conducted as needed (e.g., for institutions not listed in the Directory) to confirm eligibility and confirm contact information obtained from the IPEDS header files before study information is mailed. Once the contact information is verified, we will prepare and send an information packet to the chief administrator of each sampled institution. A copy of the letter and brochure can be found in appendix D. The materials provide information about the purpose of the study and the nature of subsequent requests. In addition to the hardcopy materials, we will send an email to the chief administrator, copying the previous campus coordinator (if still at the institution), the IPEDS Keyholder, and the Director of Institutional Research to make them aware of the NPSAS:20 data collection. Several versions of the chief administrator letter will be used, tailored to the institution's situation: (1) one letter for institutions that were sampled for NPSAS:16, NPSAS:18-AC, the student records (SR) collection for the 2012 Beginning Postsecondary Students Longitudinal Study cohort (BPS:12 SR), and/or the student records collection for the High School Longitudinal Study Second Follow-up (HSLF2 SR), and have an identified campus coordinator; (2) one for new institutions with a campus coordinator candidate identified; and (3) another for new/prior institutions for which a campus coordinator has not been identified. For the last group, institutional 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 state associations and agencies, networking with the higher education community at conferences and professional meetings, and reaching out to state government leaders. These activities are intended to promote the value of NPSAS both to data providers and data users thereby increasing interest and participation in NPSAS:20.

Figure 1. Institution contacting



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

After the Registration Page is completed, the campus coordinator will be sent a letter requesting an electronic enrollment list of all students enrolled during the academic year. The NPSAS:20 data collection includes a calibration sample, described in Supporting Statement Part B, Section 2, and the main sample. The calibration sample

institutions will be providing two enrollment lists, one in the fall and one in the spring, instead of one. The earliest enrollment lists will be due in November 2019 for the calibration sample. For the main sample, enrollment lists will be collected from January 2020 to July 2020. 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 two weeks prior to a deadline, and touch-base emails typically sent when 3-4 weeks have passed with no outbound contact from study staff (see appendix D). After enrollment lists are received and validated by the contractor for completeness and quality, the campus coordinator will be sent a “thank you” email acknowledging appreciation for their time and effort.

### **Alternate Enrollment List Submission Methods**

Two alternate submission methods will be available to campus coordinators who report a lack of time or resources needed to complete the full enrollment list. The first is compiling an enrollment list with a reduced set of critical data elements (See Appendix D, pp D-33 to D-36 for a list of elements). The second is submitting files the institution already compiles for the National Student Clearinghouse (NSC) Enrollment Reporting service. This option will be suggested only to institutions participating in the NSC program. These alternate submission methods are designed to collect data needed for sampling while improving response rates and decreasing burden on the institutions. In the final weeks of the enrollment list data collection period, submitting a further reduced set of data elements (First Name, Last Name, Social Security Number, Undergraduate/Graduate) will be offered to institutions that have not yet participated to maximize response.

### **Spanish Contact Materials**

Select contact materials have been translated into Spanish and will be sent to institution staff at institutions in Puerto Rico. The contact materials include the letters sent to the chief administrator and coordinator as well as the study brochure and the Quick Guide to NPSAS:20 (see appendix D).

#### **b. Matching to Administrative Databases**

Information about NPSAS:20 sampled students will be matched with their data from several administrative databases. The administrative data sources for NPSAS:20 will be NSLDS, CPS including FAFSA, NSC, VBA, ACT and SAT test scores, and student records obtained directly from postsecondary institutions.<sup>20</sup> 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.

#### **c. Postsecondary Data Portal (PDP)**

NPSAS:20 data collection will utilize NCES’ Postsecondary Data Portal (PDP) website. The PDP is used across NCES postsecondary institution data collections. The flexible design of the website allows it to be used for multiple studies that are 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 when NPSAS:20 will begin collecting enrollment lists and student records.

---

<sup>20</sup> Data from the NSC, VBA, ACT, and SAT scores are pending contracts or agreements with those organizations. If NCES is unable to secure an agreement with any of these organizations, a change memo will be submitted by September 2019 for NSC or VBA, and by October 2020 for ACT and SAT.

There are two types of content on the PDP: general-purpose content and study-specific content. General-purpose pages provide overview information about NCES postsecondary studies and use of the website. These pages are identified in appendix D as the “pre-login” pages. Once a user logs in, they see pages with study-specific content. These pages are identified in appendix D as the “after login” content. The NPSAS:20 study-specific content includes FAQs about NPSAS:20 and instructions for providing data (appendix D), and the student records instrument. Institutions see study-specific PDP content only for the study or studies for which they have been sampled.

### **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, entering a code that is 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 user name and will be assigned to one e-mail address. Upon account creation, the new user will be sent a temporary password by the PDP. Upon logging in for the first time, the new user will be required to create a new password. The system automatically will log out the user 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.

### **4. Tests of Procedures or Methods**

The tests of procedures or methods as part of NPSAS:20 will be described in the forthcoming Student Data Collection Package to be submitted in the summer of 2019. These tests include using a calibration sample instead of a field test to inform the design regarding incentive structure and nonresponse follow-up strategies.

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

NPSAS:20 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, Mr. Ted Socha, Dr. Elise Christopher, and Dr. Gail Mulligan. NCES’s prime contractor for NPSAS:20 is RTI International (Contract# 91990018C0039), and subcontractors include Leonard Resource Group; HR Directions; KForce Government Solutions, Inc.; Research Support Services; EurekaFacts; Strategic Communications, Inc.; and Activate Research. Dr. Anthony Jones, Dr. Sandy Baum, and Dr. Stephen Porter are consultants on the study. The following staff members at RTI are working on the statistical aspects of the study design: Mr. Peter Siegel, Dr. Jennifer Wine, Mr. Stephen Black, Mr. Darryl Cooney, and Dr. T. Austin Lacy. Principal professional RTI staff, not listed above, who are assigned to the study include: Ms. Kristin Dudley, Ms. Jamie Wescott, Ms. Tiffany Mattox, Mr. Austin Caperton, Mr. Jeff Franklin, Dr. Nicole Tate, Mr. Johnathan Conzelmann, Dr. Rachel Burns, Mr. Michael Bryan, and Dr. Josh Pretlow.