

2020/22 BEGINNING POSTSECONDARY STUDENTS (BPS:20/22) FULL-SCALE STUDY

Supporting Statement Part B

OMB # 1850-0631 v.19

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

November 2021

Table of Contents

Preface.....	1
Section 1 – Respondent Universe.....	1
Section 2 – Statistical Methodology.....	1
Section 3 – Methods for Maximizing Response Rate.....	7
Section 4 – Tests of Procedures and Methods.....	11
Section 5 – Reviewing Statisticians and Individuals Responsible for Designing and Conducting the Study	18

Preface

This submission requests clearance for the 2020/22 Beginning Postsecondary Students Longitudinal Study (BPS:20/22) full-scale study data collection materials and procedures. BPS:20/22 is the first follow-up of sample members from the 2019-20 National Postsecondary Student Aid Study (NPSAS:20) who began their postsecondary education during the 2019-20 (full-scale sample) or 2018-19 (field test sample) academic year. For details on the NPSAS:20 sampling design see NPSAS:20 Supporting Statement Part B (OMB# 1850-0666 v.25). Specific plans are provided below for the BPS:20/22 cohort.

Section 1 – Respondent Universe

Included in this section is information describing the respondent universe and any sampling or other respondent selection method that will be used.

The respondent universe for BPS:20/22 consists of all students who began their postsecondary education for the first time during the 2019-20 academic year at any Title IV-eligible postsecondary institution in the United States.

The BPS:20/22 full-scale cohort will be comprised of students who first enroll in postsecondary education after high school during the 2019-20 academic year. The BPS:20/22 full-scale sample will include students from the NPSAS:20 full-scale sample who were identified as confirmed or potential 2019-20 academic year first-time beginner students based on survey, institution, or other administrative data.

Section 2 – Statistical Methodology

Statistical Methodology

The target population for the 2020/22 Beginning Postsecondary Students Longitudinal Study (BPS:20/22) full-scale consists of all students who began their postsecondary education for the first time during the 2019–20 academic year at any Title IV-eligible postsecondary institution in the United States. Identification of the BPS:20/22 full-scale sample requires a multi-stage process that began, first, with selection of the 2019-20 National Postsecondary Student Aid Study (NPSAS:20) full-scale sample of institutions and was followed next by selection of students within those institutions. The BPS:20/22 full-scale sample is comprised of students from the NPSAS:20 full-scale sample who were determined to be first-time beginners (FTBs), or were potential FTBs, as indicated by the NPSAS institution or administrative data.

NPSAS:20 Full-scale Sample

The NPSAS:20 institution (first stage) sampling frame included all levels (less-than-2-year, 2-year, and 4-year) and control classifications (public, private nonprofit, and private for-profit) of nearly all Title IV eligible postsecondary institutions in the 50 states, the District of Columbia, and Puerto Rico. The institution sampling frame used institution data collected from various surveys of the Integrated Postsecondary Education Data System (IPEDS). An institution was NPSAS-eligible if, during the 2019-20 academic year, the institution:

- offered an educational program designed for persons who have completed secondary education;
- offered at least one academic, occupational, or vocational program of study lasting at least 3 months or 300 clock hours;
- offered courses that were open to more than the employees or members of the company or group (e.g., union) that administered the institution;
- was located in the 50 states, the District of Columbia, or Puerto Rico;
- was not a U.S. service academy (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) due to their unique funding/tuition base); and
- had a signed Title IV participation agreement with the U.S. Department of Education (an institution that has a written 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).

The NPSAS:20 institution sampling frame was constructed from the Integrated Postsecondary Education Data System (IPEDS) 2018-19 Institutional Characteristics Header, 2018-19 Institutional Characteristics, 2017-18 12-Month Enrollment, and 2017 Fall Enrollment files. The institution strata used for the sampling design were based on 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 (includes all eligible institutions that IPEDS classifies as public 4-year institutions, including those that are non-doctorate-granting, primarily sub-baccalaureate institutions);and
- all other institutions, including:
 - public less-than-2 year;
 - private nonprofit (all levels); and
 - private for-profit (all levels).

The sample design allowed NPSAS:20 to have state-representative undergraduate student samples for public 2-year and public 4-year institutions as well as overall. From this point forward, the word “state” will refer to the 50 states, the District of Columbia, and Puerto Rico. In addition, the sample was nationally representative for both undergraduate and graduate students. The NPSAS:20 institution sample consisted of a census of all public 2-year and all public 4-year institutions and a random sample of institutions from the “all other institutions” stratum. Within the “all other institutions” stratum, the goal was to sample at least 30 institutions per state, when there are 30 institutions in this stratum in a state, so that institutions in this stratum were sufficiently represented within the state and national samples.

The following criteria were used to determine institution sample sizes within the “all other institutions” stratum:

- In states with 30 or fewer institutions in the “all other institutions” strata, a census of these institutions was selected.
- 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 selected a census of institutions. We arbitrarily chose 36 institutions as the cutoff to avoid high sampling fractions. This cutoff resulted in taking a census of institutions in states that had between 31 and 36 institutions in the “all other institutions” strata. Based on the latest IPEDS data, there were only three states (Mississippi, Nebraska, and Nevada) that had between 31 and 36 institutions in the “other” stratum.
- In states with more than 36 institutions in the “all other institutions” strata, a sample of 30 institutions was selected.

Within the “all other institutions” stratum, institutions were selected using sequential probability minimum replacement (PMR) sampling (Chromy 1979), which resembles stratified systematic sampling with probabilities proportional to a composite measure of size. This is the same methodology that has been used since NPSAS:96. Institution measure of size was determined using undergraduate and graduate student enrollment counts and FTB counts from the IPEDS 2017-18 12-Month Enrollment and 2017 Fall Enrollment files, respectively (OMB# 1850-0666 v.25.). Composite measure of size (Folsom, Potter, and Williams 1987) sampling was used to ensure that target sample sizes were achieved within institution and student sampling strata, while also achieving approximately equal student weights across institutions. All eligible students from sampled institutions comprised the student sampling frame.

Within the “all other institutions” stratum, additional implicit stratification was accomplished by sorting the sampling frame by the following classifications, as appropriate (OMB# 1850-0666 v.25.):

- Control and level of institution;
- Historically Black Colleges and Universities (HBCUs) indicator;
- Hispanic-serving institutions (HSIs) indicator (no longer available from IPEDS, so we created an HSI proxy following the definition of HSI as provided by the U.S. Department of Education (<https://www2.ed.gov/programs/dueshsi/definition.html>) and using IPEDS Hispanic enrollment data);
- Carnegie classifications of postsecondary institutions; and
- the institution measure of size.

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

From the approximately 3,110 institutions selected for the NPSAS:20 full-scale data collection, 98 percent met eligibility requirements; of those, approximately 72 percent provided enrollment lists.

The second stage of the NPSAS:20 sample specification was the selection of a stratified sample of individuals within sampled institutions. NPSAS-eligible undergraduate and graduate students were those who were enrolled in the NPSAS institution in any term or course of instruction between July 1, 2019 and April 30, 2020 and who were:

- 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 but had 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®) or other high school completion program. The GED® credential is a high school equivalency credential earned by passing the GED® test, which is administered by GED Testing Service (<http://www.gedtestingservice.com/ged-testing-service>).

There were 11 student sampling strata as follows:

- undergraduate students who were potential FTBs;
- other undergraduate students not classified as potential FTBs;
- graduate students who were veterans;
- master's degree students in science, technology, engineering, and mathematics (STEM) programs;
- master's degree students in education and business programs;
- master's degree students in all other programs;
- doctoral-research/scholarship/other graduate students in STEM programs;
- doctoral-research/scholarship/other graduate students in education and business programs;

- doctoral-research/scholarship/other graduate students in other programs;
- doctoral-professional practice students; and
- other graduate students not captured in the above categories.

When students met the criteria to be classified into multiple stratum, they were assigned following the list in hierarchical order (e.g., a STEM master's student who was also a veteran would fall into the “graduate students who are veterans” category). Several student subgroups were sampled at rates different than their natural occurrence within the population due to specific analytic objectives. The following groups were oversampled:

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

The NPSAS:20 full-scale sample was randomly selected from the frame with students sampled at fixed rates according to student sampling strata and institution sampling strata. Sample yield was monitored, and sampling rates were adjusted when necessary. Sampling rates were adjusted to maintain sample yield targets, such as when institution enrollment lists were trending larger or smaller than expected. The full-scale sample achieved a size of approximately 380,100 students; about 173,360 of which were asked to complete a survey, and around 206,740 of which were not asked to complete a survey. The administrative sample of 380,100 students was randomly selected first and the 173,360 students requested to complete a survey were subsampled from the larger administrative sample. The latter group of 206,740 are referred to as an admin-only sample since only administrative data were collected for them (they were not fielded for the survey data collection). Student records and administrative data were attempted to be collected for all sampled students.

Identification of FTBs

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. To address this concern, participating institutions were asked to provide additional information for all eligible students and matching to administrative databases was utilized to further reduce false positives and false negatives prior to sample selection.

In addition to an FTB indicator, we requested that enrollment lists provided by institutions (or institution systems) include degree program, class level, date of birth, enrollment in high school (or completion program) 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, were not classified as potential FTBs for sampling. Additionally, students who were 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 were not eligible for sampling. If the FTB indicator was not provided for a student on the list but the student was 18 years old or younger and did not appear to be dually enrolled, the student was classified as a potential FTB for sampling. Otherwise, if the FTB indicator was not provided for a student on the list and the student was over the age of 18, then the student was sampled as an “other undergraduate,” (but such students would be included in the BPS cohort if identified during the student survey as an FTB).

Prior to sampling, all students listed as potential FTBs were matched to National Student Loan Data System (NSLDS) records to determine if any had 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 could 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 could not exclude all listed FTBs with prior enrollment but significantly improved the accuracy of the listing prior to sampling, yielding fewer false positives.

Simultaneously with NSLDS matching, all potential FTBs were also matched 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. After NSLDS and CPS matching, a subset of the remaining potential FTBs were matched to the National Student Clearinghouse (NSC) for further narrowing of FTBs based on the presence of evidence of earlier enrollment. Due to the cost of matching individuals to the NSC we only targeted individuals in institution sectors that historically had high false positive rates. Potential FTBs over the age of 18 in the public 2-year and for-profit sectors were targeted for this match because these sectors either had high false-positive rates in NPSAS:12 or had large NPSAS:20 sample sizes.

In setting the NPSAS:20 FTB selection rates, we considered the false-positive rates, based on the NPSAS:12 survey, which had an overall unweighted false positive weight of 22 percent and an unweighted false negative rate of 4.6 percent. NPSAS:12 was examined as the reference as it was the most recent NPSAS administration with a BPS cohort. Based on confirmed FTB status from the NPSAS:20 survey, we found the observed false positive rate of the final potential FTB indicator to be 21.7 percent with a false negative rate of 2.2 percent. These rates are similar to those observed from NPSAS:12 which followed the same administrative matching approach to refine the institution-provided FTB flags.

After NPSAS:20 data collection, the potential FTB flag was further refined by matching all potential FTBs to the NSLDS again to catch any individuals who did not match through the early cycle (generally due to missing student identifiers at the time of the first match). In addition, this second NSLDS match was used to remove any students who responded to the NPSAS:20 student survey and self-identified as an FTB, but the NSLDS match indicated otherwise.

2020/22 Beginning Postsecondary Students Longitudinal Study (BPS:20/22) Full-scale
Supporting Statement Part B

BPS:20/22 Full-Scale Sample

NPSAS:20 consisted of an administrative sample of 380,100 students. The NPSAS:20 survey sample was a subset of approximately 173,360 students from the total student sample. The remaining individuals who were part of the administrative sample but were not selected for the survey sample are referred to as administrative-only (hereafter referred to as 'admin-only') sample or students. For the BPS:20/22 cohort, the full-scale sampling frame is based on the larger NPSAS:20 administrative sample. This frame contains survey confirmed FTBs as well as survey nonrespondent and admin-only potential FTBs.

The BPS:20/22 frame only includes confirmed and potential FTB students who are defined as study respondents in NPSAS:20 and who were not found to be deceased during NPSAS:20 data collection. A study respondent is any individual who is a survey or administrative student respondent. All confirmed FTB students are also NPSAS:20 study respondents since NPSAS survey respondents are also NPSAS:20 study respondents. The BPS:20/22 full-scale sample consists of three different groups based on their NPSAS:20 sample along with their administrative and survey response statuses. The groups are (1) NPSAS:20 survey respondents who are confirmed FTBs, (2) NPSAS:20 survey nonrespondents who are potential FTBs and NPSAS administrative student respondents, and (3) NPSAS:20 admin-only students who are potential FTBs and NPSAS administrative student respondents. As stated earlier, all three groups are considered study respondents. Table 1 shows the distribution of the BPS:20/22 frame.

Table 1. BPS:20/22 full-scale frame by NPSAS:20 sample and response status

NPSAS:20 full-scale status	Count
Total study respondent FTBs	51,910
Confirmed FTBs	
Survey respondents	26,470
Potential FTBs who are administrative respondents	
Survey nonrespondents	14,710
Admin-only students ¹	10,730

1- 370 admin-only students in institutions without NPSAS survey students are not included

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

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

Some NPSAS:20 admin-only potential FTB students (approximately 370) who are considered administrative student respondents were sampled from institutions that were unable to provide accurate enrollment list information necessary for student survey contacting. These students were therefore not able to be included in the student survey portion of the sample and will not be included in the BPS:20/22 sampling frame. These institutions were not found to significantly differ from the remaining institutions or admin-only potential FTBs and thus should

not impact sample representativeness. The decision to include admin-only FTB students is based on two primary factors: 1) field test data showing significantly higher response rates for the admin-only group compared to survey nonrespondents and 2) analysis results suggesting minimal loss of precision which will still allow the sample to meet precision goals.

The BPS:20/22 full-scale sample consists of approximately 37,330 total students including approximately 26,470 NPSAS:20 survey respondents and approximately 10,860 potential FTB students who did not complete the survey but who were administrative student respondents. The potential FTB students are split nearly evenly between NPSAS:20 survey nonrespondents and admin-only students.

In addition, the BPS:20/22 sample is designed with a goal to be state representative for a subset of states. The minimum desired sample sizes for each state subgroup are based on the ability to measure a relative change of 20 percent in proportions across rounds and an assumed design effect of 3. To meet this goal, 741 survey respondents per desired state are necessary by the end of BPS:20/25 data collection.

The base NPSAS:20 administrative sample that BPS:20/22 follows was designed to be state representative. The distribution of confirmed and potential FTBs from NPSAS:20 was reviewed by state to arrive at the final decision to target representativeness for California, Florida, Georgia, New York, North Carolina, Pennsylvania, and Texas. An oversample of approximately 330 students is required to achieve the desired sample size for Georgia. The decision to oversample students for Georgia was made as Georgia was close to the target sample size, had enough additional administrative student respondents to be oversample and was considered an important state to target.

All NPSAS:20 survey respondents who are confirmed FTBs are sampled with certainty. The approximately 25,440 potential FTBs who did not complete a survey, but who are administrative respondents, are explicitly stratified by “survey nonrespondent” and “admin-only student” as well as control and level of institution. A sample of approximately 10,530 students is split evenly between survey nonrespondents and admin-only students and was proportionally allocated across control and level of institution. Additionally, the oversample of approximately 330 students for Georgia is allocated by sampling all potential FTBs within Georgia. Within the explicit strata, a simple random sample of students is selected. Table 2 displays the final BPS:20/22 sample by control and level of institution including the oversample of approximately 330 students within Georgia while table 3 details the sample by state.

2020/22 Beginning Postsecondary Students Longitudinal Study (BPS:20/22) Full-scale
Supporting Statement Part B

Table 2. BPS:20/22 sample by control and level of institution

Institution characteristics	Total	Confirmed and potential FTB administrative student respondents from NPSAS:20 Administrative Sample		
		NPSAS:20 Survey Respondents	NPSAS:20 Survey Nonrespondents	NPSAS:20 Admin-Only Respondents
Total	37,330	26,470	5,510	5,350
Institution type				
Public				
Less-than-2-year	400	270	120	<5
2-year	12,370	9,000	2,210	1,160
4-year non-doctorate-granting primarily sub-baccalaureate	2,610	2,010	550	50
4-year non-doctorate-granting primarily baccalaureate	2,290	1,850	260	180
4-year doctorate-granting	7,840	4,750	790	2,300
Private nonprofit				
Less-than-4-year	290	190	100	<5
4-year non-doctorate-granting	2,940	2,000	220	720
4-year doctorate-granting	3,600	2,400	330	880
Private for-profit				
Less-than-2-year	980	780	150	50
2-year	1,600	1,280	300	10
4-year	2,430	1,940	480	<5

NOTE: Detail may not sum to totals because of rounding. Potential FTB's are individuals who did not complete a NPSAS survey who appeared to be FTB's in enrollment or NPSAS:20 admin data.

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2020/22 Beginning Postsecondary Students Longitudinal Study (BPS:20/22) Full-scale.

2020/22 Beginning Postsecondary Students Longitudinal Study (BPS:20/22) Full-scale
Supporting Statement Part B

Table 3. BPS:20/22 sample by state

State	Confirmed and potential FTB administrative student respondents from NPSAS:20 Administrative Sample			
	Total	NPSAS Survey Respondents	NPSAS Survey Nonrespondents	NPSAS Admin- Only Respondents
Total	37,330	26,470	5,510	5,350
California	2,750	2,230	500	20
Florida	1,970	1,580	350	40
Georgia ¹	1,410	860	390	170
North Carolina	1,370	970	230	180
New York	2,030	1,660	280	90
Pennsylvania	1,340	940	170	230
Texas	2,340	1,810	410	120
All other states	24,120	16,410	3,190	4,510

¹ Georgia includes an oversample of an additional 250 NPSAS survey nonrespondents and 80 NPSAS admin-only respondents.

NOTE: Detail may not sum to totals because of rounding. Potential FTB's are individuals who did not complete a NPSAS survey who appeared to be FTB's in enrollment or NPSAS:20 admin data.

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2020/22 Beginning Postsecondary Students Longitudinal Study (BPS:20/22) Full-scale.

Based on observations from the BPS:20/22 field test and previous full-scale BPS data collections, we expect a response rate of roughly 82 percent for NPSAS:20 survey respondents, 22 percent for survey nonrespondents and approximately 57 percent for the admin-only students. Table 4 displays the BPS:20/22 full-scale sample size by NPSAS:20 data collection outcome group and the expected yield of completed interviews.

Table 4. BPS:20/22 expected completes by NPSAS:20 data collection outcome

NPSAS:20 Outcome	Sample Size	Eligibility Rate	Expected Response Rate	Expected Completes
Overall	37,330	0.94	0.72	25,030
NPSAS Survey Respondents	26,470	1.00	0.82	21,710
NPSAS Survey Nonrespondents	5,510	0.78	0.22	950
NPSAS Admin-Only Respondents	5,350	0.78	0.57	2,380

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

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2020/22 Beginning Postsecondary Students Longitudinal Study (BPS:20/22) Full-scale.

Section 3 – Methods for Maximizing Response Rate

Included in this section is information describing the methods to be used to maximize response and to deal with issues of non-response.

Achieving high response rates in the BPS:20/22 full-scale study data collection will depend on successfully identifying and locating sample members and being able to contact them and gain their cooperation. As was used successfully in prior NCES longitudinal studies, shortly before data collection begins, we will send an address update/initial contact mailing/e-mail to remind sample members of their inclusion in the study. The following sections outline additional methods for maximizing response to the BPS:20/22 full-scale data collection.

a. 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.

- Tracing activities conducted prior to the start of data collection will include batch database searches, such as to the National Change of Address (NCOA), for cases with enough contact information to be matched. To handle cases for which contact information is invalid or unavailable, project staff will conduct additional advance tracing through proprietary interactive databases to expand on leads found.
- Hard copy mailings, e-mails, and text messages will be used to maintain ongoing contact with sample members, prior to and throughout data collection. We will send a panel maintenance mailing in October 2021 to request that sample members update their contacting information (previously approved under the BPS:20/22 field test clearance package, OMB # 1850-0631 v.18).
- Immediately prior to the start of data collection, we will send an initial contact mailing to sample members to request that they update their contact information. A follow-up reminder e-mail will be sent approximately 2 weeks after that mailing to remind them to respond. In addition, we will send a letter to announce the start of data collection. The announcement will include a request that sample members complete the web survey and will provide each sample member a Study ID and password, the study website address, and a toll-free number to the help desk. Sample members who did not participate in the NPSAS:20 survey will receive \$2 cash (or PayPal if a good address is not available) with the data collection announcement. After the data collection announcement mailing, an e-mail message with the same information will also be sent.
- The telephone locating and interviewing stage will include 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 interviewing 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 conducting a LexisNexis e-mail search for nonrespondents toward the end of data collection.

b. Training for Data Collection Staff

Telephone data collection will be conducted using the contractor's virtual call center, which allows the contractor to retain experienced staff who can successfully work from home. Telephone data collection staff will include Performance Team Leaders (PTLs) and Data Collection Interviewers (DCIs). Training programs, administered through Zoom, are critical to maximizing response rates and collecting accurate and reliable data.

PTLs, who are responsible for all supervisory tasks, will attend project-specific training for PTLs, 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. PTLs will also receive training in the following areas: providing direct supervision of virtual staff 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 DCIs 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. DCIs will receive project-specific training on telephone interviewing and answering questions from web participants regarding the study or related to specific items within the interview. Bilingual interviewers will receive a supplemental training that will focus on Spanish contacting and interviewing procedures. At the conclusion of training, all BPS 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.

c. Case Management System

The BPS:20/22 full-scale survey will be conducted using a single web-based survey instrument for both web (including mobile devices) and CATI data collection. Data collection activities will be monitored through a CATI case management system, which is equipped with the numerous capabilities, including: online 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 based upon information obtained from respondents); sample management module for tracking case progress and status; and an 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 reduces the need for tracking appointments and helps ensure punctual interviewing. The scheduler automatically calculates the delivery

time of the case in reference to the appropriate time zone.

- Sorting non-appointment cases according to parameters and priorities set by project staff is another feature of the scheduling module. For instance, priorities may be set to give first preference to cases within certain sub-samples or geographic areas; or cases may be sorted to establish priorities based on prior round response 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.
- 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 cases with language barriers that require interviewers with specialized skills.
- 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.
- The scheduler can flag problem cases for supervisor attention. 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 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.

d. 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. The survey instrument can be found in Appendix E.

In addition to the functional capabilities of the case management system and web survey instruments described above, our efforts to achieve the desired response rate will include using established procedures proven effective in other large-scale studies. These include:

- Providing multiple response modes, including mobile-friendly self-administered and interviewer-administered options.
- Offering incentives to encourage response.
- Employing experienced DCIs who have proven their ability to contact and obtain cooperation from a high proportion of sample members.
- Training the DCIs 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 a completed interview at the initial contact while allowing

respondent flexibility in scheduling appointments to be interviewed.

- Providing assurance of confidentiality procedures, including requiring respondents to answer security questions before obtaining and resuming access to the survey and the survey automatically logging out of a session after 10 minutes of inactivity.
- Thoroughly reviewing all refusal cases and making special conversion efforts whenever feasible (see next section e).

e. Refusal Aversion and Conversion

Recognizing and avoiding refusals is important to maximize the response rate, and interviewer training will cover this and other topics related to obtaining cooperation. PTLs will closely monitor DCIs at the beginning of outbound calling, and provide re-training, as necessary. In addition, supervisors will review daily interviewer production reports produced by the CATI system to identify and retrain any DCIs who are producing unacceptable numbers of refusals or other problems.

Refusal conversion efforts 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.

Section 4 – Tests of Procedures and Methods

Included in this section is information describing any tests of procedures or methods that will be undertaken.

During the course of this data collection, the following experiment(s) will be undertaken. The BPS:20/22 field test included two sets of experiments: data collection experiments focused on survey participation to reduce nonresponse error and the potential for nonresponse bias, and questionnaire design experiments focused on minimizing measurement error to improve data quality. The full-scale data collection design described below will implement the tested approaches, revised based on the BPS:20/22 field test results described in Section 4.a.

a. Summary of BPS:20/22 Field Test Data Collection Design and Results

The BPS:20/22 field test contained two data collection experiments and two questionnaire design experiments. The results of these field test experiments are summarized below. For detailed results of the BPS:20/22 field test experiments, see Appendix D.

The data collection experiments explored the effectiveness of 1) offering an extra incentive for early survey completion, and 2) sending survey reminders via text messages. Results from these data collection experiments provide insight in preparation for the full-scale study regarding the effectiveness of these interventions across three data quality indicators: survey response (operationalized using response rates), sample representativeness (assessed across age, sex, ethnicity, race, and institutional control), and data collection efficiency (operationalized as the number of the days between the start of the experiment and survey completion).

The “early bird” incentive experiment investigated the effectiveness of giving respondents an additional \$5 incentive if they completed the survey within the first three weeks of data collection (experimental group) versus no additional incentive (control group). Response rates at the end of data collection did not differ across the early bird group (63.9 percent) and the control group (63.4 percent; $X^2 = 0.08$, $p = .78$). Both the early bird and control groups had similar representativeness across age, sex, ethnicity, race, and institutional control. At the end of data collection, respondents in the early bird group took significantly fewer days (28.1 days) than respondents in the control group (30.9 days) to complete the survey ($t(2,231.7) = 2.09$, $p < 0.05$). However, this difference is small (2.8 days), and not long enough to allow for any significant cost savings in the data collection process (e.g., via fewer reminder calls, texts, or mailings). Therefore, the use of an early bird incentive in the BPS:20/22 full-scale data collection is not recommended.

The reminder mode experiment compared the effectiveness of using text message reminders (experimental group) versus telephone call reminders (control group). Response rates at the end of data collection for the text message group (29.6 percent) and the telephone group (31.5 percent) did not significantly differ ($X^2 = 0.83$, $p = 0.36$). In the telephone reminder group, the percentage of white respondents (74.1 percent) significantly differed from the percentage of white nonrespondents (64.5 percent; $X^2 = 7.40$, $p < 0.01$), indicating a potential source of nonresponse bias. For the text message reminder group, there was not a significant difference between the percentage of White respondents (65.3 percent) and nonrespondents (63.3 percent) ($X^2 = 0.28$, $p = 0.60$), indicating better sample representativeness. The text message and telephone groups had similar representativeness across the remaining respondent characteristics: age, sex, ethnicity, and institutional control.

Finally, the number of days it took for respondents in the text message reminder group to complete the survey (75.5 days) was not significantly different from the telephone reminder group (77.0 days; $t(542.8) = 0.62$, $p = 0.27$). As text message reminders achieved response rates, representativeness, and efficiency that was comparable to more expensive telephone reminders, the use of text reminders (coupled with telephone reminders as described in Section 4.b) is recommended as a part of the BPS:20/22 full-scale data collection.

The questionnaire design experiments explored the effectiveness of 1) different methods for collecting enrollment data, and 2) using a predictive search database on a question collecting address information. In addition, information about the impacts of the coronavirus pandemic was collected by randomly assigning respondents one of two separate topical modules to maximize the number of questions fielded without increasing burden. Results from these questionnaire collection experiments provide insight in preparation for the full-scale study regarding the effectiveness of these methods across three data quality indicators: missingness (operationalized as item- and question-level nonresponse rate), administrative data concordance (operationalized as agreement rates between self-reported enrollment and administrative records; month-level enrollment intensity experiment only), and timing burden (operationalized as the mean complete time at the question level).

The month-level enrollment intensity experiment compared two methods for collecting enrollment information in the 2020-21 academic year: a single forced-choice grid question that displayed all enrollment intensities (i.e., full-time, part-time, mixed, and no enrollment) on one form (control group) and separate yes/no radio gates for full-time and part-time enrollment (experimental group). There were no statistically significant differences across the control and treatment conditions on rates of missingness (0 percent and 0.03 percent missing, respectively ($t(636) = 1.42$, $p = 0.1575$)) or agreement rates with administrative enrollment data (70.0 percent agreement and 70.6 percent agreement, respectively ($t(1311.1) = 0.24$, $p = 0.8119$)). On

average, the treatment group took significantly longer to complete the enrollment question (17.2 seconds) than the control group (10.5 seconds; $t(1312.8) = 15.47, p < .0001$), though this difference is expected given the additional screen respondents must navigate in the experimental group. As the experimental question did not represent a clear improvement over the original forced-choice grid, the use of the original question is recommended BPS:20/22 full-scale data collection.

The predictive search address database experiment explored the utility of suggesting USPS-standardized addresses to respondents as they entered their address into the survey. This analysis compares address entry for the same set of respondents across the BPS:20/22 field test (using the database-assisted predictive search method) and the NPSAS:20 full-scale survey (using traditional, manual address entry). Overall, 98 percent of respondents provided a complete permanent address using the manual method in NPSAS:20, compared to 85 percent using the predictive search method in BPS:20/22 field test ($t(2023.5) = 13.88, p < .0001$). However, it should be noted that addresses obtained using the predictive search method were error-free (0 percent of FTB check addresses were undeliverable), while 1.2 percent of addresses obtained using the manual method were undeliverable. Also, additional improvements to the survey instrument (e.g., soft check validations for incomplete addresses) may further reduce rates of missingness for the predictive search method. Finally, on average, respondents took longer to provide their address using manual entry (29.5 seconds) compared to the predictive search system (27.1 seconds; $t(3055.7) = 4.09, p < .0001$). Given the higher quality data resulting from the predictive search method, the potential to improve the predictive search method via instrument adjustments, and the significant reduction in completion time compared to manual entry, the continuation of predictive search method is proposed for BPS:20/22 full-scale.

Given the impact of the coronavirus pandemic on higher education, researchers have expressed interest in using BPS:20/22 data to examine these impacts on postsecondary students. BPS:20/22 field test respondents were randomly assigned into two groups that received one of two modules. Each module measured similar constructs, however, module one consisted of survey questions from NPSAS:20 that measured student academic, social, and personal experiences related to the coronavirus pandemic, and module two collected a new set of constructs, including changes in enrollment and borrowing, changes in academic engagement, and access to support resources, that may be of analytic value to researchers and policymakers. Across both modules, the average item nonresponse rate was 2 percent. Module one had an average nonresponse rate of 3 percent, significantly higher than the 0.6 percent nonresponse rate of module two ($t(836.72) = 5.16, p < .0001$). Regardless of module assignment, the coronavirus pandemic questions took respondents an average of 2.7 minutes to complete. The BPS:20/22 full-scale survey instrument will administer a subset of the questions from both field test coronavirus pandemic modules, based upon field test performance and TRP feedback. The coronavirus pandemic module for the full-scale maintains the burden goal of three minutes.

b. BPS:20/22 Full-scale Data Collection Design

The data collection design proposed for the BPS:20/22 full-scale study builds on the designs implemented in past BPS studies, as well as the National Postsecondary Student Aid Study (NPSAS) and the Baccalaureate and Beyond (B&B) studies. Additionally, results from the BPS:20/22 field test Data Collection Experiments (Appendix D) inform recommendations for the BPS:20/22 full-scale data collection design.

A primary goal of the full-scale design is to minimize the potential for nonresponse bias that could be introduced into BPS:20/22, especially bias that could be due to lower response rates among NPSAS:20 nonrespondents. Another important goal is to reduce the amount of time and cost of data collection efforts.

To accomplish these goals, the plan is to achieve at least a 70 percent response rate. Doing so will minimize potential nonresponse bias, optimize statistical power, and enable sub-group analyses. The sample will be divided into two groups and differential data collection treatments will be implemented based on prior round response status. A similar approach was successfully implemented in the BPS:20/22 field test, and the latest B&B studies where more reluctant sample members received a more aggressive protocol (for an experimental comparison see B&B:16/17 field test <https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2020441>).

For the BPS:20/22 full-scale design, the following sample groupings will be used:

- NPSAS:20 survey respondents: Sample members who responded to NPSAS:20 and self-identified that they began their postsecondary education between July 1, 2019 and April 30, 2020 will receive a default data collection protocol (n = 26,470).
- NPSAS:20 survey nonrespondents and administrative-only cases: NPSAS:20 administrative student respondents who are potential 2019-20 academic year FTBs will receive an aggressive data collection protocol. This group includes NPSAS:20 survey nonrespondents (n = 5,510) and NPSAS:20 administrative-only sample (cases who were never invited to complete the NPSAS:20 survey; n = 5,350) who are potential 2019-20 academic year FTBs based on administrative data. The goal of this treatment is to convert reluctant sample members (i.e., NPSAS:20 survey nonrespondents) and sample members who have never been contacted (i.e., administrative-only cases) to participate in the study as early in data collection as possible. Table 5 below presents the type and timing of interventions to be applied in data collection by groups and protocol. The details of these interventions are described below.

Table 5. 2020/22 Beginning Postsecondary Students Full-Scale data collection protocols, by data collection phase and group assignment

	Data Collection Group Assignments	
	Default Protocol	Aggressive Protocol
Sample	<ul style="list-style-type: none"> • NPSAS:20 interview respondents (26,470) 	<ul style="list-style-type: none"> • NPSAS:20 interview non-respondents (5,510) • NPSAS:20 admin-only (5,350)
Data Collection Protocols		
Prior to data collection	<ul style="list-style-type: none"> • Panel maintenance • Prenotification/greeting (from NCES) 	<ul style="list-style-type: none"> • Panel maintenance • Prenotification/greeting (from NCES)
Early completion phase	<ul style="list-style-type: none"> • DC announcement mail, text, and e-mail • Mode tailoring with NPSAS:20 completion mode 	<ul style="list-style-type: none"> • DC announcement mail, text, and e-mail \$2 cash prepaid • Targeted CATI begins 2 weeks after mail outs – continued throughout
Production phase 1	<ul style="list-style-type: none"> • Targeted light CATI outbound begins • Postcard, e-mail, text message reminders – continued throughout 	<ul style="list-style-type: none"> • Postcard, e-mail, text message reminders – continued throughout
Production phase 2	<ul style="list-style-type: none"> • \$10 boost post-paid 	<ul style="list-style-type: none"> • \$20 boost post-paid • Abbreviated survey offered
Nonresponse Conversion Phase	<ul style="list-style-type: none"> • Abbreviated survey offered 	
Total incentives	<ul style="list-style-type: none"> • \$30 baseline • \$10 boost • Maximum = \$40 	<ul style="list-style-type: none"> • \$2 pre-paid • \$45 baseline • \$20 boost • Maximum = \$67

The duration of each phase of data collection will be determined based on phase capacity—the time at which a subgroup's estimates remain stable regardless of additional data collection efforts. For example, during the early completion phase, key metrics are continually monitored, and when they stabilize over a period of time, cases are then transferred to the next phase. Phase capacity will be determined based on a series of individual indicators within each data collection protocol. For example, response rates and other level of effort indicators over time accounting for covariates, such as institution control, will be assessed.

Incentives. The baseline incentive for the default protocol will be \$30 with a \$10 incentive boost in Production Phase 2, leading to a maximum possible total incentive of \$40. The baseline incentive for the aggressive protocol will be \$45. An experiment conducted in BPS:12/14 showed that a \$45 baseline incentive yielded the highest response rates (Hill et al. 2016). However, this experiment was underpowered to detect differences from \$30 in the lower propensity response groups (Wilson et al. 2015). Nonetheless, implementing a higher baseline incentive is recommended given the \$30 baseline incentive and the \$10 incentive boost from NPSAS:20 was not enough to encourage prior year nonrespondents to participate. Further, the \$40 BPS:20/22 field test incentive yielded a response rate of only 25.3 percent among these “aggressive protocol” sample members. The baseline incentive will be paid in addition to a possible \$2 prepaid incentive (see prepaid incentive section below), and a \$20 incentive boost (see nonresponse conversion incentive section below). The maximum possible total incentive is \$67 in the aggressive data collection protocol. Results from the BPS:20/22 field test showed that offering sample members an early bird incentive did not significantly improve response rates or representativeness by the end of data collection, nor did it practically improve data collection efficiency (Appendix D). Therefore, early bird incentives will not be used in the full-scale study for either the default or aggressive protocols.

Beyond the baseline incentives, both data collection protocols employ similar interventions, although the timing and intensity of these interventions differ across groups. Interventions occur sooner in the aggressive protocol and are more intense.

Prenotification. The first mailing that individuals in the default and aggressive data collection protocols will receive is a greeting card. This mailing is aimed to increase the perceived legitimacy of the upcoming survey request (e.g., Groves et al. 1992) in both data collection groups and announce the incentive amounts. Greeting cards, in particular, have been shown to significantly increase response rates in longitudinal studies (Griggs et al. 2019) and this method will be used as a precursor to the invitation letter. The greeting card will be mailed a few weeks in advance of data collection.

\$2 prepaid incentive. Cash prepaid incentives have been shown to significantly increase response rates in both interviewer-administered and self-administered surveys. These prepaid incentives increase the perceived legitimacy of the survey request and therefore reduce the potential for nonresponse bias (e.g., Church 1993; Cantor et al. 2008; Goeritz 2006; Medway and Tourangeau 2015; Messer and Dillman 2011; Parsons and Manierre 2014; Singer 2002). During the early completion phase in the B&B:16/17 field test, prepaid incentives (\$10 via check or PayPal) in combination with telephone prompting also significantly increased response rates by 4.4 percentage points in the aggressive protocol group. Given these positive findings combined with general recommendations in the literature (e.g., Singer and Ye 2013; DeBell et al. 2019), a small \$2 cash prepaid ‘visible’ incentive, or, where necessary due to low address quality, a \$2 prepaid PayPal incentive announced on a separate index card will be sent to all cases in the aggressive protocol for BPS:20/22 full-scale (see results from the B&B:16/20 calibration experiment – Kirchner et al. 2021). Sample members will be notified of this prepaid incentive in the data collection prenotification, and it will be included

in the data collection announcement letter.

Mode tailoring. The leverage-saliency theory suggests that respondents have different hooks that drive their likelihood of survey participation (Groves et al. 2000); thus, offering a person the survey mode (e.g., web, mail, telephone) that they prefer may increase their likelihood of responding. This is further supported by empirical evidence that shows offering people their preferred mode speeds up their response and is associated with higher participation rates (e.g., Olson et al. 2012). Using the NPSAS:20 survey completion mode as a proxy for mode preference, the BPS:20/22 full-scale early completion phase will approach sample members in the default protocol with their mode of completion for NPSAS:20. Specifically, while all sample members in the default protocol will receive identical data collection announcement letters and e-mails, those who completed the NPSAS:20 survey by telephone (4.3 percent) will be approached by telephone from the start of data collection. Likewise, those who completed the NPSAS:20 main study survey online will not be contacted by telephone before a preassigned outbound telephone data collection date.

(Light) outbound CATI calling and text messaging. The results from the BPS:20/22 field test showed that there were no statistically significant differences in the response rates for the text message reminder and the telephone only group at the end of the experimental period (Appendix D). As a result, both data collection groups will receive early text message reminders combined with prioritized telephone calls. Telephone calls will be prioritized to individuals for whom no cell phone number exists, those who opt out of the text message reminders, and those sample members who will be prioritized based on other criteria (e.g., from lower performing sectors). Text messages from sample members will be answered with an automated text response, with the possibility of two-way text messaging (i.e., interviewers respond to text message questions sent by sample members) in some cases.

Sample members in the default group who qualify for telephone calls will receive a light CATI protocol. Light CATI involves a minimal number of phone calls, used mainly to prompt web response (as opposed to regular CATI efforts that involve more frequent phone efforts, with the goal to locate sample members and encourage their participation). In the B&B:16/17 field test, introduction of light CATI interviewing appeared to increase production phase response rates in the default protocol. Although one should use caution when interpreting these results – group assignment in B&B:16/17 field test was not random but instead compared NPSAS:16 “early” and “late” respondents– the findings are consistent with the literature which has shown that web surveys tend to have lower response rates compared to interviewer-administered surveys (e.g., Lozar Manfreda et al. 2008). Attempting to survey sample members by telephone also increases the likelihood of initiating locating efforts sooner. B&B:16/17 field test results showed higher locate rates in the default protocol (93.7 percent), which had light CATI, compared to a more relaxed protocol without light CATI (77.8 percent; $p < 0.001$). For the BPS:20/22 full-scale data collection, light CATI will be used in the default protocol once CATI begins in Production Phase 1. Additionally, all cases in the aggressive protocol will receive earlier and more intense telephone prompting than eligible cases in the default group.

Incentive boosts. 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). These boosts are especially common in large federal surveys during their nonresponse follow-up phase (e.g., The Center for Disease Control and Prevention's National Survey of Family Growth) and have been implemented successfully in other postsecondary education surveys (e.g., HSL: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. Therefore, a \$10 incentive boost

increase to the BPS:20/22 baseline incentive is planned during Production Phase 2 for all remaining nonrespondents in the default data collection protocol, before the abbreviated survey is offered in the nonresponse conversion phase. Remaining nonrespondents in the aggressive data collection protocol will be offered a \$20 incentive boost increase to the baseline incentive before the abbreviated survey (both offered in Production Phase 2). This is because the \$10 incentive boost in NPSAS:20 did not show any effect on this group. If necessary, incentive boosts may be targeted only at certain groups of nonrespondents to achieve response goals (e.g., targeting nonrespondents from certain states to ensure representativeness, targeting aggressive group nonrespondents to reduce the potential for nonresponse bias).

Abbreviated survey. Obtaining responses from all sample members is an important assumption of the inferential paradigm. The leverage-saliency theory (Groves et al. 2000) and the social exchange theory (Dillman et al. 2014) suggest that the participation decision of an individual is driven by different survey design factors or perceived cost of participating. As such, reducing the perceived burden of participating by reducing the survey length may motivate sample members to participate.

During the B&B:16/17 field test, prior round nonrespondents were randomly assigned to one of two groups: 1) prior round nonrespondents who were offered the abbreviated survey during the production phase (i.e., before the nonresponse conversion phase), and 2) prior round nonrespondents who were offered the abbreviated survey during the nonresponse conversion phase (i.e., after the production phase). At the end of the production phase, prior round nonrespondents who received the abbreviated survey had a higher overall response rate (22.7 percent) than those who were not offered the abbreviated during that phase (12.1 percent; $t(2,097) = 3.67, p < 0.001$). Further, at the end of data collection, prior round nonrespondents who were offered the abbreviated survey during the earlier production phase had a significantly higher response rate (37 percent) than prior round nonrespondents who were not offered the abbreviated survey until the nonresponse conversion phase (25 percent) ($t(2,097) = 3.52, p = .001$). These results indicate that offering an abbreviated survey to prior round nonrespondents during the production phase (i.e., earlier in data collection) significantly increases response rates. The B&B:08/12 and B&B:08/18 full-scale studies also demonstrated the benefit of an abbreviated survey. Offering the abbreviated survey to prior round nonrespondents increased overall response rates of that group by 18.2 (B&B:08/12) and 8.8 (B&B:08/18) percentage points (Cominole et al. 2015). In NPSAS:20, 14.4 percent of those offered the abbreviated survey completed it. Therefore, an abbreviated survey option will be offered to all sample members in the BPS:20/22 full-scale study. For the aggressive protocol, the abbreviated survey will be offered during Production Phase 2, which is the latter half of the production phase of data collection. For the default protocol, the abbreviated survey will be offered as the last step in nonresponse conversion.

Other interventions. While all BPS studies are conducted by NCES, the data collection contractor, RTI International, has typically used the study-specific e-mail “@rti.org” to contact and support sample members. Changing the e-mail sender to the NCES project officer or the RTI project director may increase the perceived importance of the survey and help personalize the contact materials, thereby potentially increasing relevance. Switching the sender during data collection also increases the chance that the survey invitation is delivered to the sample member rather than to a spam filter.

As a result of the above experiment(s), detailed information on field results can be found in Appendix D.

Section 5 – Reviewing Statisticians and Individuals Responsible for Designing and Conducting the Study

Included in this section is the name and telephone number of individuals consulted on statistical aspects of the design and the name of the agency unit, contractor(s), grantee(s), or other persons who will actually collect and/or analyze the information for the agency.

BPS:20/22 is being conducted by NCES/ED. The following statisticians at NCES are responsible for the statistical aspects of the study: Dr. David Richards, Dr. Tracy Hunt-White, Dr. Elise Christopher, and Dr. Gail Mulligan. NCES's prime contractor for BPS:20/22 is RTI International (RTI). The following staff members at RTI are working on the statistical aspects of the study design: Dr. Joshua Pretlow, Dr. Jennifer Wine, Dr. Nestor Ramirez, Mr. Darryl Cooney, Mr. Michael Bryan, Dr. T. Austin Lacy, Dr. Emilia Peytcheva, and Mr. Peter Siegel. Principal professional RTI staff not listed above, who are assigned to the study include: Ms. Ashley Wilson, Ms. Kristin Dudley, Mr. Jeff Franklin, Ms. Chris Rasmussen, and Ms. Donna Anderson.

References

Cantor, D., O'Hare, B.C., and O'Connor, K.S. (2008). The Use of Monetary Incentives to Reduce Nonresponse in Random Digit Dial Telephone Surveys. In Lepkowski, J.M., Tucker, N.C., Brick, J.M., de Leeuw, E., Japac, L., Lavrakas, P.J., Link, M.W., and Sangster, R.L. (eds.). *Advances in Telephone Survey Methodology*. New York: Wiley.

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.

Church, A.H. (1993). Estimating the Effect of Incentives on Mail Survey Response Rates: A Meta-Analysis. *Public Opinion Quarterly*, 57(1): 62-79.

Cominole, M., Shepherd, B. and Siegel, P. (2015). 2008/12 Baccalaureate and Beyond Longitudinal Study (B&B:08/12). Data file documentation. NCES 2015-141. Washington, DC: National Center for Education Statistics.

DeBell, M., Maisel, N., Edwards, B., Amsbary, M., and Meldener, V. (2019). Improving Survey Response Rates with Visible Money. *Journal of Survey Statistics and Methodology*. Online First. Retrieved from: <https://academic.oup.com/jssam/advance-article/doi/10.1093/jssam/smz038/5610622>

Dillman, D., Smyth, J., and Christian, L. (2014). *Internet, Phone, Mail, and Mixed Mode Surveys: The Tailored Design Method*. John Wiley & Sons Inc. Hoboken, New Jersey.

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.

Goeritz, A.S. (2006). Incentives in web studies: Methodological issues and review. *International Journal of Internet Science*, 1: 58-70.

Griggs, A., Powell, R., Keeney, J., Waggy, M., Harris, K., Halpern, C. and Dean, S. (2019) Research Note: A Prenotice Greeting Card's Impact on Response Rates and Response Time, *Longitudinal and Life Course Studies*, 10(4): 421-432.

Groves, R.M., Cialdini, R., and Couper, M. 1992. Understanding the Decision to Participate in a Survey. *Public Opinion Quarterly*, 56(4):, 475-495.

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.

Groves, R.M., Singer, E. and Corning, A. (2000). Leverage-Saliency Theory of Survey Participation. Description and Illustration. *Public Opinion Quarterly*, 64: 299-308.

Hill, J., Smith, N., Wilson, D., and Wine, J. (2016). 2008/12 2012/14 Beginning Postsecondary Students Longitudinal Study (BPS:12/14). NCES 2016-062. Washington, DC: National Center for Education Statistics.

Kirchner, A., Emilia, P., and N. Tate. (2021, May). Can Prepaid PayPal Incentives Be as Effective as Prepaid Cash Incentives? Paper presented at the Annual Conference of the American Association for Public Opinion Research.

Lozar Manfreda, K., Bosnjak, M., Berzelak, J., Haas, I., and Vehovar, V. (2008). Web Surveys Versus Other Survey Modes. A Meta-Analysis Comparing Response Rates. *International Journal of Market Research*, 50(1): 79-104.

Medway, R.L. and Tourangeau, R. (2015). Response Quality in Telephone Surveys. Do Prepaid Incentives Make a Difference? *Public Opinion Quarterly*, 79(2): 524-543.

Messer, B.L., and Dillman, D.A. (2011). Surveying the General Public Over the Internet Using Address-Based Sampling and Mail Contact Procedures. *Public Opinion Quarterly*, 75(3): 429-457.

Olson, K., Smyth, J. D., and Wood, H. (2012). Does Giving People Their Preferred Survey Mode Actually Increase Survey Participation? An Experimental Examination. *Public Opinion Quarterly*, 76: 611-635.

Parsons, L., and Manierre, M.J. (2014). Investigating the Relationship among Prepaid Token Incentives, Response Rates, and Nonresponse Bias in a Web Survey. *Field Methods*, 26(2): 191-204.

Singer, E. (2002). The Use of Incentives to Reduce Nonresponse in Household Surveys. In Groves, R.M., Dillman, D. A., Eltinge, J.L., Little, R.J.A. (eds.), *Survey Nonresponse*. New York: Wiley.

Singer, E. and Ye, C. (2013). The Use and Effects of Incentives in Surveys. *Annals of the American Academy of Political and Social Science*, 645(1): 112-141.

Wilson, D., Shepherd, B., and J. Wine. (2015, May). The Use of a Calibration Sample in a Responsive Survey

Design. Paper presented at the Annual Conference of the American Association for Public Opinion Research, Hollywood, Florida.