2015–16 National Postsecondary Student Aid Study (NPSAS:16)

Data File Documentation

MAY 2018

Jennifer Wine
Peter Siegel
Rob Stollberg
RTI International

Tracy Hunt-White
Project Officer
National Center for Education Statistics

NCES 2018-482
U.S. DEPARTMENT OF EDUCATION

ies NATIONAL CENTER FOR EDUCATION STATISTICS
Institute of Education Sciences

U.S. Department of Education
Betsy DeVos
Secretary

Institute of Education Sciences
Mark Schneider
Director

National Center for Education Statistics
James L. Woodworth
Commissioner

The National Center for Education Statistics (NCES) is the primary federal entity for collecting, analyzing, and reporting data related to education in the United States and other nations. It fulfills a congressional mandate to collect, collate, analyze, and report full and complete statistics on the condition of education in the United States; conduct and publish reports and specialized analyses of the meaning and significance of such statistics; assist state and local education agencies in improving their statistical systems; and review and report on education activities in foreign countries.

NCES activities are designed to address high-priority education data needs; provide consistent, reliable, complete, and accurate indicators of education status and trends; and report timely, useful, and high-quality data to the U.S. Department of Education, the Congress, the states, other education policymakers, practitioners, data users, and the general public. Unless specifically noted, all information contained herein is in the public domain.

We strive to make our products available in a variety of formats and in language that is appropriate to a variety of audiences. You, as our customer, are the best judge of our success in communicating information effectively. If you have any comments or suggestions about this or any other NCES product or report, we would like to hear from you. Please direct your comments to

NCES, IES, U.S. Department of Education
Potomac Center Plaza (PCP)
550 12th Street, SW
Washington, DC 20202

May 2018

The NCES Home Page address is http://nces.ed.gov.
The NCES Publications and Products address is http://nces.ed.gov/pubsearch.

This publication is only available online. To download, view, and print the report as a PDF file, go to the NCES Publications and Products address shown above.

This report was prepared for the National Center for Education Statistics under Contract No. ED-05-CO-0033 with RTI International. Mention of trade names, commercial products, or organizations does not imply endorsement by the U.S. Government.

Suggested Citation
Wine, J., Siegel, P., Stollberg, R. (2018) *2015–16 National Postsecondary Student Aid Study (NPSAS:16) Data File Documentation* (NCES 2018-482). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Retrieved [date] from http://nces.ed.gov/pubsearch.

Content Contact
National Center for Education Statistics
NCES.Info@ed.gov
(800) 677-6987

Acknowledgments

We are greatly indebted to the students who generously participated in the survey. Their willingness to take the time to share their information and experiences made the 2015–16 National Postsecondary Student Aid Study (NPSAS:16) a success.

We gratefully acknowledge the assistance of the staff members at postsecondary institutions who provided data for the student and transcript data collections. We also extend special thanks to members of the Technical Review Panel (TRP) who provided their expertise to help shape NPSAS:16.

Contents

PAGE

PAGE

Appendices

List of Tables

TABLE PAGE

TABLE PAGE

List of Figures

Chapter 1. Overview of NPSAS:16

The National Postsecondary Student Aid Study (NPSAS) is a complex, nationally representative cross-sectional study of students attending postsecondary institutions eligible for student financial aid from the federal government. NPSAS covers topics pertaining to student enrollment in postsecondary education, with a focus on how individuals and families finance postsecondary education. It includes a student survey as well as the collection of data from the institutions in which the study students are enrolled and from other relevant data sources, including U.S. Department of Education records on student loan and grant programs and student financial aid applications. The result is a comprehensive dataset of student-level demographic and enrollment data with federal records on various forms of financial aid.

This report describes the design, methods, and results of the 2015–16 National Postsecondary Student Aid Study (NPSAS:16), conducted by the U.S. Department of Education's National Center for Education Statistics (NCES). The following legislation authorizes this and previous cycles of NPSAS, as well as the two longitudinal studies derived from NPSAS—the Beginning Postsecondary Students Longitudinal Study (BPS) and the Baccalaureate and Beyond Longitudinal Study (B&B):

- the Higher Education Act of 1965, as amended by the Higher Education Opportunity Act of 2008, 20 U.S.C. § 1015(a) (2012) and
- the Education Sciences Reform Act of 2002, 20 U.S.C. §§ 9541 to 9548 (2012).

Since 1987, NPSAS has been fielded every 3 to 4 years—most recently during the 2011–12 academic year—in response to the need for current information on financial aid programs. The regularity of NPSAS administration makes it possible to examine the impact of changes in federal policy concerning higher education over time. With respect to federal student aid, eligibility criteria change, grant and loan amounts fluctuate, and the balance between various aid options can shift dramatically. A recurring study such as NPSAS is essential to understanding those changes, particularly as they affect how students and families pay for college.

Chapter 1 of this report provides an overview of the background and purpose of NPSAS, as well as the study design, schedule, and products. Chapter 2 describes the sampling design and the steps NPSAS statisticians used to select the institution and

student samples. Chapter 3 describes the design, outcomes, and evaluation activities associated with institution data collection. Chapter 4 provides details on the student interview design, data collection, outcomes, and evaluations. Chapter 5 includes information on the student administrative records matching activities and outcomes. Chapter 6 contains a description of post-data collection data file processing and editing. Chapter 7 includes information on weighting, imputation, bias analysis, and variance estimation.

Tables and figures throughout this report present relevant analyses from the full-scale study. Unless otherwise indicated, a probability level of .05 was used for all tests of significance conducted for NPSAS:16 evaluations. Due to rounding, row and column entries in tables may not sum to their respective totals and reported percentages may differ somewhat from those that would result from the rounded numbers. Rounding is used to ensure the confidentiality of study respondents.

1.1 Background and Purpose

NPSAS is a comprehensive, nationwide study of how students and their families pay for postsecondary education. It features a nationally representative sample of undergraduate and graduate students enrolled in Title IV eligible postsecondary education institutions in the United States. The institution sampling frame includes public institutions and private institutions (both for-profit and nonprofit) and spans less-than-2-year institutions to 4-year colleges and universities.

NPSAS collections traditionally serve as the base-year data collection for one of two longitudinal studies, BPS and B&B. NPSAS:16 is the base-year data collection for the B&B:16 cohort of baccalaureate-completing college students, for which three follow-up collections are planned—in 2017 (B&B:16/17), in 2020 (B&B:16/20), and in 2026 (B&B:16/26). Consequently, subsets of questions in the NPSAS:16 student interview focus on the experience of B&B-eligible students in their last year of postsecondary education, including student debt accrual and repayment status, entry to graduate school, and the transition to employment.

1.2 Overview of NPSAS:16 Study Design

The data collected for NPSAS:16 come from three sources: (1) postsecondary institutions, (2) students, and (3) administrative data records. To facilitate selection of a nationally representative sample, the target population included all students enrolled in Title IV eligible postsecondary institutions during the 2015–16 academic year in the 50 states, the District of Columbia, and Puerto Rico.

NPSAS staff contacted institutions to request student-level information on enrollment and financial aid. Many of the required student financial aid data elements requested from institutions were also available for verification from the Central Processing System (CPS) and the National Student Loan Data System (NSLDS). CPS gathers information from the Free Application for Federal Student Aid (FAFSA) to determine federal aid eligibility. NSLDS contains student-level data on Pell Grants and federal student loans. NPSAS staff obtained these data through file matching with both CPS and NSLDS data to reduce the data collection burden on sampled institutions and sampled students. Early in the institutional data collection process, institutions confirmed participation in Title IV financial aid programs for study eligibility purposes and provided student enrollment lists for sampling purposes. Once NPSAS staff sampled students from the institution-provided enrollment lists, data were collected from students using a mobile enhanced web-based interview.

To supplement the institution and student interview data collections, NPSAS staff gathered additional data for the NPSAS:16 student sample from a variety of administrative data sources. These included the previously mentioned data obtained from CPS and NSLDS, as well as from the National Student Clearinghouse (NSC), ACT, the College Board, and the Veterans Benefits Administration (VBA).

1.3 Schedule and Products

Table 1 shows the schedule for the major activities of the full-scale study.

Table 1. Schedule of major activities for the NPSAS:16 full-scale study: 2015–18

NPSAS:16 activity	Start date	End date
Contacts with institutions to request student enrollment lists	Oct. 8, 2015	Feb. 26, 2016
Student enrollment list collection	Jan. 11, 2016	Jul. 10, 2016
Select student sample	Jan. 18, 2016	Jun. 1, 2016
Collect student data from institutional records	Feb. 10, 2016	Nov. 7, 2016
Student survey self-administered web-based data collection	Feb. 9, 2016	Nov. 8, 2016
Conduct telephone interviews with students	Mar. 2, 2016	Nov. 7, 2016
Process data, construct data files	Jan. 22, 2016	Oct. 2, 2017
Prepare/update reports	Aug. 1, 2016	Sept. 1, 2018

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

The following reports and web tables will be available on the NCES website at https://nces.ed.gov/surveys/npsas/:

- 2015–16 National Postsecondary Student Aid Study (NPSAS:16): Student Financial Aid Estimates for 2015–16;
- What Is the Price of College? Total, Net, and Out-of-Pocket Prices by Type of Institution in 2015–16;
- Web Tables—Student Financing of Undergraduate Education: 2015–16 (series);
- Web Tables—Profile of Undergraduate Students: 2015–16; and
- Web Tables—Profile of Graduate Students and Graduate Financial Aid Estimates: 2015–16.

NPSAS micro-level data files, associated codebooks, and data file documentation are available to researchers who have obtained a restricted-use data license from NCES. To apply for a restricted-use data license, visit the NCES website at https://nces.ed.gov/statprog/instruct.asp. Further information on the process for obtaining a restricted-use data license is available in the *NCES Restricted-Use Data Procedures Manual* at https://nces.ed.gov/statprog/rudman/.

The public may use NCES web tools (e.g., QuickStats, PowerStats, and TrendStats), found at https://nces.ed.gov/datalab, to access and analyze NPSAS:16 restricted-use data without having a restricted-use license. These tools permit analysis without disclosing data file contents to the user and suppress or flag any estimates that fail to meet reporting standards. QuickStats is an intuitive graphical tool that can generate simple tables and graphs. PowerStats offers greater analytic capabilities and can produce complex tables or estimate simple regression models. TrendStats allows users to produce averages, medians, and percentages over time for variables repeated across studies.

Chapter 2. Sampling Design

This chapter describes the target population, sampling design, and sampling methods for NPSAS:16. All documented procedures and methods were developed and refined in consultation with a Technical Review Panel (TRP) comprised of nationally recognized experts in higher education, staff from the National Center for Education Statistics (NCES), and representatives from other federal agencies.[1]

2.1 Respondent Universe

NPSAS:16 used a two-stage sampling design. The first stage involved the selection of institutions. In the second stage, students were selected from within sampled institutions. To construct the full-scale institution sampling frame for NPSAS:16, NPSAS statisticians used institution data collected from various surveys of the Integrated Postsecondary Education Data System (IPEDS). The student sampling frame included all students who met eligibility requirements from the sampled institutions. In the rare instance of a sampled institution lacking enrollment information, statisticians imputed the missing enrollment information using IPEDS imputation procedures.[2]

The NPSAS:16 institution (first stage) sampling frame (described below) included 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 United States. To be eligible for NPSAS:16, institutions must have met the following requirements:

- 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 administers the institution;

[1] See appendix A for a complete list of TRP participants.
[2] See https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2016111 for further detail on imputation in IPEDS.

- been located in the 50 states, the District of Columbia, or Puerto Rico;
- not been a U.S. service academy institution; and
- have signed the Title IV participation agreement with the U. S. Department of Education.[3]

Institutions providing only avocational, recreational, or remedial courses or only in-house courses for their own employees or members were ineligible. 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) were also excluded because of the academies' unique funding/tuition base.

The institution eligibility conditions for NPSAS:16 were consistent with the most recent iterations of NPSAS. The requirement that an institution must be eligible to distribute federal Title IV student aid was first implemented with NPSAS:2000. In NPSAS:2000 it was determined that there was sufficient comparability in survey design to ensure that important comparisons with data from previous NPSAS cycles could be made (Riccobono et al. 2002). Institutions that offered only correspondence courses, provided these same institutions were also eligible to distribute federal Title IV student aid, were first included in NPSAS:04. Finally, while institutions in Puerto Rico were not included in NPSAS:87 and NPSAS:12, they are included in NPSAS:16 and all other administrations of NPSAS.

The student (second stage) sampling frame is described below, and the requirements for NPSAS student eligibility have largely remained constant over time. For NPSAS:16, the target population consisted of all eligible students who were enrolled at any time between July 1, 2015 and June 30, 2016[4] at eligible postsecondary institutions in the United States 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 determined by the institution to be eligible for Title IV aid; or (4) an occupational or vocational program that required at least 3 months or

[3] 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.
[4] So as to not delay data collection, enrollment lists covered the period of July 1, 2015 through April 30, 2016. The date of April 30 was selected to include virtually all students enrolled prior to the summer term.

300 clock hours of instruction to receive a degree, certificate, or other type of formal award;

- not currently enrolled in high school; and
- not solely enrolled in a high school completion program.

2.2 Institution Sample

NPSAS statisticians created the NPSAS:16 full-scale institution sampling frame in a manner different than the three previous NPSAS studies, creating separate institution frames for the field test and the full-scale study. The field-test institution frame was constructed from the IPEDS 2013–14 Institutional Characteristics Header, 2013–14 Institutional Characteristics, 2012–13 12-month Enrollment, and 2012–13 Completions files. The full-scale institution frame was constructed from the same survey files for the following academic year. Creating two separate institution frames ensured a more accurate and current full-scale institution sample because each frame was constructed using the most up-to-date files.

To avoid overburdening institutional systems, those systems with two or more postsecondary institutions organized under the control of a single administrative entity and institutions likely to be selected with certainty (i.e., with a probability of selection equal to one) were removed from the field-test sampling frame. Similarly, most of the institutions selected for the field-test sample were removed from the full-scale frame.[5] The weights for the full-scale sample institutions were adjusted so that the sum of the weights would represent the full population of eligible institutions.

NPSAS statisticians selected 2,000 institutions using a variation of probability proportional to size (PPS) sampling called sequential probability minimum replacement (PMR) sampling (Chromy 1979). A composite size measure (Folsom, Potter, and Williams 1987) was used to help achieve self-weighting samples[6] 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. PMR sampling generally allows for institutions to be selected multiple times. Instead of allowing this, NPSAS statisticians ensured that all institutions with a probability of being selected more than once were instead included in the sample one time with

[5] During the full-scale study, the sampling design was revised to oversample public 4-year, non-doctorate-granting, primarily subbaccalaureate institutions. To ensure sufficient counts in the sector, 23 of these institutions were in both the field test and full-scale studies.
[6] Self-weighting samples have equal weights within sampling domains.

certainty. Institution composite measures of size were determined using enrollment data from the most recent IPEDS 12-month Enrollment and Completions files. Greater detail regarding the sampling process can be found in appendix B.

The 11 institution strata were based on institution level, control, level of offering, and highest level of offering[7]

1. public less-than-2-year;
2. public 2-year;
3. public 4-year, non-doctorate-granting, primarily subbaccalaureate;
4. public 4-year, non-doctorate-granting, primarily baccalaureate;
5. public 4-year, doctorate-granting;
6. private nonprofit, less-than-4-year;
7. private nonprofit, 4-year, non-doctorate-granting;
8. private nonprofit, 4-year, doctorate-granting;
9. private for-profit, less-than-2-year;
10. private for-profit, 2-year; and
11. private for-profit, 4-year.

As shown above, the stratum of public 4-year, non-doctorate-granting institutions was split into two strata—public 4-year institutions that were primarily subbaccalaureate and those that were primarily baccalaureate (strata 3 and 4). The subbaccalaureate institutions were usually community colleges that predominantly awarded subbaccalaureate degrees while offering bachelor's degrees in only a small number of select fields. Recent trends in enrollment show that over 40 percent of students in public 4-year, non-doctorate-granting institutions are enrolled at primarily subbaccalaureate institutions. Splitting the public 4-year, non-doctorate-granting institutions into two strata, rather than combining them, allows for oversampling and controlling the sample size of the subbaccalaureate institutions and students in them, including the baccalaureate recipients.

[7] The institution strata can be aggregated by control or level for the purposes of reporting institution counts.

The institution sampling rates and the numbers of institutions selected for each of the 11 institution strata are reported in table 2. Within each institution stratum, additional implicit stratification was accomplished by the following classifications:[8] (1) Historically Black Colleges and Universities (HBCUs) status; (2) Hispanic Serving Institutions (HSIs) status;[9] (3) INSTCAT (institution category derived using the level of offerings reported on the IPEDS Institutional Characteristics component and the number and level of awards that were reported on the IPEDS Completions component); (4) Carnegie classifications of degree-granting postsecondary institutions;[10] (5) the Office of Business Economics Region from the IPEDS Header file (Bureau of Economic Analysis of the U.S. Department of Commerce Region); (6) state and system for states with large systems (e.g., the SUNY and CUNY systems in New York, the state and technical colleges in Georgia, and the California State University and University of California systems in California); and (7) the institution measure of size. This implicit stratification helped ensure that the sample was approximately proportional to the population for these measures.

Table 2 shows institution sampling rates and the number of institutions sampled, by institution stratum.

[8] *Implicit stratification* is the process in which strata are created during the sampling process by sorting the data, rather than creating the strata prior to sampling and selecting an independent sample from each stratum.
[9] HSI status no longer exists in IPEDS. An HSI proxy was created using IPEDS Hispanic enrollment data.
[10] Some Carnegie categories were collapsed for the purposes of implicit stratification.

Table 2. Size of universe, institution sampling rates and number of institutions sampled, by institution stratum: 2015–16

Institution stratum[1]	Size of universe[2]	Sampling rate	Sample size
Total	**6,920**	**28.9**	**2,000**
Public			
Less-than-2-year	240	9.3	20
2-year	1,010	37.1	380
4-year, non-doctorate-granting, primarily subbaccalaureate	110	65.4	70
4-year, non-doctorate-granting, primarily baccalaureate	180	53.9	100
4-year, doctorate-granting	350	100.0	350
Private nonprofit			
Less-than-4-year	260	7.6	20
4-year, non-doctorate-granting	890	36.5	330
4-year, doctorate-granting	640	41.7	270
Private for-profit			
Less-than-2-year	1,630	4.3	70
2-year	910	13.2	120
4-year	690	40.7	280

[1] Institution stratum reflects institution categorization as determined from the 2013–14 Integrated Postsecondary Education Data System (IPEDS) files; some changes in this classification were identified when using more recent IPEDS files for weighting.
[2] Based on 2014–15 IPEDS data.
NOTE: Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

Table 3 shows the counts of sampled, eligible, and participating institutions, as well as the weighted and unweighted participation rates by control and level of institution.[11] Almost all the 2,000 sampled institutions met the eligibility requirements. Of those 2,000, approximately 1,750 provided enrollment lists. Overall, the NPSAS:16 institution response rate was commensurate with that of previous rounds of NPSAS.

[11] Unless otherwise indicated, references to "institution type," "institution stratum," or "institution characteristics" are hereafter interchangeable with control and level of institution. Control and level of institution are based on information from the sampling frame, which was formed from the IPEDS 2014–15 Institutional Characteristics Header, 2014–15 Institutional Characteristics, 2013–14 Completions, and 2013–14 12-month Enrollment files.

Table 3. Number of sampled and eligible institutions and number and percentage of institutions providing enrollment lists, by control and level of institution: 2015–16

Control and level of institution[1]	Sampled institutions	Eligible institutions	Institutions providing lists		
			Number	Unweighted percent	Weighted percent[2]
All institutions	**2,000**	**1,990**	**1,750**	**88.0**	**89.6**
Control of institution					
Public	920	920	830	90.2	90.2
Private nonprofit	610	600	530	87.9	88.2
Private for-profit	480	470	400	83.7	88.1
Level of institution					
Less-than-2-year	100	90	70	75.5	75.2
2-year	510	510	450	87.3	88.2
4-year, non-doctorate-granting	730	730	630	86.8	89.9
4-year, doctorate-granting	660	660	610	91.5	91.3
Control and level of institution					
Public less-than-2-year	20	20	20	77.3	77.9
Public 2-year	380	380	330	88.0	88.5
Public 4-year, non-doctorate granting, primarily sub-baccalaureate	70	70	70	92.9	95.3
Public 4-year, non-doctorate granting, primarily baccalaureate	100	100	90	90.6	89.7
Public 4-year, doctorate-granting	350	350	330	92.6	92.0
Private nonprofit, 2-year or less	20	20	20	94.4	94.2
Private nonprofit, 4-year, non-doctorate-granting	330	330	280	86.8	88.2
Private nonprofit, 4-year, doctorate-granting	270	270	240	89.2	88.2
Private for-profit, less-than-2-year	70	70	50	74.3	74.3
Private for-profit, 2-year	120	120	100	83.9	83.1
Private for-profit, 4-year	280	280	240	85.5	92.2

[1] Control and level of institution are based on data from the sampling frame, which was formed from the IPEDS 2014–15 Institutional Characteristics Header file.
[2] The weight used for this column is a base weight.
NOTE: Percentages are based on the unrounded count of eligible institutions within the row under consideration. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

2.3 Student Sample

Each sampled institution verified as NPSAS eligible was asked to provide a complete list of students who satisfied all NPSAS eligibility conditions. These lists included information needed to identify students for matching to administrative records, classify students to create sampling strata, and locate students to conduct the student survey. The student sample was randomly selected via stratified systematic sampling from lists of students enrolled between July 1, 2015 and April 30, 2016 at the sampled institutions. The following data items were requested from each sampled institution to form the student sampling frame;

- name;
- Social Security number (SSN);
- student ID number (if different from SSN);
- student level (undergraduate, masters, doctoral-research/scholarship/other, doctoral-professional practice, other graduate);
- class level of undergraduates (first year, second year, etc.);
- date of birth (DOB);
- Classification of Instructional Programs (CIP) or major;
- undergraduate degree program;
- high school/completion program completion date (month and year);
- baccalaureate recipient indicator (for students who have already received their bachelor's degree at the NPSAS institution since July 1, 2015);[11]
- potential baccalaureate recipient indicator (for students who are expected to receive their bachelor's degree at the NPSAS institution by June 30, 2016);[12]
- enrollment in high school (or completion program);
- date of first enrollment (at the postsecondary level);
- veteran status;
- grade point average (GPA);
- number of credits accumulated;
- account overdue (student owes fee that would prevent bachelor's degree award);
- race;
- ethnicity;
- sex;
- first-time graduate student at the NPSAS institution indicator; and
- contact information (local and permanent street address and phone number and school and home e-mail address).

[12] Splitting baccalaureate receipt into two items is based on the field test. It made providing baccalaureate information easier for institutions that could not identify the potential baccalaureate recipients and helped with QC checks against IPEDS counts for institutions that could not identify the potential baccalaureate recipients.

The 17 student sampling strata were

1. potential baccalaureate recipients who are veterans;
2. potential baccalaureate recipients from science, technology, engineering, and mathematics (STEM) programs;
3. potential baccalaureate recipients from teacher education programs;
4. potential baccalaureate recipients from business programs;
5. potential baccalaureate recipients from other programs;
6. other undergraduate students who are veterans;[13]
7. other undergraduate students;[12]
8. graduate students who are veterans;
9. first-time graduate students;
10. master's degree students in STEM programs;
11. master's degree students in education and business programs;
12. master's degree students in other programs;
13. doctoral-research/scholarship/other students in STEM programs;
14. doctoral-research/scholarship/other students in education and business programs;
15. doctoral-research/scholarship/other students in other programs;
16. doctoral-professional practice students; and
17. other graduate students.[14]

If students fell into multiple strata, such as students who were veterans or students with double majors, the ordering of the strata above was used to prioritize the stratification. Several student subgroups were intentionally sampled at rates differing from their natural occurrence within the population because of specific analytic objectives. Due to sheer numbers, sampling certain student groups (see undersampled student groups below) at a typical rate would have made it difficult to draw inferences about the experiences of other bachelor's degree, master's degree,

[13] *Other undergraduate students* are defined as any undergraduate student not classified as a potential baccalaureate recipient.
[14] *Other graduate* students are those who are not enrolled in a degree program, such as students just taking graduate courses.

and doctoral students, respectively. Table 4 shows the oversampled and undersampled student groups

Table 4. Oversampled and undersampled student groups: 2015–16

Oversampled student groups:	Undersampled student groups:
• potential baccalaureate recipients who are veterans • potential baccalaureate recipients from STEM programs • potential baccalaureate recipients from teacher education programs • other undergraduate students who are veterans • graduate students who are veterans • first-time graduate students • master's degree students in STEM programs • doctoral-research/scholarship/other students in STEM programs • students and potential baccalaureate recipients in public 4-year, non-doctorate-granting institutions that are primarily subbaccalaureate • undergraduate students at all award levels enrolled in for-profit institutions • master's degree students enrolled in for-profit institutions	• potential baccalaureate recipients from business programs • master's degree students in education and business programs • doctoral-research/scholarship/other students in education and business programs

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

To identify and sample veterans, NPSAS staff sent SSNs from the student enrollment lists to the Veterans Benefits Administration (VBA) for record matching. Students who were identified as veterans were placed in the appropriate veteran stratum.

Student SSNs from enrollment lists were also matched to National Student Loan Data System (NSLDS) data. The match results identified federal student financial aid recipients. Within each student stratum, individuals were sorted by whether or not they received federal aid and then systematically sampled such that the number of aided and unaided sampled students approximately matched the population proportions of aided and unaided students within the institution and student strata. This implicit stratification was done to help produce more accurate financial aid estimates. Substantial differences in the number of sample members receiving federal student loans had been observed in NPSAS:12 between the full-sample estimates and the poststratified estimates. This led to increased weight variation but, more importantly, could have led to bias in the final weighted estimates. Greater detail on VBA and NSLDS matching can be found in chapter 5.

As student lists were received from institutions, students were selected by means of stratified systematic sampling with predetermined sampling rates that varied by

student stratum. To eliminate cross-institution duplication, the SSNs of those selected from an institution were compared with SSNs of students who had already been selected from other institutions. Multiplicity adjustments in the sample weighting (described in more detail in chapter 7) accounted for students who had more than one chance of being selected because they had attended multiple institutions during the 2015–16 academic year.

Initial student sampling rates were calculated for each sample institution using sampling rates designed to generate approximately equal probabilities of selection within the institution-by-student sampling strata (see appendix B). In certain instances, NPSAS statisticians modified sampling rates as follows:

- Student sampling rates were increased for each institution to yield at least 10 students (if possible) to ensure sufficient yield for variance estimation.
- Student sampling rates were decreased, with few exceptions, if an institution sample size was greater than 300 students.[15]
- Student sampling rates were adjusted higher or lower based on expected yield calculations for institutions where the sample had not yet been selected.

These adjustments to the initial sampling rates resulted in some additional variability in the student sampling rates and increased survey design effects (variance inflation, see section 7.3). Table 5 shows the expected and achieved sample sizes by student sampling type and institution stratum.

[15] Because of their large enrollments, there were approximately 10 institutions that had a student sample size greater than 300.

Table 5. Expected and achieved numbers of sample students, by student type and control and level of institution: 2015–16

Control and level of institution	All students			Undergraduates									Graduate students		
				All			Potential baccalaureate			Other					
	Number expected	Number achieved	Percent achieved	Number expected	Number achieved	Percent achieved	Number expected	Number achieved	Percent achieved	Number expected	Number achieved	Percent achieved	Number expected	Number achieved	Percent achieved
All institutions	**126,320**	**122,030**	**96.6**	**105,260**	**99,080**	**94.1**	**37,590**	**37,890**	**100.8**	**67,670**	**61,190**	**90.4**	**21,050**	**22,950**	**109.0**
Public less-than-2-year	700	400	57.0	700	400	57.0	†	†	†	700	400	57.0	†	†	†
Public 2-year	21,780	18,210	83.6	21,780	18,180	83.4	†	50	†	21,780	18,130	83.2	†	30	†
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	5,750	5,850	101.7	5,670	5,790	102.0	2,620	2,780	106.0	3,060	3,010	98.6	80	60	75.0
Public 4-year, non-doctorate-granting, primarily baccalaureate	7,060	7,090	100.4	5,360	5,380	100.3	2,620	2,640	100.6	2,750	2,750	100.1	1,700	1,710	100.5
Public 4-year, doctorate-granting	25,980	26,830	103.3	20,860	20,630	98.9	9,740	9,210	94.5	11,120	11,430	102.8	5,120	6,190	121.0
Private nonprofit, 2-year or less	890	990	111.6	890	990	111.6	†	†	†	890	990	111.6	†	†	†
Private nonprofit, 4-year, non-doctorate-granting	12,040	11,300	93.9	9,210	8,730	94.8	5,000	4,730	94.7	4,210	3,990	94.8	2,830	2,580	91.1
Private nonprofit, 4-year, doctorate-granting	14,010	14,080	100.5	9,060	8,310	91.7	5,570	5,450	97.8	3,500	2,860	81.8	4,950	5,780	116.7
Private for-profit, less-than-2-year	3,440	2,610	75.8	3,440	2,610	75.8	†	#	†	3,440	2,600	75.7	†	†	†
Private for-profit 2-year	7,100	6,540	92.0	7,100	6,540	92.0	†	#	†	7,100	6,530	91.9	†	#	†
Private for-profit 4-year	27,560	28,140	102.1	21,180	21,540	101.7	12,050	13,040	108.2	9,130	8,500	93.1	6,380	6,600	103.4

† Not applicable.
# Rounds to zero.
NOTE: Some institution classifications of student type on the enrollment lists (e.g., potential baccalaureate, undergraduate, or graduate) were updated over the course of student interviewing; the statistics presented in this table are based on the original student sampling frame classification, not on the student's final classification. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

Table 6 shows the expected and achieved sample sizes, reported by student stratum. The achieved sample size of 122,030 was lower than the targeted 126,320 because many of the enrollment lists, particularly from for-profit institutions, contained fewer students than expected based on initial IPEDS counts. Sampling rates were adjusted, as described above, but the sample size still fell short of the target. Overall, more potential baccalaureate recipients and graduate students were selected into the sample than planned (for further details about sample allocation, see appendix B). Table 7 shows the initial classification of the student sample by student type and control and level of institution.

Table 6. Expected and achieved NPSAS:16 student samples, by student stratum: 2015–16

Student stratum[1]	Students sampled		
	Number expected[2]	Number achieved[3]	Percent acheived[4]
Total	**126,320**	**122,030**	**96.6**
Potential baccalaureate students	37,590	37,890	100.8
Veterans	3,000	4,030	134.2
Science, technology, engineering, and mathematics (STEM) majors	7,800	7,670	98.3
Teaching majors	5,000	4,000	80.1
Business majors	3,500	3,610	103.1
Other majors	18,290	18,590	101.6
Other undergraduate students	67,670	61,190	90.4
Veterans	4,000	4,870	121.8
Other	63,670	56,320	88.5
Graduate students	21,050	22,950	109.0
Veterans	2,850	3,370	117.9
First-time graduate students	3,000	1,560	52.0
Master's degree students in STEM programs	2,000	1,860	93.0
Master's degree students in education or business programs	2,000	2,250	112.3
Master's degree students in other programs	3,500	3,140	89.8
Doctoral—research/scholarship/other students in STEM programs	2,000	2,590	129.4
Doctoral—research/scholarship/other students in education or business programs	1,600	3,050	190.6
Doctoral—research/scholarship/other students in other programs	1,600	2,150	134.4
Doctoral—professional practice	2,000	2,130	106.5
Other graduate[6]	500	860	171.2

[1]Some institution classifications of student type on the enrollment lists (e.g., potential baccalaureate, undergraduate, or graduate) were updated over the course of student interviewing; the statistics presented in this table are based on the original student sampling frame classification, not on the student's final classification.
[2] Based on sample allocation and IPEDS 2013–14 12-month Enrollment and Completions counts.
[3] The student sample was drawn from 1,750 eligible institutions that provided enrollment lists.
[4] Percentage reported reflects the ratio of "achieved" to "expected."
[5] Other graduate students are those who are not enrolled in a degree program, such as students just taking graduate courses.
NOTE: Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

Table 7. Initial classification of NPSAS:16 student sample, by student type and control and level of institution: 2015–16

Control and level of institution	Total sample[1]		Student type[2,3]					
			Potential baccalaureate		Other undergraduate		Graduate	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
All institutions	**122,030**	**100.0**	**37,890**	**100.0**	**61,190**	**100.0**	**22,950**	**100.0**
Control of institution								
Public	58,370	47.8	14,670	38.7	35,710	58.4	7,990	34.8
Private nonprofit	25,510	20.9	10,180	26.9	6,990	11.4	8,340	36.3
Private for-profit	38,150	31.3	13,050	34.4	18,490	30.2	6,620	28.8
Level of institution								
Less-than-2-year	3,170	2.6	#	#	3,170	5.2	†	†
2-year	25,570	21.0	50	0.1	25,480	41.6	40	0.2
4-year, non-doctorate-granting	43,500	35.6	20,820	55.0	16,550	27.0	6,130	26.7
4-year, doctorate-granting	49,790	40.8	17,010	44.9	15,990	26.1	16,790	73.1
Control and level of institution								
Public less-than-2-year	400	0.3	†	†	400	0.6	†	†
Public 2-year	18,210	14.9	50	0.1	18,130	29.6	30	0.1
Public 4-year, non-doctorate-granting	5,850	4.8	2,780	7.3	3,010	4.9	60	0.3
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	7,090	5.8	2,640	7.0	2,750	4.5	1,710	7.4
Public 4-year, non-doctorate-granting, primarily baccalaureate	26,830	22.0	9,210	24.3	11,430	18.7	6,190	27.0
Private nonprofit, 2-year or less	990	0.8	†	†	990	1.6	†	†
Private nonprofit, 4-year, non-doctorate-granting	11,300	9.3	4,730	12.5	3,990	6.5	2,580	11.2
Private nonprofit, 4-year, doctorate-granting	14,080	11.5	5,450	14.4	2,860	4.7	5,780	25.2
Private for-profit, less-than-2-year	2,610	2.1	#	#	2,600	4.3	†	†
Private for-profit, 2-year	6,540	5.4	#	#	6,530	10.7	#	#
Private for-profit, 4-year	28,140	23.1	13,040	34.4	8,500	13.9	6,600	28.8

Rounds to zero.
† Not applicable.
[1] The student sample was drawn from the 1,750 eligible institutions that provided enrollment lists.
[2] Some institution classifications of student type on the enrollment lists (e.g., potential baccalaureate, undergraduate, or graduate) were updated over the course of student interviewing; the statistics presented in this table are based on the original student sampling frame classification, not on the student's final classification.
[3] The five potential baccalaureate strata have been combined. The two other undergraduate strata have been combined. The graduate veterans stratum, first-time graduate stratum, three master's strata, four doctoral strata, and other graduate stratum have been combined.
NOTE: Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

Chapter 3. Institution Data Collection Design, Outcomes, and Evaluation

The following chapter describes the design, implementation, and outcomes of institution data collection for NPSAS:16. It includes detail on the institution contacting process, an overview of the technical systems put in place to assist in data collection efforts, and evaluation of enrollment list and student data quality.

3.1 Institution Data Collection Design and Systems

NPSAS:16 institution data were collected in several stages, using systems designed for the contacting sample members and data collection processes. NPSAS project staff were trained using an Institution Contacting System (ICS) to record data on any communications with institution staff. A Postsecondary Data Portal (PDP) website was created as the data collection repository, a single location for consolidating information about the study and contact information for the help desk and project staff.

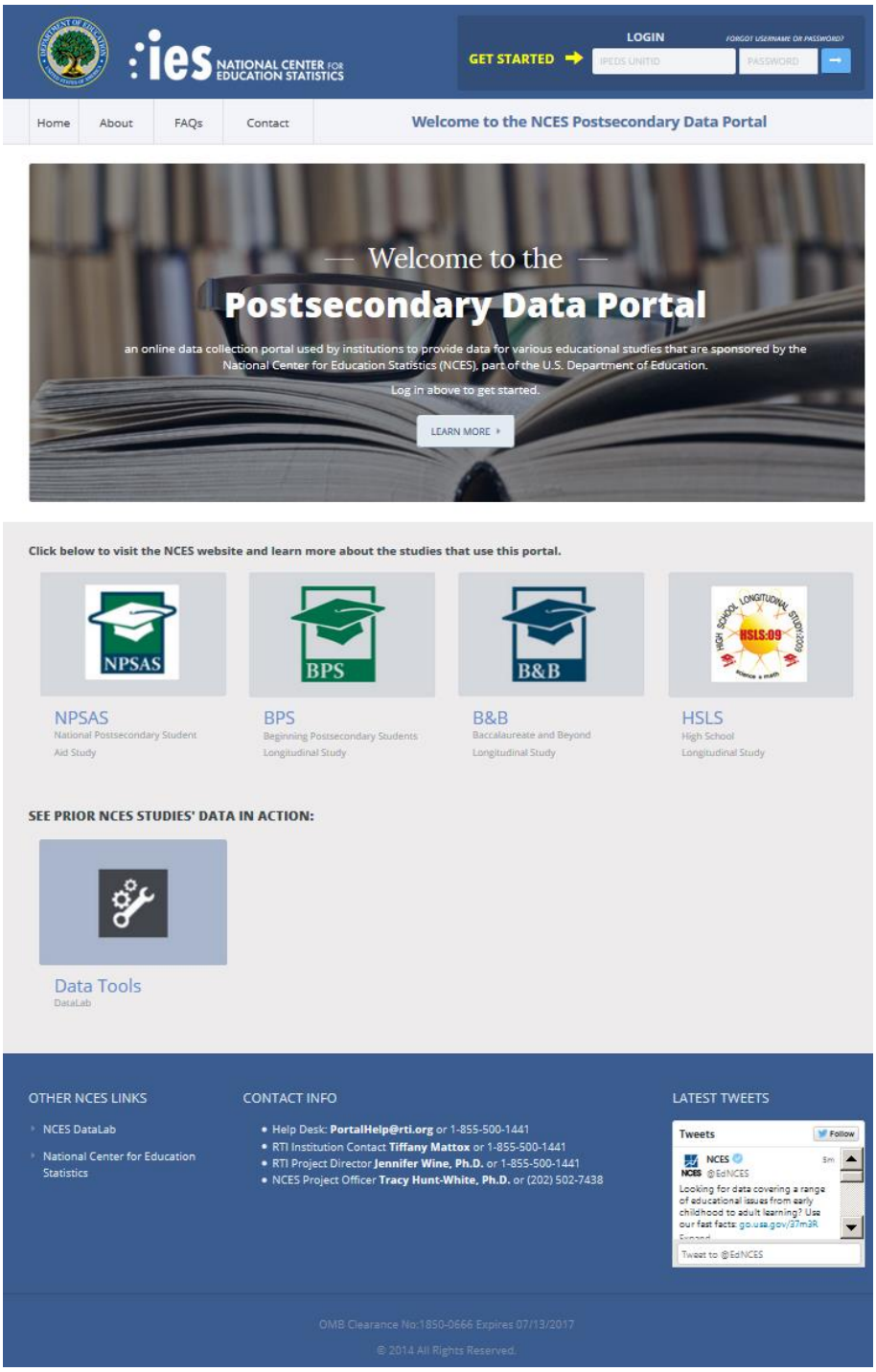
3.1.1 Institution Contacting System

Project staff used the ICS for scheduling and to track NPSAS data collection participation at the institution level. The ICS served as a record-keeping system for all outbound and incoming communications with institutions, regardless of format (telephone call, e-mail, U.S. mail). The reporting functions of the ICS allowed project staff to view the overall progress of institution recruitment, enrollment list collection, and student records collection.

3.1.2 Institution Website

All institution data collection was conducted through the PDP. The PDP is a secure platform for uploading requested electronic data and provides reliable, user-friendly access to both general and study-specific documents. From the PDP, institution representatives could access a frequently asked questions (FAQs) section that spanned NCES studies; a tutorial on use of the PDP; and contact information for the help desk, RTI staff, and NCES staff. Figure 1 shows the home page of the PDP website.

Figure 1. NCES Postsecondary Data Portal home page: 2015–16



SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

3.1.3 Student Records Collection System

The NPSAS:16 student records instrument consisted of five sections:

1. Institution Information, which collected terms of enrollment at the NPSAS institution in the academic year.

2. General Student Information, which collected student characteristics and contact information.

3. Enrollment, which collected the degree program, major(s), class level, and enrollment intensity at the NPSAS institution.

4. Budget, which included budgeted costs of attendance.

5. Financial Aid, which collected federal, state, institution, graduate, and any private or other financial aid awards received by the student.

The data elements collected in each of these sections were consistent with items collected in the NPSAS:16 field test (see appendix O) and in prior NPSAS administrations. The full-scale instrument was refined with input from the NPSAS:16 Technical Review Panel (TRP; see https://edsurveys.rti.org/npsas16trp/ for a summary of the meetings), focus groups conducted with institutions that participated in prior student record collections, and the results of the NPSAS:16 field test. Instrument changes were intended to ensure consistency of data elements across NCES postsecondary studies that collect student records data; improve the clarity of item definitions; and enhance the usability of the PDP for participating institutions. For example, based on focus group feedback, the class level item was revised to remove response options for 2nd- and 3rd-year graduate students, which institutions reported is difficult for them to assess beyond the first year. Instead, institutions were offered response options for "1st-year graduate student" and "beyond 1st-year graduate student." In addition, new items were added to the instrument, including SAT and ACT test scores, credit/clock hours in program and cumulative credit/clock hours completed, and degree completion date. Appendix C includes a list of all items collected in the NPSAS:16 full-scale student records instrument.

The NPSAS:16 student records instrument could be completed in three modes:

1. Web mode, in which institution staff used drop-down boxes and text-entry fields to key data directly on the PDP website, one student at a time.

2. Excel mode, in which institutions downloaded a preformatted Excel spreadsheet template from the PDP, keyed or copied student data into a

spreadsheet template offline, and then uploaded the completed template to the PDP website.

3. Comma-separated values (CSV) mode, in which institutions downloaded customized file specifications from the PDP website, prepared data files offline according to the file specifications, and then uploaded completed files to the PDP website.

Institutions could choose any of these modes, or use a combination of them, to provide student records data.

Between the NPSAS:16 field test and full-scale collection, changes were made to the PDP website to improve its usability and the quality of the data collected. For example, a session timeout warning was implemented, which would launch a pop-up dialog box on the PDP to alert users that their log-in sessions were about to expire. In addition, an abbreviated student records instrument was prepared, which was offered as a refusal conversion and avoidance tool to institutions that expressed reluctance to provide the full set of student records data elements.

3.2 Institution Contacting, Recruitment, and Student Enrollment List Acquisition

At the outset of institution data collection, NPSAS staff contacted the sampled institutions to secure their participation in the study. They asked institutions to designate an institution coordinator to act as a primary point of contact for the submission of student enrollment lists. These activities are described in the following sections.

3.2.1 Institution Contacting and Recruitment

NPSAS staff began notifying sampled institutions of the impending student record collection in October 2015, roughly 4 months before the earliest enrollment list submission deadlines in January 2016. Early notification was intended to provide institutions with study requirements, deadlines, and enough time to allocate the staff and resources needed to submit data on schedule. This notice was also intended to provide institutions with sufficient time for internal review and approval and resolution of any potential obstacles to participation.

NPSAS project staff trained eleven people to contact institutions, three of whom had experience as institution contactors in prior NPSAS studies (NPSAS:12 or the NPSAS:16 field test). Their training included an overview of the NPSAS:16 full-scale

study, guidance in building strong working relationships with institution coordinators (ICs), and instruction in assisting with data collection and submission using the PDP. Institution contactors identified institutions as potentially ineligible—including closed institutions and those not open to the general public. They also identified instances where sampled institutions had merged with other institutions (whether sampled or not sampled). The identified institutions were reviewed by NPSAS project staff to confirm if the institution was ineligible for the study

To encourage participation and confirm the legitimacy of the study in advance of data collection, institution contactors provided chief administrators with a list of postsecondary organizations and associations that have endorsed NPSAS (see appendix D). The same organizations endorsed NPSAS in 2012 and for the NPSAS:16 field test. All correspondence with institutions, including letters, brochures, and the project website featured the endorsement list.

Once institution contactors had verified contact information, they sent chief administrators a packet of information about the study (reproduced in appendix E). The information packet included the following materials:

- a cover letter printed on NCES letterhead providing background information on NPSAS and requesting that the chief administrator designate an IC via the PDP website;
- website access instructions;
- a brochure that summarized NPSAS objectives and provided background information and key findings from past NPSAS cycles; and
- a schedule and flow chart of all NPSAS data collection activities.

Two days after NPSAS staff sent the mailing to chief administrators, institution contactors made their first calls to chief administrators' offices to prompt for designation of ICs. If chief administrators were unable or unwilling to log in to the website to designate a coordinator, they could provide that information over the telephone.

Once the institution named a coordinator, institution contactors confirmed study participation and set a deadline for submission of the student enrollment list. NPSAS staff customized deadlines according to the institution's term structure. For institutions with distinct terms, ICs were asked to provide the start and end dates for the term that included April 30, 2016. Institution contactors set the institution's deadline for 2 weeks after the start of that term. For institutions with continuous

enrollment, institution contactors asked the ICs to provide lists by May 15. Contactors communicated as needed with ICs to offer reminders of the scheduled due date and to find out if they had any questions.

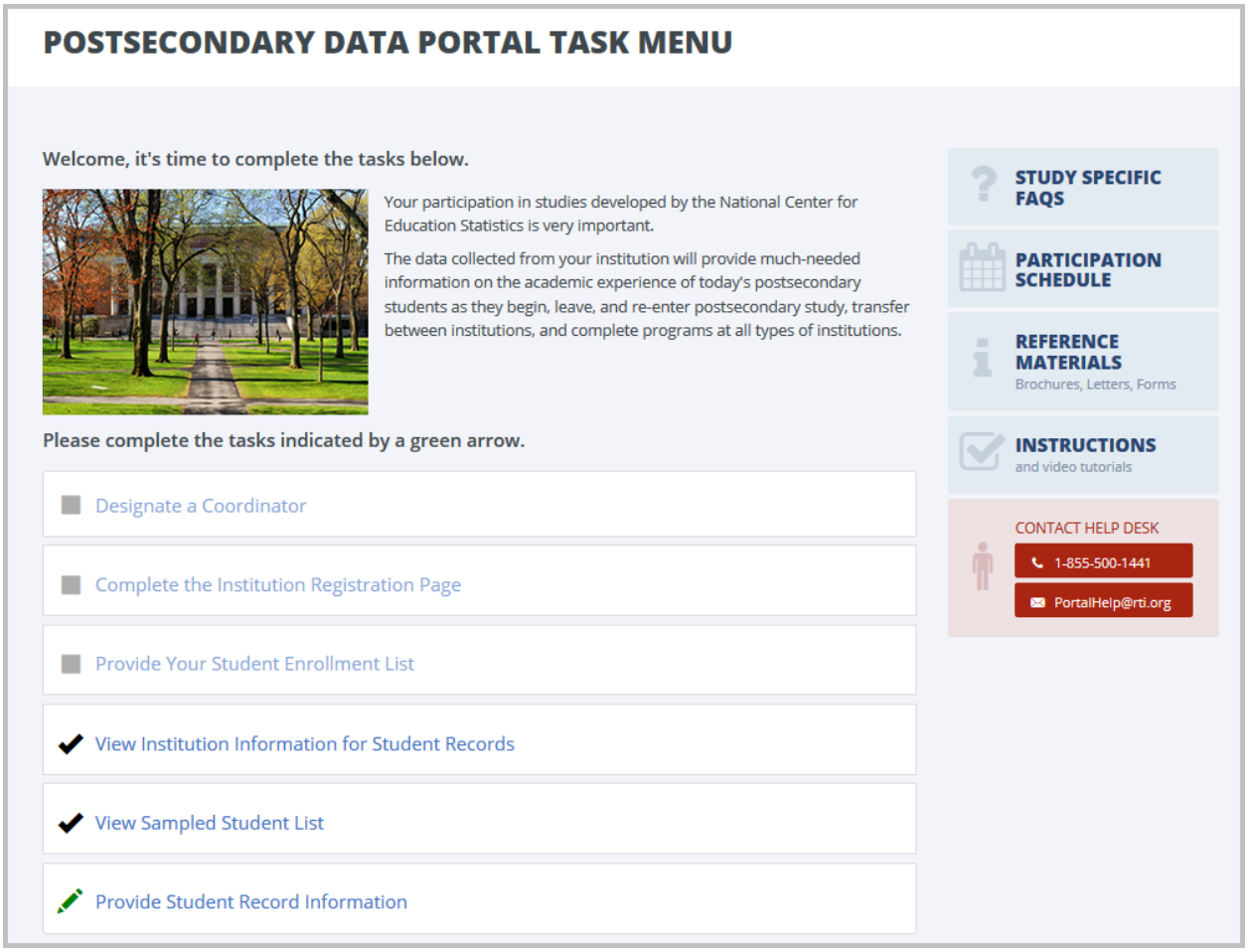
After an IC was designated and deadlines were communicated, institution contactors sent the following materials to ICs (see appendix H):

- a cover letter describing the study with information on how to access the PDP website and complete the Institution Registration Page (IRP);
- a brochure describing the study; and
- a schedule and flow chart of all NPSAS data collection activities.

Institution contactors then followed up by telephone to confirm receipt and prompt for completion of the IRP. After ICs completed the IRP, institution contactors asked them to review the variables on the student list to gain an understanding of the size and scope of the request. They were encouraged to contact the NPSAS help desk with any questions or concerns.

Institution contactor staff continued their follow-ups, as appropriate, to ensure timely completion of the student enrollment list request. All ICs were prompted by telephone prior to their scheduled deadlines and sent a reminder via e-mail. For convenience, the e-mail prompt contained the institution's log-in information and a link to the PDP website. Once logged in, an IC could view a task menu indicating the stages of data collection already completed (denoted by a black check mark) and outstanding tasks (indicated by a green pencil). This design allowed institutions to identify the tasks that were not yet completed and monitor their overall progress. See figure 2 for an example of the PDP task menu.

Figure 2. PDP website task menu: 2015–16



SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

NPSAS project staff identified large campus systems with centralized record keeping at the start of data collection using IPEDS reporting data. The ICs for these systems were given the option of reporting for constituent institutions individually or at the system level. These reporting options greatly increased the efficiency of data collection and reduced burden for these large campus systems. Project staff worked with these systems directly to provide guidance on reporting and to accommodate any ad hoc quests.

3.2.2 Student Enrollment List Acquisition

As described in section 3.2.1, institutions were formally asked to provide enrollment list information for all students enrolled at any time between July 1, 2015 and

April 30, 2016.[16] The PDP website provided ICs with complete instructions for preparing a student list. Institution contactors clarified or elaborated on the instructions in follow-up telephone conversations as necessary.

As part of the enrollment list acquisition process, institutions were encouraged to upload their student enrollment lists using the secure upload interface on the PDP website. If an institution could not upload data due to firewall issues or other technical limitations, enrollment lists could be e-mailed as compressed, encrypted files. Because of the potential risk to data security, the sending of physical copies of the lists was not permitted.

Once ICs submitted enrollment lists, NPSAS project staff performed several checks on the quality and completeness before selecting the student sample. These included verifying that institutions used a readable format and that key data needed for sampling and initial locating (e.g., baccalaureate indicator, SSN, contact information) were provided. If staff detected problems with lists during quality checks, they contacted institutions to resolve any issues.

3.2.3 Institution Recruitment and Student List Acquisition Outcomes

Of the total sample of 1,990 eligible institutions, 95 percent initially agreed to participate by designating an IC. Eighty-one percent of eligible institutions completed the IRP. One of the purposes of the IRP is to help determine the due date for the student enrollment list. Several large institutions were asked not to complete the IRP because their list due dates were negotiated with them directly. Approximately 88 percent of the eligible institutions provided usable student enrollment lists. Thirteen lists were rejected, and omitted from these counts, for having either too many missing items for the list to be usable or obvious errors (e.g., mismatched e-mail addresses). Approximately 38 percent of the lists arrived during the first 2 months of list collection, a rate consistent with the NPSAS:12 collection. Eighty-nine percent of the 1,500 institutions that had previously participated in a NPSAS data collection provided enrollment lists for NPSAS:16, a rate statistically different from the rate among institutions that had not previously participated (84 percent), $\chi^2(1, n = 1{,}990) = 10.09$, $p < .05$.

The percentage of institutions providing enrollment lists across strata ranged from 74 percent to 94 percent. The lowest participation rates were among the private for-

[16] The NPSAS:16 target population consisted of all eligible students enrolled at any time between July 1, 2015 and June 30, 2016. However, most institutions provided enrollment lists that covered the period of July 1, 2015 through April 30, 2016. The date of April 30 was selected to include virtually all students enrolled prior to the summer term without delaying data collection.

profit, less than 2-year sector (74 percent) and public, less-than-2-year institutions (77 percent). Table 8 presents enrollment list collection results by institution level, control, and type.

Table 8. Enrollment list receipt, by control and level of institution: 2015–16

Control and level of institution	Total eligible institutions	Institution-level response rate		Sampled students
		Number	Percent	
Total	**1,990**	**1,750**	**88.0**	**122,030**
Control of institution				
Public	920	830	90.2	58,370
Private nonprofit	600	530	87.9	25,510
Private for-profit	470	400	83.7	38,150
Level of institution				
Less-than-2-year	90	70	75.5	3,170
2-year	510	450	87.3	25,570
4-year, non-doctorate-granting	730	630	86.8	43,500
4-year, doctorate-granting	660	610	91.5	49,790
Control and level of institution				
Public				
Less-than-2-year	20	20	77.3	400
2-year	380	330	88.0	18,210
4-year, non-doctorate-granting, primarily subbacalaureate	70	70	92.9	5,850
4-year, non-doctorate-granting, primarily baccalaureate	100	90	90.6	7,090
4-year, doctorate-granting	350	330	92.6	26,830
Private nonprofit				
Less-than-4-year	20	20	94.4	990
4-year, non-doctorate-granting	330	280	86.8	11,300
4-year, doctorate-granting	270	240	89.2	14,080
Private for-profit				
Less-than-2-year	70	50	74.3	2,610
2-year	120	100	83.9	6,540
4-year	280	240	85.5	28,140

NOTE: All percentages are unweighted and based on the number of eligible institutions within the row under consideration. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

NPSAS staff evaluated enrollment lists for accuracy and completeness in several ways, including comparing institution-provided data to the IPEDS data for the same institutions. Staff then contacted institutions that had submitted student counts with discrepancies to reconcile the data. Approximately 76 percent of the enrollment lists had no problems identified during quality checks.

Several data elements new to NPSAS administrations were added to the enrollment list request in the NPSAS:16 field test and full-scale study. GPA, number of credits,

and account overdue status were requested to potentially help verify baccalaureate recipients during sampling. Over 92 percent of the institutions provided GPA and credits, but overdue status proved harder for institutions to provide, with about 78 percent of institutions reporting data for this variable. Sex and race/ethnicity were requested to help with weighting and nonresponse bias analysis, and over 96 percent of institutions provided these items. A class level response option indicating enrollment in the first year of a graduate program was requested to determine if institutions could provide accurate information regarding first-time graduate student status. Approximately 89 percent of institutions provided data for this variable. Table 9 shows the percentage of students for whom these new data elements were provided by sampled institution control and level.

Table 9. Institutions providing GPA, number of credits, account overdue status, sex, first-time graduate student indicator, and race/ethnicity, by control and level of institution: 2015–16

Control and level of institution	GPA		Number of credits		Account overdue status		Sex		First-time graduate student indicator		Race/ethnicity	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Total	**1,620**	**92.3**	**1,620**	**92.5**	**1,370**	**78.1**	**1,690**	**96.4**	**1,560**	**89.1**	**1,690**	**96.5**
Public												
Less-than-2-year	10	41.2	10	41.2	10	70.6	20	100.0	10	76.5	20	100.0
2-year	320	97.9	320	97.9	230	68.0	330	99.4	260	79.2	330	99.1
4-year, non-doctorate-granting, primarily subbaccalaureate	60	95.4	60	95.4	50	80.0	60	95.4	60	87.7	60	93.9
4-year, non-doctorate-granting, primarily baccalaureate	80	93.1	80	95.4	60	70.1	80	95.4	80	95.4	80	96.6
4-year, doctorate-granting	310	93.9	310	94.2	260	80.7	310	95.4	310	94.8	310	95.4
Private nonprofit												
Less-than-4-year	10	70.6	10	52.9	10	82.4	20	100.0	10	64.7	20	100.0
4-year, non-doctorate-granting	270	94.7	270	95.7	250	88.7	280	98.2	260	91.8	280	98.6
4-year, doctorate-granting	230	95.0	230	96.2	190	81.2	240	99.2	240	98.3	240	99.6
Private for-profit												
Less-than-2-year	30	61.5	30	50.0	40	71.2	50	98.1	40	75.0	50	98.1
2-year	90	86.9	90	90.9	60	57.6	90	92.9	80	81.8	90	92.9
4-year	210	89.8	210	89.8	200	85.6	210	89.8	210	89.4	210	90.3

NOTE: All percentages are unweighted and based on the number of eligible institutions within the row under consideration. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

Project staff also examined institution participation by selected classifications categories used for implicit stratification (see section 2.2), including 2010 Carnegie classification categories (table 10). Of the 1,750 institutions that provided enrollment lists in NPSAS:16, 190 did not have a Carnegie classification. Of those with a known Carnegie classification, participation ranged from a high of 290 institutions classified as master's (larger programs) to five classification categories with participation numbers that rounded to zero.

Table 10. Number and percentage of participating NPSAS:16 institutions, by 2005 Carnegie institutional classification: 2015–16

2005 Carnegie institutional classification	Number	Percent
All institutions	**1,750**	**100.0**
Not classified	190	10.6
Associate's—public rural-serving small	10	0.5
Associate's—public rural-serving medium	60	3.3
Associate's—public rural-serving large	70	3.8
Associate's—public suburban single campus	40	2.5
Associate's—public suburban multicampus	60	3.2
Associate's—public urban single campus	20	1.0
Associate's—public urban multicampus	90	5.0
Associate's—public special use	#	0.1
Associate's—private nonprofit	10	0.6
Associate's—private for-profit	90	4.9
Associate's—public 2-year under 4-year	10	0.5
Associate's—public 4-year, primarily associate's	20	1.3
Associate's—private nonprofit, 4-year, primarily associate's	#	0.1
Associate's—private for-profit, 4-year, primarily associate's	20	1.1
Research (very high research activity)	100	5.7
Research (high research activity)	80	4.7
Doctoral/research universities	70	3.7
Master's (larger programs)	290	16.3
Master's (medium programs)	100	5.9
Master's (smaller programs)	50	3.0
Baccalaureate colleges-arts and sciences	90	5.4
Baccalaureate colleges-diverse fields	100	5.6
Baccalaureate/associate's colleges	50	2.6
Special focus—theological	10	0.5
Special focus—medical	30	1.5
Special focus—other health professions	20	1.3
Special focus—engineering	#	0.2
Special focus—other technology	10	0.8
Special focus—business/management	20	0.9
Special focus—art, music, and design	50	2.7
Special focus—law	10	0.6
Special focus—other special-focus	#	0.1
Tribal colleges	#	0.1

Rounds to zero.
NOTE: Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

Similarly, table 11 shows the number of HBCUs, also used for implicit stratification, participating in the current and prior NPSAS rounds. Forty HBCUs participated in NPSAS:16.

Table 11. Participation of Historically Black Colleges and Universities: 2015–16

Cycle participated	Number of HBCUs participating	HBCUs as a percentage of total number of participating institutions
NPSAS:87	20	1.9
NPSAS:90	20	1.5
NPSAS:93	30	2.6
NPSAS:96	20	1.9
NPSAS:2000	20	2.3
NPSAS:04	30	2.1
NPSAS:08	40	2.3
NPSAS:12	30	2.0
NPSAS:16	40	2.0

NOTE: HBCUs = Historically Black Colleges and Universities. NPSAS = National Postsecondary Student Aid Study. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

3.3 Student Records Data Collection

Once institutions sent student enrollment lists, NPSAS project staff created the student sample as detailed in chapter 2 and began to collect institution record data for sample members. The following section describes student records collection and outcomes.

3.3.1 Student Records Collection From Institutions

After a student sample was selected for a particular institution, NPSAS staff sent the designated IC an information packet on the student records collection process. These packets included instructions for accessing the PDP website and a Quick Guide to Providing Student Records Data (see appendix E). The secure website contained a list of the sampled students, customized for each institution, and instructions and system requirements required for web data entry or file upload. Specific instructions on how to construct the requested data files (either by template or programmatically) were also available. Several features of the website—help text, a help desk telephone number, and an e-mail link for problem reports—were included to help institutions provide data. Help desk project staff were made available to provide assistance if institution staff had questions or encountered problems.

3.3.2 Student Records Collection Outcomes

Of the 1,750 institutions with sampled students, 93 percent provided student records data.[17] Most institutions opted for Excel mode (62 percent), 30 percent uploaded CSVs, and the remaining 8 percent entered data into the PDP student records interface. Table 12 shows student records collection results by control and level of institution and student type. From the institutions that provided student records data, NPSAS staff obtained student-level data for 93 percent of eligible sample members. This total included approximately 92 percent of the total undergraduate students in the sample and 97 percent of the graduate students. NPSAS project staff collected student records for 95 percent of the students identified as potentially eligible for inclusion in the B&B longitudinal follow-up study (see section 4.4.7).

Table 12. Student record collection methods, by control and level of institution: 2015–16

Control and level of institution	Institutions providing enrollment lists	Institutions providing student records		Web mode		Excel mode		CSV mode	
		Number	Percent	Number	Percent	Number	Percent	Number	Percent
Total	**1,750**	**1,640**	**93.3**	**130**	**8.2**	**1,010**	**61.5**	**500**	**30.3**
Control of institution									
Public	830	790	94.9	60	8.0	480	60.6	250	31.3
Private nonprofit	530	510	96.4	60	11.2	410	80.6	40	8.2
Private for-profit	400	340	85.8	10	4.1	120	34.8	210	61.1
Level of institution									
Less-than-2-year	70	60	90.1	10	20.3	40	64.1	10	12.5
2-year	450	410	91.9	30	8.3	250	61.1	130	31.3
4-year, non-doctorate-granting	630	580	91.9	40	7.6	340	59.1	190	33.0
4-year, doctorate-granting	610	580	96.2	40	7.4	370	64.0	170	28.8

NOTE: Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

[17] Five percent of the responding institutions completed the abbreviated student records instrument.

Table 13. Student records collection results, by control and level of institution and student type: 2015–16

Control and level of institution and student type	Institutions providing enrollment lists	Institutions providing student records		Total eligible students[1]	Student records collected	
		Number	Percent		Number	Percent
Total	**1,750**	**1,640**	**93.3**	**119,550**	**110,930**	**92.8**
Control of institution						
Public	830	790	94.9	56,850	53,470	94.1
Private nonprofit	530	510	96.4	25,170	24,320	96.6
Private for-profit	400	340	85.8	37,530	33,130	88.3
Level of institution						
Less-than-2-year	70	60	90.1	3,050	2,610	85.4
2-year	450	410	91.9	24,510	21,920	89.4
4-year, non-doctorate-granting	630	580	91.9	42,730	39,100	91.5
4-year, doctorate-granting	610	580	96.2	49,260	47,300	96.0
Control and level of institution						
Public						
Less-than-2-year	20	20	100.0	370	360	98.4
2-year	330	310	93.1	17,350	15,780	90.9
4-year, non-doctorate-granting, primarily subbaccalaureate	70	60	93.8	5,610	5,300	94.4
4-year, non-doctorate-granting, primarily baccalaureate	90	80	96.6	6,950	6,590	94.8
4-year, doctorate-granting	330	320	96.3	26,570	25,440	95.8
Private nonprofit						
Less-than-4-year	20	20	100.0	960	950	99.4
4-year, non-doctorate-granting	280	270	96.5	11,140	10,750	96.4
4-year, doctorate-granting	240	230	95.8	13,910	13,420	96.5
Private for-profit						
Less-than-2-year	50	50	86.5	2,520	2,080	82.6
2-year	100	90	86.9	6,360	5,360	84.1
4-year	240	200	85.2	27,810	24,900	89.5
Student type						
Total undergraduate	†	†	†	95,020	87,210	91.8
Potential B&B student[2]	†	†	†	33,760	32,190	95.4
Other undergraduates	†	†	†	62,510	56,210	89.9
Graduate	†	†	†	24,530	23,710	96.7

† Not applicable.
[1] Total eligible students sampled from 1,750 institution enrollment lists.
[2] Students receiving baccalaureate in 2015–16, count includes graduate students who earned baccalaureate during 2015–16.
NOTE: B&B = Baccalaureate and Beyond Longitudinal Study. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

3.4 Institution Data Evaluation

The following section describes the evaluation processes followed to assess data collection outcomes and quality of the collected data.

3.4.1 Evaluation of Enrollment List Quality

Project staff evaluated enrollment lists for the presence of selected key variables, including contact information, SSN, DOB, and, for the first time in NPSAS, high school graduation date. As shown in table 14, about 99 percent of the enrollment lists used for sampling included street addresses, about 98 percent included e-mail addresses, and about 97 percent included SSNs. NPSAS staff used high school graduation date to identify ineligible students on the enrollment lists, including students concurrently enrolled in high school who were initially identified as first-time beginning students. About 82 percent of the lists used for sampling contained high school graduation date.

Table 14. Institutions providing student contact information, Social Security number, date of birth, and high school graduation date, by control and level of institution: 2015–16

Control and level of institution	Address		Social Security number		Date of birth		E-mail address		Telephone number		High school graduation date	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Total	**1,740**	**99.1**	**1,700**	**96.8**	**1,680**	**96.1**	**1,720**	**97.9**	**1,660**	**94.7**	**1,440**	**82.0**
Public												
Less-than-2-year	20	94.1	20	100.0	20	100.0	10	82.4	20	88.2	10	58.8
2-year	330	100.0	330	98.2	330	99.4	330	99.4	330	98.8	310	94.6
4-year, non-doctorate-granting, primarily subbaccalaureate	70	100.0	60	93.9	60	93.9	60	95.4	60	95.4	60	92.3
4-year, non-doctorate-granting, primarily baccalaureate	90	100.0	90	98.9	90	97.7	90	97.7	80	94.3	80	89.7
4-year, doctorate-granting	330	99.7	310	93.9	310	94.8	320	96.3	300	92.7	270	83.2
Private nonprofit												
Less-than-4-year	20	100.0	20	94.1	20	100.0	20	100.0	20	100.0	10	70.6
4-year, non-doctorate-granting	280	99.3	280	98.6	280	97.5	280	99.7	270	94.0	220	78.0
4-year, doctorate-granting	240	98.3	230	95.4	240	98.3	240	99.2	230	94.6	180	77.0
Private for-profit												
Less-than-2-year	50	96.2	50	94.2	50	98.1	50	88.5	50	100.0	30	59.6
2-year	100	100.0	100	99.0	90	90.9	100	97.0	90	95.0	70	67.7
4-year	230	97.9	230	97.5	210	90.3	230	98.7	220	91.5	190	80.5

NOTE: All percentages are unweighted and based on the number of eligible institutions within the row under consideration. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

3.4.2 Evaluation of Student Records Collection Activities and Data Quality

Throughout student records data collection, NPSAS staff assisted sampled institutions with addressing any questions or issues with file uploads during student records data collection. Generally, institutions did not encounter significant obstacles when providing student records data, as demonstrated by the high proportion of participating institutions.

NPSAS staff reviewed student records data for data quality, including item-level completeness. Table 15 shows the completion rates of key student records data elements by data collection mode (web, CSV, or Excel). Variation in item-level response can generally be attributed to differences in institutions' data management systems, which vary in the information and level of detail they record, the format of the data available, and how easily data can be retrieved. However, the data elements listed in table 16 are available at a large majority of institutions, and the key student records data elements had a high level of completion overall.

Items with the highest completion rates were degree program (about 99 percent) and student class level (approximately 98 percent). For institutions, these are critical data elements for determining students' eligibility for financial aid, and these data are usually readily accessible from their data management systems. Of the key data elements listed in table 14, high school completion type had the lowest completion rate (about 72 percent). Some institutions reported that their campus data systems indicated whether the student had a high school credential, but did not indicate the type of credential received.

Table 15. Student records item-level completion rates, by data element and primary mode: 2015–16

	Total		Primary mode					
			Web mode		CSV upload		Excel upload	
Data element	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Total	**105,640**	**100.0**	**31,010**	**29.4**	**18,280**	**17.3**	**56,350**	**53.3**
Student characteristics								
Sex	103,470	97.9	29,010	93.5	18,170	99.4	56,290	99.9
Marital status	97,710	92.5	27,940	90.1	17,750	97.1	52,020	92.3
Citizenship	94,440	89.4	26,920	86.8	17,940	98.1	49,580	88.0
High school completion type	76,260	72.2	24,570	79.2	15,980	87.4	35,700	63.4
Race	83,160	78.7	20,430	65.9	15,330	83.9	47,390	84.1
Ethnicity	100,650	95.3	26,590	85.7	18,070	98.9	55,980	99.4
Enrollment								
Degree program	104,840	99.2	30,990	99.9	18,170	99.4	55,690	98.8
Student class level	103,410	97.9	30,680	98.9	18,200	99.6	54,530	96.8
Residency for tuition purposes	82,610	78.2	23,450	75.6	15,140	82.8	44,030	78.1
Total tuition and fees charged	96,210	91.1	26,640	85.9	17,530	95.9	52,040	92.4
Budget								
Tuition and fees	82,770	78.3	20,600	66.4	14,340	78.4	47,830	84.9
Financial aid								
Any aid received	80,310	76.0	27,670	89.2	17,320	94.7	35,320	62.7

NOTE: All percentages are unweighted and based on the number of eligible students within the row under consideration. Includes the total of all nonmissing responses, including responses of "Unknown." Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

3.5 Institution Data Collection Conclusions

NPSAS:16 project staff conducted institution recruiting and contacting for student enrollment list acquisition from March 4, 2015 through July 8, 2016. The overall response rate was 88 percent, a rate comparable to previous NPSAS cycles.

Of the 1,750 institutions with sampled students, 94 percent provided student records data. The high proportion of institutions providing student records data indicates that there were no major issues in complying with the institution data request.

Chapter 4. Student Interview Design, Data Collection, Outcomes, and Evaluation

Chapter 4 provides an overview of the interview design and associated systems for NPSAS:16 and describes the efforts to locate and contact sample members. The concluding sections include an evaluation of the student interview items.

4.1 Student Interview Design and Systems

The NPSAS:16 full-scale student interview, administered between February and November of 2016, included many items that had been part of previous NPSAS cycles to allow for trend comparisons among cohorts. Input from the study's expert TRP helped identify new data elements. For the complete list of TRP members, see appendix A. After conducting cognitive interviews and analyzing findings from the NPSAS:16 field test, NPSAS staff finalized new data elements and response options. For a summary of findings from the NPSAS:16 field test, see appendix O.

The data elements for NPSAS:16 included survey elements grouped by seven key content areas: Enrollment, Education Experiences, Financial Aid, Employment, Income and Expenses, Background, and Locating. The following are brief descriptions of these key content areas. For a complete list of data elements, see appendix F.

Enrollment items determined eligibility for the NPSAS study and identified members of the B&B cohort. The student interview collected extensive information on enrollment at the sampled institution (referred to hereafter as the NPSAS institution) during the 2015–16 academic year, including degree type, enrollment intensity, undergraduate or graduate year or level, and expected date of degree completion.

Education Experiences items gathered information on both high school and postsecondary experiences. For high school experiences, items included an estimate of GPA; patterns of high school coursetaking, such as Advanced Placement (AP) or International Baccalaureate (IB) courses; and taking of the SAT and ACT exams. For

postsecondary educational experiences, data were also collected on GPA, major, and online coursetaking at the NPSAS institution.

Financial Aid items collected information about sources of aid for the 2015–16 academic year, such as loans, grants or scholarships, employer assistance, and veterans benefits, etc. A historical focus of NPSAS studies has been student borrowing behavior, so the interview collected information on the amounts borrowed both for the 2015–16 academic year and cumulatively. Those students who did not apply for financial aid in the 2015–16 academic year received a question as to why they had not applied. Respondents with graduate-level assistantships, fellowships, or traineeships received items about pay and further details related to these school jobs.

Employment items captured information about all paid employers between July 2015 and June 2016, earnings, and average hours worked per week. Because K–12 teaching has been an ongoing focus in the NPSAS family of studies, respondents received questions on whether they had any experience teaching kindergarten through 12th grade (K–12), planned to become a K–12 teacher, or had prepared for a teaching career. B&B-eligible respondents received questions related to future employment, including work plans for the upcoming year and perceptions about their future occupation and wages.

The ***Income and Expenses*** section of the NPSAS:16 interview collected information such as respondents' annual income; spouse's annual income; number of children and other dependents; and monthly costs of dependents, including child care and dependent college expenses. Additional items included monthly rent or mortgage amount, homelessness and dependency status, and whether the student received untaxed benefits during the 2015–16 academic year. Information collected about the families of dependent respondents included parents' marital status and income.

Background items obtained information about student demographic characteristics, such as date of birth, marital status, sex, U.S. citizenship, immigration status, and race and ethnicity. The background section also included items on family members of respondents, including spouses' enrollment in postsecondary education in the 2015–16 academic year and siblings' college attendance.

The ***Locating*** section collected contacting information for follow-up studies.

Figure 3 below depicts the key content areas and principal topics in the full-scale survey. For the complete NPSAS:16 student interview, see appendix G.

Figure 3. NPSAS:16 full-scale survey by content area and topics: 2015–16

[1] Only B&B-eligible respondents received these items.
NOTE: NPSAS = National Postsecondary Student Aid Study. B&B = Baccalaureate and Beyond Longitudinal Study. GPA = Grade point average.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

4.1.1 Survey Mode Administration

For the NPSAS:16 full-scale study, a single survey instrument was administered in two user modes: web (nonmobile and mobile) and telephone. The web survey was designed with device-optimized pages based on HTML5 web standards that ensured that instrument formatting would scale properly for all screen sizes. Screen-size optimization allowed respondents to complete the survey on a desktop or laptop computer while providing a mobile-friendly design for respondents who wished to complete the survey on a tablet or smartphone.

For the telephone survey, an interviewer accessed the web instrument through a computer-assisted telephone interviewing case management system (CATI-CMS), which assigned cases and provided the appropriate screens and scripts for the interviewer to use during the survey. (For more information on how NPSAS staff used the case management system, see section 4.2.) On-screen instructions provided telephone interviewers with guidance on administering each question (e.g., whether the interviewer should read response options aloud, when to probe). To minimize mode effects, NPSAS project staff incorporated the following features into the interview to provide web respondents with assistance similar to that provided by a trained telephone interviewer:

- a help text button on every form (or web screen) to define key terms and clarify question intent;
- prompts to correct out of range or incorrectly formatted responses;
- conversion text to encourage responses to unanswered critical items; and
- prompts to encourage response if a sample member left three consecutive questions unanswered.

4.1.2 Coding Systems

Assisted coding systems (coders) programmed within the NPSAS survey standardized the collection and coding of several pieces of information. NPSAS staffed designed coders to simplify data entry for four survey items with potentially complex strings for answers: postsecondary institutions attended during the 2015–16 academic year, the respondent's last high school, major or field of study at the NPSAS institution, and respondents' intended future occupation. The respondent (or telephone interviewer) entered text strings into a coder, which launched a keyword search of an underlying database and returned a list of possible matches for

selection. The following are descriptions of the individual coding systems and sources:

- The ***Postsecondary Institution coder*** linked to the complete set of postsecondary institutions contained in the Integrated Postsecondary Education Data System (IPEDS), https://nces.ed.gov/ipeds/, developed by the National Center for Education Statistics (NCES). For this coder, data from prior years supplemented data from the 2012–13 Institution Characteristics Header file. This coder covered any postsecondary institutions the respondent attended, other than the NPSAS institution, during the 2015–16 academic year. For any institutions not listed in the database, follow-up questions asked respondents to provide the control (e.g., public or private) and level (e.g., 4-year or 2-year) of the institution. The coder retained any initially entered text strings that yielded no IPEDS matches.

- The ***High School coder*** database contained data from the 2013–14 administration of the Private School Universe Survey (https://nces.ed.gov/surveys/pss/) and the preliminary 2013–14 school year file of the Common Core of Data for public schools (http://nces.ed.gov/ccd/). The database was supplemented with PSS and CCD data from prior years. For schools not identified within the high school coder, the coder retained the entered text string and asked respondents to supply the school control, district or county name, and the highest and lowest grade levels taught at the school. Students who identified as home schooled or as having last attended a foreign high school were not administered the high school coder form.

- The ***Major coder*** used the 2010 Classification of Instructional Programs (CIP) taxonomy, also developed by NCES (https://nces.ed.gov/ipeds/cipcode). For any majors or fields of study not found in the CIP database, respondents selected a general major area and a specific discipline.

- The ***Occupation coder*** linked to the 2014 Occupational Information Network Online (O*NET OnLine) database (https://onetonline.org). For any occupations not listed in the database, the respondents provided a general occupational area, specific occupational area, and a detailed classification area for the occupation.

4.1.3 Survey Design System

NPSAS staff created the NPSAS:16 survey instrument using a proprietary web-based system in which staff developed the instrument for review, testing, and subsequent modifications. Staff stored all information relating to the instrument in a structured query language (SQL) server database made accessible through a web interface.

4.2 Student Interview Data Collection

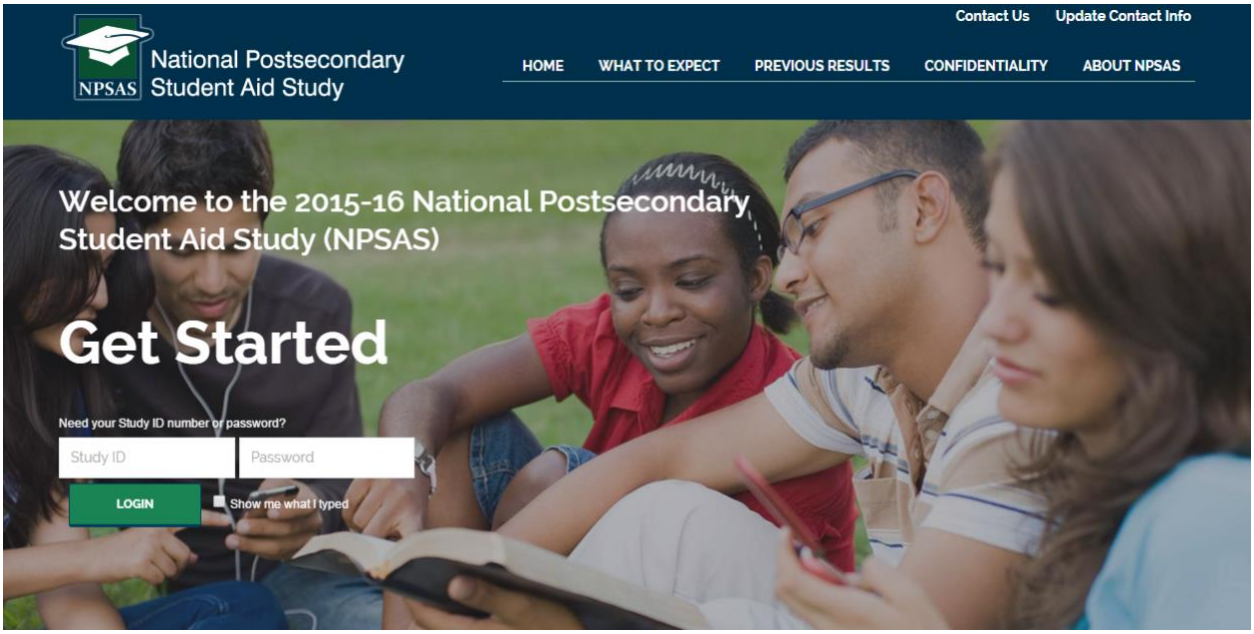
The primary mode for NPSAS:16 student data collection was by web survey available through the study website. Sample members also had the option of completing the survey with an interviewer trained in computer-assisted telephone interviewing (CATI) methods. A help desk was available to provide additional information and support to sample members.

4.2.1 Study Website

Communications sent to sample members included a link to the home page for the NPSAS:16 study website (figure 4), where they could log in to update contact information and complete the student interview. Other navigation options included links to information on the study, data use, confidentiality assurances, and selected findings from previous studies. The "Contact Us" page provided contact information for the study help desk and project staff at RTI, as well as links to the main NCES and RTI websites.

The NPSAS:16 study website, following NCES web policies, used a three-tier security approach to ensure the security of all collected data. The first tier of security included a secure login/password combination provided to sample members before the start of data collection. The second tier of security encrypted all entered data with Secure Sockets Layer (SSL) technology. The third tier of security stored all collected data in a secured database housed on a machine physically separate from the web server.

Figure 4. Home page for NPSAS:16 study website: 2015–16



SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

4.2.2 Training of Interview Data Collection Staff

Members of the NPSAS:16 interview data collection team held one of four roles: telephone interviewer, quality expert (QE), quality control supervisor (QCS), or intensive tracing staff which are further described following table 16. All data collection staff, regardless of role, completed a comprehensive training program before beginning work on the study (see appendix I for training agendas). Additionally, all data collection staff completed a general training program on topics such as proper interviewing techniques, confidentiality procedures, and sample member rights. Table 16 shows the training schedule of the interview data collection staff.

Table 16. Training of data collection staff: 2015–16

Staff trained	Time period	Number of staff trained
Telephone interviewers, QCSs, and QEs	January 26–28, March 8–10, April 26–28, May 10–12, May 24–26, July 19–21, August 2–4, August 23–25	194
Intensive tracing staff	March 23, April 14, May 11, June 16, July 7	40

NOTE: QCS = quality control supervisor. QE = quality expert.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

Telephone interviewers. Telephone interviewers acted as the primary point of contact for sample members, conducting telephone interviews and employing strategies to avert or convert refusals. Telephone interviewers also served as help desk agents to respond to sample member concerns, reset passwords when needed, and address incentive receipt inquiries and issues. Telephone interviewers familiarized themselves with the survey instrument and received training specific to each interview question. They developed proficiency with the interview through mock interviews, hands-on practice with case management systems, and instruction on conversational interviewing techniques. Training materials included a telephone interviewer manual and associated materials addressing survey administration and conversational interviewing. Project staff certified telephone interviewers after they conducted a mock interview and provided appropriate and accurate responses to NPSAS:16 frequently asked questions. Weekly quality circle meetings of QEs and telephone interviewers were held to review proper administration of the survey and ad hoc topics related to NPSAS:16 or general interview protocol. Project staff asked trainees for feedback in identifying training needs or topics for future quality circle meetings.

QEs. QEs supervised telephone interviewers, performing day-to-day monitoring responsibilities and providing constructive feedback and coaching to interviewers after monitoring live or recorded NPSAS interviews. QEs attended interviewer training to learn survey basics and interviewing conventions. In addition, they assumed general monitoring responsibilities and were provided with an interviewing manual and a file compilation of screens and text in CATI and the web interview, including help text.

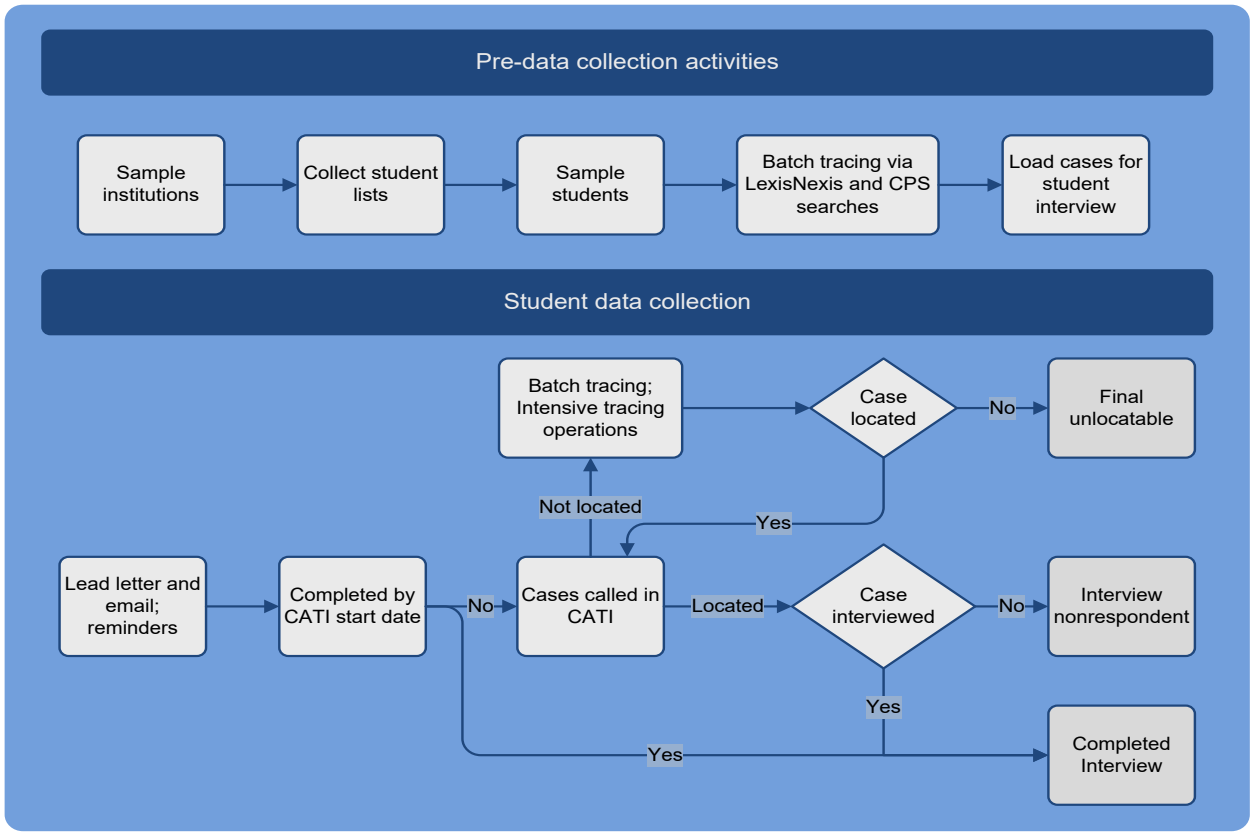
QCSs. QCSs supervised all staff, coordinating monitoring of telephone-interviewer performance and production, providing guidance to interviewers, and troubleshooting problems as they arose. The QCSs also attended telephone-interviewer trainings so they would be familiar with the interview and all aspects of its administration, allowing them to better identify any areas that needed improvement.

Intensive tracing staff. Intensive tracing staff completed a 16-hour program on tracing procedures with an additional 2 hours of project-specific training, including the tracing techniques most appropriate for locating NPSAS:16 sample members. Tracing staff received additional training on refusal aversion techniques and case review.

4.2.3 Locating and Contacting Sample Members

Before the start of data collection, several batch locating databases were used to update (or confirm) sample member contact information received during collection of institution enrollment lists. At the start of data collection, staff sent a mailing and an e-mail to sample members. Once outbound telephone efforts began, specially trained intensive tracing staff conducted additional batch tracing and intensive tracing for sample members who could not be located by telephone. Once sample members were located, interviewers contacted and invited them to complete the interview. See figure 5 for a diagram of locating activities.

Figure 5. NPSAS:16 sample member locating activities



NOTE: CPS = Central Processing System. CATI = computer-assisted telephone interviewing.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

After institutions were sampled and student lists were collected, the following steps were taken before data collection began, with respect to tracing:

Step 1: In batch tracing, tracing staff sent cases with at least one valid address to LexisNexis (formerly FirstData) to access the U.S. Postal Service (USPS)

National Change of Address database (NCOA) for matching. Survey staff updated records with new or updated address information based on the match.

Step 2: Before the first mailing, staff sent cases that still had no good mailing address after NCOA matching to LexisNexis's Single Best Address search. While NCOA only provides information for people who registered a change of address with the USPS, Single Best Address can provide new addresses, including those not registered with NCOA. Single Best Address uses a name and Social Security number (SSN) to search multiple data sources, using progressive search logic to return the most current address available.

Step 3: Because NCOA and Single Best Address only provide address information, staff sent sample member information to LexisNexis's PhoneAppend telephone number lookup service. LexisNexis carries approximately 718.8 million current and historical phone numbers, of which 80 percent are likely cell phones, 15 percent are residential landlines, and 5 percent are business landlines. PhoneAppend returns a single telephone number based on a search by name, street address, and ZIP code.

Step 4: In addition to the LexisNexis searches, staff sent cases with a valid SSN to CPS for record matching. CPS contains information on students who have applied for financial aid using FAFSA. NPSAS staff then compared records obtained from CPS to existing contact data, updating locating information when necessary.

Data collection mailings and e-mails. Using the addresses updated in batch tracing, staff sent mailings to all addresses identified for sample members. Mailings proceeded on a flow basis as institutions provided sample member information and as batch tracing procedures provided additional contact information. All mail correspondence was via USPS mail and contained a lead letter and study brochure. The lead letter notified sample members of the start of data collection and the incentive they were eligible to receive for completing the survey. The letter also included unique login information for the web survey instrument and encouraged participation during the early response period. The brochure provided information about the purpose of the study, confidentiality and security concerns, and study contact information. Staff sent additional mailings such as postcards, letters, and flyers periodically as reminders to complete the study.

Staff sent mail communications to all sample member e-mail addresses collected from institutions and updated via batch tracing procedures. E-mails also went out on a flow basis and provided sample members with a link to complete the survey, as

well as unique login information. See appendix H for examples of the mailing and e-mail contact materials sent to sample members.

CATI locating. Telephone interviewers attempted to conduct a telephone interview with any sample members who did not complete a web interview. Interviewers called the number with the best likelihood of reaching the sample member, as determined by the automated calling system. If the sample member could not be reached at the listed number, the interviewer attempted to gather locating information from the contact who answered the call. If it was not possible to obtain updated contact information by phone, the interviewer used all other information available about the sample member and other contacts to locate the sample member. When the interviewer had exhausted all other sources of tracing data, intensive tracing was initiated.

Intensive tracing. Staff relegated cases that could not be located by other methods to intensive tracing. Intensive tracing cases included those with no telephone number to load into CATI or for which all known numbers had failed. Intensive tracing was a two-stage process, utilizing both public-domain and proprietary databases.

The first stage of intensive tracing identified sample members in consumer databases (e.g., LexisNexis, Experian, and Accurint) by SSN. If staff found a new telephone lead, they sent the case back to the telephone interview queue for follow-up by telephone interviewers. If the search resulted in a new address only, directory assistance searches were conducted to locate a telephone number for the contact. This approach minimized the effort required to locate cases and the time that cases were unavailable for data collection efforts.

Cases that could not be located during initial tracing efforts went to the second tier of intensive tracing. Tracing staff conducted a manual review of each case and determined the appropriate next steps based on the leads developed from prior tracing and contacting activities. On a case-by-case basis, tracing staff performed the following activities to obtain current contact information:

- Accurint database searches for sample members, parents, and other contacts;
- LexisNexis database searches including FastData reverse phone, SSN search, address search, and name search;
- Experian social search;
- running matches with public records (e.g., driver's license searches through state departments of motor vehicles);

- searching institution websites for campus/alumni directories and class or personal web pages; and
- other ad hoc methods, such as calling individuals with the same unusual surname in small towns or rural areas to see if they were relations of or knew the sample member.

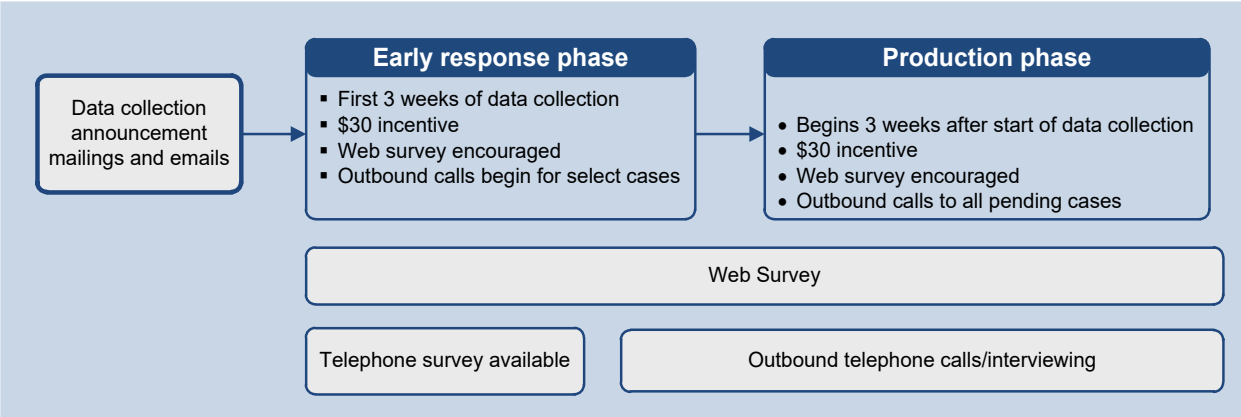
If, after exhausting all possible leads for locating sample members, staff set these cases to a status of "unlocatable." If the sample member did not complete an interview by the end of data collection, that sample member was classified as a nonrespondent.

4.2.4 Interviewing

Data collection for the NPSAS:16 interview consisted of early-response and production phases (figure 6). Regardless of when they chose to respond, sample members could access both the web and telephone versions of the survey. (For the discussion that follows, respondents will be classified as either web or telephone respondents, with the exception of the results presented in table 25, and the discussion there, where results are further broken down to include not just web respondents, but a subset of that group, the mobile completers. There are only two modes of completion, however: web and telephone.)

The early-response phase began in February 2016 with communications to sample members encouraging completion of the web survey. Respondents could then opt to call the help desk at any time to complete the interview over the telephone, but project staff limited outbound telephone contacts during this phase. The 3-week early-response phase began in waves, based on when institutions sent sample member information and staff completed batch tracing procedures. During the remainder of data collection (the production phase), interviewers called sample members to encourage survey completion by web or telephone. Project staff also sent multiple reminder mailings and e-mails throughout the data collection period to encourage sample members to participate. Table 17 shows the timing for outbound telephone contacting of sample members.

Figure 6. Data collection phases: 2015–16



SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

Table 17. Beginning of outbound telephone calls, by control and level of institution and response phase: 2015–16

Control and level of institution	Beginning of outbound telephone calls to sample members
Public	
Less-than-2-year	Early-response phase, 15 days after initial contact mailing
2-year	Early-response phase, 15 days after initial contact mailing
4-year, non-doctorate-granting, primarily subbaccalaureate	Production phase, 3 weeks + 1 day after initial contact mailing
4-year, non-doctorate-granting, primarily baccalaureate	Production phase, 3 weeks + 1 day after initial contact mailing
4-year, doctorate-granting	Production phase, 3 weeks + 1 day after initial contact mailing
Private nonprofit	
2-year-or-less	Early-response phase, 15 days after initial contact mailing
4-year, non-doctorate-granting	Production phase, 3 weeks + 1 day after initial contact mailing
4-year, doctorate-granting	Production phase, 3 weeks + 1 day after initial contact mailing
Private for-profit	
Less-than-2-year	Early-response phase, 11 days after initial contact mailing
2-year	Early-response phase, 11 days after initial contact mailing
4-year	Production phase, 3 weeks + 1 day after initial contact mailing

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

Once outbound calling began, interviewers started locating sample members, gaining their cooperation, answering any possible questions about the study, and conducting interviews. Upon successfully reaching sample members, interviewers encouraged them to complete the interview immediately via CATI. Alternatively, an interviewer could e-mail secure credentials for the web interview to sample members who preferred to complete the survey later. Interviewers followed up with sample members by telephone 5 days after contact if the sample member had selected the web option but not completed the survey.

Next, automated call scheduling assigned cases to interviewers by type and priority, best day and time of day to call; priority was given to scheduled appointments. The scheduler organized cases into queues based on factors such as prior contact status (e.g., cases that were recently contacted or had never been contacted), refusal status, and appointments set during a prior contact attempt. The scheduler also automatically ordered numbers to call by prioritizing lines most likely to result in contacting the sample member. Staff added new numbers continuously, based on contact attempts; batch and intensive tracing efforts; and updates received through mailings, e-mails, and help desk call-ins. The call scheduler reprioritized telephone numbers based on new information as it became available.

Late in data collection, two abbreviated versions of the interview were made available to selected sample members as part of responsive design efforts. For further detail on responsive design, see chapter 7. The total number of eligible sample members offered one of the abbreviated interviews was about 47,670. Of those offered, almost 24 percent ultimately completed an abbreviated interview. The abbreviated interviews included fewer questions and therefore required less time to complete—approximately 15 minutes and 10 minutes, respectively. The abbreviated interview questions focused on the key data that could classify a sample member as a study member, as described in section 4.4.7. As of the end of data collection, approximately 15 percent of the 78,860 NPSAS:16 interview respondents had completed an abbreviated interview (table 18).

Table 18. Abbreviated interview offer, by control and level of institution and student type: 2015–16

Control and level of institution and student type	Eligible sample	Offered abbreviated interview		Abbreviated interview respondent	
		Number	Percent of eligible sample	Number	Percent of offered abbreviated interview
Total	**119,550**	**47,670**	**39.9**	**11,600**	**24.3**
Control of institution					
Public	56,850	21,460	37.7	4,930	23.0
Private nonprofit	25,170	7,530	29.9	1,800	23.9
Private for-profit	37,530	18,680	49.8	4,870	26.1
Level of institution					
Less-than-2-year	3,050	1,810	59.3	330	18.3
2-year	24,520	11,520	47.0	2,290	19.9
4-year, non-doctorate-granting	42,730	17,670	41.3	4,690	26.5
4-year, doctorate-granting	49,260	16,680	33.9	4,290	25.7
Control and level of institution					
Public					
Less-than-2-year	370	180	47.6	40	21.0
2-year	17,360	7,380	42.5	1,430	19.4
4-year, non-doctorate-granting, primarily subbaccalaureate	5,610	2,320	41.3	510	22.1
4-year, non-doctorate-granting, primarily baccalaureate	6,950	2,510	36.2	660	26.3
4-year, doctorate-granting	26,570	9,070	34.1	2,280	25.2
Private nonprofit					
Less-than-4-year	960	480	50.4	90	18.7
4-year, non-doctorate-granting	11,140	3,450	31.0	770	22.4
4-year, doctorate-granting	13,910	4,100	29.5	1,040	25.4
Private for-profit					
Less-than-2-year	2,520	1,520	60.2	270	18.1
2-year	6,360	3,770	59.2	790	20.8
4-year	27,810	12,900	46.4	3,710	28.8
Student type					
Total undergraduate	95,020	40,480	42.6	9,790	24.2
Potential B&B	33,760	13,080	38.7	3,240	24.8
Other undergraduate	62,520	27,940	44.7	6,640	23.8
Graduate	23,280	6,650	28.6	1,720	25.8

NOTE: B&B = Baccalaureate and Beyond Longitudinal Study. For Student type, the 1,260 students who are classified as both Potential B&B and Graduate are all included in the Potential B&B count and excluded from the Graduate count. The Total undergraduate count excludes the 1,260 Potential B&B students who are also classified as Graduate. Respondent count includes eligible students who met the criteria for qualification as a student interview respondent, which required completing at least a partial interview. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

4.3 Data Collection Quality Control

Quality control procedures used in student interview data collection included frequent monitoring of recorded interviews, a help desk to answer questions about

the study or assist sample members with completion of the web interview, quality circle meetings to facilitate communication among staff members, and debriefing meetings to identify areas for potential improvement.

4.3.1 Interview Monitoring

QEs regularly monitored telephone interviews during NPSAS:16 data collection to meet a series of data-quality objectives:

- identification of items in the interview that posed problems for interviewers and/or respondents;
- reduction in the number of interviewer errors;
- improvement in interviewer performance through reinforcement of effective strategies; and
- assessment of data quality.

QEs recorded feedback on standardized monitoring forms, evaluating interviewers on their professionalism, question administration, conversational interviewing, and familiarity with the instrument. Interviewers received regular feedback from monitoring sessions, and quality circle meetings frequently incorporated issues identified during monitoring to improve the overall quality of telephone interviews. Supervisory staff used segments of recorded interviews as training aids during project trainings and meetings.

4.3.2 Help Desk

In addition to the study information available on the study website, NPSAS:16 staff implemented a help desk to respond to sample members on matters ranging from general inquiries, to interview completion assistance, to incentive status updates. Staff confirmed contact information for the sample member for each call received, recording a description of the problem and the resolution for future reference. If technical difficulties prevented sample members from completing the web interview, rather than suggesting a reattempt of the web interview, help desk staff connected the callers with telephone interviewers to continue the survey via a telephone interview. Two common types of help desk incidents were requests to retrieve login credentials and requests to complete the interview over the telephone. For the convenience of sample members, a "Forgot Password?" feature on the study website enabled automated retrieval of credentials, conditional upon answering requisite security questions.

4.3.3 Quality Circle Meetings

As part of supervisory responsibilities, QEs met with telephone interviewers for regular quality circle meetings designed to facilitate communication between project staff. Frequently covered topics included:

- clarification of questions and item responses from the survey instrument;
- reinforcement of successful interviewing and refusal conversion techniques;
- guidelines for providing detailed case comments;
- strategies for gaining cooperation from sample members and other contacts;
- data security protocols; and
- study progress.

Project staff summarized meeting discussions and provided summaries to interviewers for review and to serve as an ongoing resource for the duration of data collection.

4.3.4 Debriefing

After NPSAS:16 data collection ended, project staff met with telephone interviewers and call-center supervisory staff to learn more about their experiences during the study. The debriefing meetings were designed to encourage reflection on the completed data collection and consider improvements for successive studies. The following is a summary of lessons learned from the debriefing sessions.

Telephone interviewers appreciated hands-on training in practicing interviews. In response to feedback from prior studies, NPSAS:16 training had included more of this type of activity for interviewers to gain experience with the case management and contacting systems and the survey instrument. Interviewers found that reviewing refusal aversion strategies and frequently asked questions helped them develop strategies to gain cooperation from reluctant sample members and "gatekeepers." Gatekeepers are parents or other contacts who answered telephone call attempts to sample members. The interviewers also emphasized the importance of training geared toward handling hostile sample members.

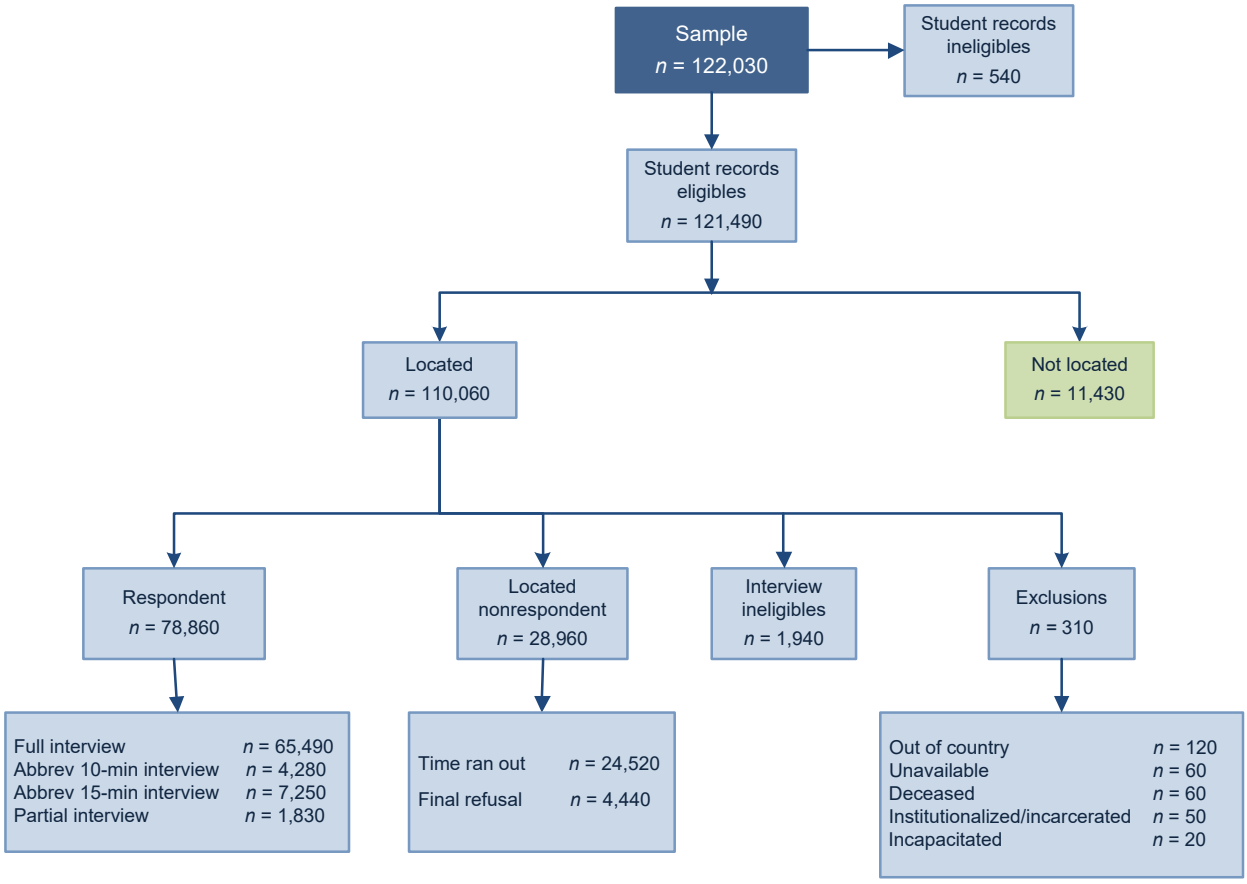
In addition, telephone interviewers reported that the resources provided in the interview, such as help text and conversion text, helped them administer the interview successfully. They also felt that recorded interviews used during

monitoring feedback sessions and quality circle meetings helped to improve their interviewing techniques. Finally, interviewers provided feedback to improve the flow of the abbreviated interview in future studies. Project staff prepared a summary of the debriefing meetings for consideration when planning future studies.

4.4 Student Interview Data Collection Outcomes

To assess student interview data collection outcomes, NPSAS staff reviewed the number of sample members located and interviewed, the time required to complete the interview, the time spent contacting sample members, and the rate of conversion for interview refusals. As indicated in chapter 2, students had to meet certain criteria to be eligible for NPSAS (enrolled in the NPSAS year, enrolled in a Title IV eligible program, not concurrently enrolled in high school, etc.). NPSAS staff asked the institutions to provide only eligible students on enrollment lists, but occasionally, ineligible students were sampled. Upon closer examination, and in cooperation with the sampled institution, those students identified as ineligible during the student record collection process were removed from the denominator for calculating student record response rates. Overall, NPSAS staff located approximately 90 percent (n = 110,060) of NPSAS:16 student records eligible sample members. Of the sample members located, approximately 72 percent (n = 78,860) responded. Of the 119,550 total eligible sample members, approximately 66 percent responded. See figure 7 below for overall locating and interviewing results.

Figure 7. NPSAS:16 overall locating and interviewing results: 2015–16



NOTE: Located case total includes an additional 350 cases later found to be student record ineligibles. Sample sizes rounded to the nearest 10. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

4.4.1 Student Locating Results

Locate rates by control and level of institution, shown in table 19, ranged from a high of approximately 93 percent for students enrolled at private nonprofit, 4-year, doctorate-granting institutions to a low of about 83 percent for students enrolled at private for-profit, less-than-2-year institutions. Data collection staff located potential B&B students at a higher rate than nonpotential B&B undergraduate students (χ^2 (1, n = 98,388) = 179.04, p <.001). They located graduate students more often than undergraduate students overall (χ^2 (1, n = 122,030) = 290.05, p <.001).

Table 19. Student locating results, by control and level of institution and student type: 2015–16

Control and level of institution and student type	Total sample	Located	
		Number	Percent of total sample
Total	**122,030**	**110,410**	**90.5**
Control of institution			
Public	58,370	53,130	91.0
Private nonprofit	25,510	23,760	93.1
Private for-profit	38,150	33,530	87.9
Level of institution			
Less-than-2-year	3,170	2,620	82.7
2-year	25,570	22,430	87.7
4-year, non-doctorate-granting	43,500	39,350	90.5
4-year, doctorate-granting	49,790	46,010	92.4
Control and level of institution			
Public			
Less-than-2-year	400	340	85.9
2-year	18,210	16,250	89.2
4-year, non-doctorate-granting, primarily subbaccalaureate	5,850	5,210	89.1
4-year, non-doctorate-granting, primarily baccalaureate	7,090	6,540	92.2
4-year, doctorate-granting	26,830	24,800	92.4
Private nonprofit			
Less-than-4-year	990	850	85.3
4-year, non-doctorate-granting	11,300	10,480	92.7
4-year, doctorate-granting	14,080	13,140	93.3
Private for-profit			
Less-than-2-year	2,610	2,160	82.8
2-year	6,540	5,460	83.5
4-year	28,140	25,200	89.6
Student type[1]			
Total undergraduate	97,110	87,160	89.8
Potential B&B student	34,130	31,230	91.5
Other undergraduate	64,260	57,070	88.8
Graduate	24,920	23,250	93.3

[1] As potential B&B students can also be graduate students, the listed subtotals are not mutually exclusive. In NPSAS:16, 1,280 potential B&B students are also classified as graduate students.
NOTE: B&B = Baccalaureate and Beyond Longitudinal Study. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

Batch tracing. Matching the NPSAS sample with the CPS database, which provides information on students who have applied for federal financial aid using the FAFSA, resulted in updated or confirmed contact information for about 70 percent of the cases submitted for batch tracing. NPSAS staff then submitted all existing and updated contact information received from CPS to the NCOA database. Of the

120,040 cases sent to NCOA, NCOA returned an address for 30,350 (about 25 percent) (table 20).

As part of the NCOA batch tracing step, NPSAS staff submitted sample member information to PhoneAppend for telephone number updates. Of the 120,040 cases submitted, PhoneAppend returned 60,610 (about 51 percent) with new or confirmed telephone numbers. Before intensive tracing, NPSAS staff submitted a small group of cases to Premium Phone after exhausting all other leads. Of the 9,420 cases submitted, Premium Phone returned 4,610 (about 49 percent) with new or confirmed telephone numbers.

Table 20. Batch processing record match rates, by method of tracing: 2015–16

Method of tracing	Number of records sent	Number of records matched	Percent matched
CPS	111,910	77,810	69.5
NCOA	120,040	30,350	25.3
PhoneAppend	120,040	60,610	50.5
Premium Phone	9,420	4,610	48.9
Single Best Address	3,370	3,110	92.2
Single Best Phone	3,410	2,180	64.0
NSLDS	122,030	78,120	64.0

NOTE: CPS = Central Processing System. NCOA = National Change of Address. NSLDS = National Student Loan Data System. Percentage is based on the number of records sent for batch tracing. Because records were sent to multiple tracing sources, multiple record matches were possible. Match rate includes instances when sample member contact information was confirmed and when new information was provided. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

Intensive tracing. Staff initiated intensive tracing for those sample members who were not located in batch tracing or initial locating. Overall, 9,570 cases, or approximately 8 percent of the total sample, required intensive tracing (table 21). By type of institution, the rate requiring intensive tracing ranged from a high of about 19 percent of students at public, less-than-2-year institutions to roughly 5 percent of students at private for-profit, 4-year institutions.

Of the 9,570 cases requiring intensive tracing, about 90 percent were successfully located. Of the total located, 2,690, about 31 percent of those located, completed interviews (table 22).

Table 21. Cases requiring intensive tracing, by institution characteristics and student type: 2015–16

Control and level of institution and student type	Total sample	Cases requiring intensive tracing	
		Number	Percent of total sample
Total	**122,040**	**9,570**	**7.8**
Control of institution			
Public	58,370	4,990	8.6
Private nonprofit	25,510	2,230	8.7
Private for-profit	38,150	2,350	6.2
Level of institution			
Less-than-2-year	3,170	360	11.3
2-year	25,570	2,470	9.7
4-year, non-doctorate-granting	43,500	2,920	6.7
4-year, doctorate-granting	49,790	3,820	7.7
Control and level of institution			
Public			
Less-than-2-year	400	80	18.9
2-year	18,210	1,900	10.5
4-year, non-doctorate-granting, primarily subbaccalaureate	5,850	550	9.4
4-year, non-doctorate-granting, primarily baccalaureate	7,090	460	6.5
4-year, doctorate-granting	26,830	2,010	7.5
Private nonprofit			
Less-than-4-year	990	60	6.3
4-year, non-doctorate-granting	11,300	820	7.3
4-year, doctorate-granting	14,080	1,400	9.9
Private for-profit			
Less-than-2-year	2,610	270	10.5
2-year	6,540	520	8.0
4-year	28,140	1,500	5.3
Student type			
Total undergraduate	97,110	7,410	7.6
Potential B&B student	34,130	2,050	6.0
Other undergraduate	64,260	5,460	8.5
Graduate	24,920	2,170	8.7

NOTE: B&B = Baccalaureate and Beyond Longitudinal Study. The counts for cases requiring intensive tracing exclude cases initiated to intensive tracing that were not traced but include cases for which intensive tracing work began but work was stopped. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

Table 22. Located and interviewed rates of cases requiring intensive tracing procedures: 2015–16

Intensive tracing round	Total cases	Located in tracing operations		Interviewed	
		Number	Percent of total cases	Number	Percent of located in tracing operations
Total	**9,570**	**8,620**	**90.1**	**2,690**	**31.2**
Tracing operations—stage 1	9,570	8,200	85.6	2,600	31.7
Tracing operations—stage 2	2,110	1,520	72.2	260	16.9

NOTE: Total cases count excludes cases initiated to intensive tracing that were not traced. Tracing operations—stage 2 cases are a subset of tracing operations—stage 1 cases that required additional intensive tracing efforts; therefore, total cases are not the sum of the two totals. Interviewed count includes eligible students who met the criteria for qualification as an interview respondent, which required completing at least a partial interview. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

4.4.2 Student Interview Response Rates

Some 78,860 students, approximately 66 percent of the eligible sample of 119,550, completed the NPSAS:16 interview (table 23). Across institution level and control, response rates ranged from about 73 percent for students at private nonprofit, 4-year, doctorate-granting institutions to roughly 48 percent for students at private for-profit, less-than 2-year institutions. Potential B&B students were more likely to respond than nonpotential B&B undergraduates (67 percent compared with 63 percent) ($\chi 2$ (1, n = 96,272) = 175.2, $p < .0001$). Graduate students (about 73 percent) responded at a higher rate than undergraduate students (roughly 64 percent) ($\chi 2$ (1, n = 119,553) = 769.27, $p < .0001$).

Table 23. Student interview completion rates, by control and level of institution and student type: 2015–16

Control and level of institution and student type	Eligible sample	Total respondents	
		Number	Percent of eligible sample
Total	**119,550**	**78,860**	**66.0**
Control of institution			
Public	56,850	37,710	66.3
Private nonprofit	25,170	18,260	72.6
Private for-profit	37,530	22,890	61.0
Level of institution			
Less-than-2-year	3,050	1,470	48.0
2-year	24,520	14,160	57.7
4-year, non-doctorate-granting	42,730	28,380	66.4
4-year, doctorate-granting	49,260	34,860	70.8
Control and level of institution			
Public			
Less-than-2-year	370	210	55.7
2-year	17,360	10,430	60.1
4-year, non-doctorate-granting, primarily subbaccalaureate	5,610	3,550	63.3
4-year, non-doctorate-granting, primarily baccalaureate	6,950	4,840	69.6
4-year, doctorate-granting	26,570	18,680	70.3
Private nonprofit			
Less-than-4-year	960	540	55.9
4-year, non-doctorate-granting	11,140	7,970	71.5
4-year, doctorate-granting	13,910	10,190	73.2
Private for-profit			
Less-than-2-year	2,520	1,200	47.6
2-year	6,360	3,250	51.1
4-year	27,810	18,010	64.8
Student type			
Total undergraduate	95,020	60,840	64.0
Potential B&B student	33,760	22,540	66.8
Other undergraduate	62,520	39,060	62.5
Graduate	24,530	18,020	73.4

NOTE: Respondent count includes eligible students who met the criteria for qualification as a student interview respondent, which required completing at least a partial interview. Excludes 2,480 cases determined to be ineligible for the study using data obtained from one or more sources. B&B = Baccalaureate and Beyond Longitudinal Study. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

Completion by phase and mode. As described in section 4.2.4, the NPSAS:16 student interview occurred in two phases, an early-response phase and a production phase, and in two modes, by web and by telephone. Of the 78,860 cases that completed the interview, about 53 percent (41,470 cases) completed during the early-

response phase, and about 47 percent (37,390 cases) completed in the production phase (table 24).

Table 24. Student interview completion rates, by data collection phase, control and level of institution, and student type: 2015–16

Control and level of institution and student type	Eligible sample	Total respondents		Data collection phase			
				Early response		Production	
		Number	Percent of eligible	Number	Percent of respondents	Number	Percent of respondents
Total	**119,550**	**78,860**	**66.0**	**41,470**	**52.6**	**37,390**	**47.4**
Control of institution							
Public	56,850	37,710	66.3	19,790	52.5	17,910	47.5
Private nonprofit	25,170	18,260	72.6	10,110	55.3	8,160	44.7
Private for-profit	37,530	22,890	61.0	11,570	50.5	11,320	49.5
Level of institution							
Less-than-2-year	3,050	1,470	48.0	600	40.8	870	59.2
2-year	24,520	14,160	57.7	7,130	50.3	7,030	49.7
4-year, non-doctorate-granting	42,730	28,380	66.4	14,720	51.9	13,660	48.1
4-year, doctorate-granting	49,260	34,860	70.8	19,030	54.6	15,830	45.4
Control and level of institution							
Public							
Less-than-2-year	370	210	55.7	90	42.2	120	57.8
2-year	17,360	10,430	60.1	5,370	51.5	5,060	48.5
4-year, non-doctorate-granting, primarily subbaccalaureate	5,610	3,550	63.3	1,870	52.6	1,680	47.4
4-year, non-doctorate-granting, primarily baccalaureate	6,950	4,840	69.6	2,500	51.8	2,330	48.2
4-year, doctorate-granting	26,570	18,680	70.3	9,970	53.4	8,720	46.6
Private nonprofit							
Less-than-4-year	960	540	55.9	270	50.3	270	49.7
4-year, non-doctorate-granting	11,140	7,970	71.5	4,380	54.9	3,590	45.1
4-year, doctorate-granting	13,910	10,190	73.2	5,670	55.6	4,520	44.4
Private for-profit							
Less-than-2-year	2,520	1,200	47.6	490	40.9	710	59.1
2-year	6,360	3,250	51.1	1,510	46.4	1,740	53.6
4-year	27,810	18,010	64.8	9,360	52.0	8,650	48.0
Student type							
Total undergraduate	95,020	60,840	64.0	31,350	51.5	29,490	48.5
Potential B&B student	33,760	22,540	66.8	11,980	53.2	10,560	46.8
Other undergraduate	62,520	39,060	62.5	19,790	50.7	19,270	49.3
Graduate	24,530	18,020	73.4	10,120	56.2	7,900	43.8

NOTE: Respondent count includes eligible students who met the criteria for qualification as a student interview respondent, which required completing at least a partial interview. Excludes 2,480 cases determined to be ineligible for the study using data obtained from one or more sources. B&B = Baccalaureate and Beyond Longitudinal Study. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

Whereas the web survey was available from the start of data collection, telephone contacting efforts began 1 to 3 weeks after sample members were notified of their

inclusion in the study, by control and level of the institution. Sample members were eligible to receive a $30 incentive for completing the interview through either mode. Among respondents, 85 percent (n = 65,460) completed the interview by web and the remaining 15 percent (n = 11,570) by telephone (table 25).

Graduate students (89 percent) were more likely to complete the web survey than undergraduate students (84 percent; χ^2 (1, n = 119,553) = 351.9, $p < .001$). Potential B&B students were more likely to complete the web survey than nonpotential B&B undergraduates, 87 percent compared to 82 percent, respectively (χ^2 (1, n = 96,272) = 223.33, $p < .001$).

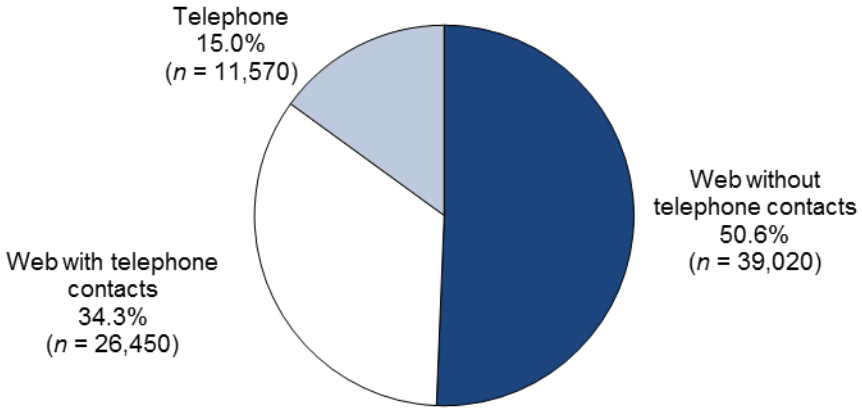
Table 25. Student interview completion rates, by mode of administration, control and level of institution, and student type: 2015–16

Control and level of institution and student type	Eligible sample	Total completes[1]		Mode of administration							
				Web total		Web nonmobile		Web mobile		Telephone	
		Number	Percent of eligible	Number	Percent of completes	Number	Percent of completes	Number	Percent of completes	Number	Percent of completes
Total	**119,550**	**77,030**	**64.4**	**65,460**	**85.0**	**48,370**	**62.8**	**17,090**	**22.2**	**11,570**	**15.0**
Control of institution											
Public	56,850	36,920	64.9	31,890	86.4	23,520	63.7	8,370	22.7	5,030	13.6
Private nonprofit	25,170	17,880	71.0	16,120	90.1	12,500	69.9	3,620	20.2	1,760	9.9
Private for-profit	37,530	22,230	59.2	17,450	78.5	12,350	55.5	5,110	23.0	4,780	21.5
Level of institution											
Less-than-2-year	3,050	1,420	46.4	960	67.5	460	32.2	500	35.3	460	32.5
2-year	24,520	13,770	56.2	10,900	79.2	6,920	50.2	3,980	28.9	2,870	20.8
4-year, non-doctorate-granting	42,730	27,680	64.8	23,210	83.8	17,210	62.2	6,000	21.7	4,470	16.2
4-year, doctorate-granting	49,260	34,160	69.4	30,390	89.0	23,780	69.6	6,610	19.3	3,770	11.0
Control and level of institution											
Public											
Less-than-2-year	370	200	53.0	140	71.4	70	35.7	70	35.7	60	28.6
2-year	17,360	10,160	58.5	8,280	81.5	5,450	53.6	2,840	27.9	1,880	18.5
4-year, non-doctorate-granting, primarily subbaccalaureate	5,610	3,470	61.8	2,940	84.7	2,150	62.1	780	22.6	530	15.3
4-year, non-doctorate-granting, primarily baccalaureate	6,950	4,740	68.2	4,160	87.7	3,090	65.2	1,070	22.5	580	12.3
4-year, doctorate-granting	26,570	18,350	69.1	16,370	89.2	12,760	69.5	3,610	19.7	1,980	10.8
Private nonprofit											
Less-than-4-year	960	520	54.0	400	77.0	210	40.4	190	36.6	120	23.0
4-year, non-doctorate-granting	11,140	7,780	69.8	6,920	88.9	5,210	66.9	1,710	22.0	870	11.1
4-year, doctorate-granting	13,910	9,990	71.8	9,100	91.1	7,260	72.7	1,840	18.4	890	8.9
Private for-profit											
Less-than-2-year	2,520	1,160	46.1	790	67.7	370	32.2	410	35.5	380	32.3
2-year	6,360	3,150	49.5	2,250	71.4	1,280	40.5	980	30.9	900	28.6
4-year	27,810	17,510	63.0	14,120	80.6	10,520	60.1	3,600	20.6	3,390	19.4
Student type											
Total undergraduate	95,020	59,390	62.5	49,690	83.7	35,390	59.6	14,300	24.1	9,700	16.3
Potential B&B student	33,760	22,070	65.4	19,130	86.7	14,440	65.4	4,690	21.2	2,940	13.3
Other undergraduate	62,520	38,060	60.9	31,210	82.0	21,460	56.4	9,760	25.6	6,850	18.0
Graduate	24,530	17,640	71.9	15,770	89.4	12,980	73.6	2,800	15.9	1,870	10.6

[1] The number of total completes excludes 1,830 partial interviews because mode of completion is not determined until the full interview is completed.
NOTE: Excludes 2,480 cases determined to be ineligible for the study using data obtained from one or more sources. B&B = Baccalaureate and Beyond Longitudinal Study. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

Telephone interviews comprised 15 percent of all completed interviews, while web interviews with telephone prompting made up about 34 percent of the interviews completed. Web interviews without telephone prompting represented about 50 percent of completed interviews (figure 8).

Figure 8. Distribution of interview respondents, by mode of administration: 2015–16



NOTE: Interviewed count excludes 1,830 partial completes. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

4.4.3 Interview Timing Burden

Interview timing. To assess the burden associated with participating in NPSAS:16, staff collected and analyzed the time required for each respondent to complete the full-scale student interview. Of special interest was the average time it took respondents to complete the interview based on mode of administration, respondent subgroup (i.e., B&B-eligible respondents, other undergraduates, and other graduates), and distinct survey path.

The NPSAS:16 full-scale student interview utilized three distinct survey paths, including a traditional full interview and two separate abbreviated (prong) interviews employed for responsive design purposes. The full NPSAS:16 student interview was composed of content areas pertinent to the study. The substantive content areas included enrollment, education experiences, financial aid, employment, income and expenses, and student background. These sections preceded the locating and incentive sections for data collection purposes. The study also implemented a two-pronged responsive survey design that employed two abbreviated sets of items. The first set of items was administered to increase study membership (prong 1), and the second was a highly-abbreviated set of items intended to reduce item variance by

interviewing select nonrespondents and to collect data on critical items (prong 2). In addition to receiving questions focused on increasing study membership, prong 1 cases received the questions administered to the prong 2 cases. Further description of prong administration and results is available in section 4.5.1.

Each web screen, or form, had an embedded time stamp to calculate the time required to complete the interview. Using the clock time on the computer or mobile device used to complete the survey, the *start time* initialized when a form first loaded, and the *end time* recorded when the respondent clicked the "Next" button on the form. Subtracting the *start time* from the *end time* yielded time spent on each form. The total instrument time was the sum of times recorded for each form.

The student survey time burden analysis did not include the following cases: partial interviews, timing outliers, and cases for which the timer kept running after the respondent exited the survey.[18] To detect total time outliers, NPSAS staff grouped interviews by student type and mode of administration. Respondents received a different number of items depending on the type of student they were, and the mode of administration affected the time to complete an interview.[19] The distribution of interview times, regardless of type and mode, was highly right-skewed, necessitating the normalization of the data using a Box-Cox power transformation (Box and Cox 1964) before determining which cases were outliers. Staff flagged cases as outliers and excluded them from the analyses using an interquartile range formula (Tukey 1977) with a multiplier of 1.5, resulting in 880 cases excluded as local (by mode) time outliers.[20] Once staff excluded these cases, they used the same transformation and outlier detection method with a multiplier of 1.5 on the entire remaining sample to detect global time outliers. Combined, these outlier detection methods led to the exclusion of 950 full-interview cases, and 180 abbreviated (prong) interviews, which was 1.4 percent of the completed cases (nonpartial interviews). Given the parameters for including and excluding cases, approximately 68,610 cases of the 79,270 total interviews (about 87 percent) were in the timing analyses (table 26). Of the total

[18] "Pseudo break-offs" occurred when the respondent exited the survey and the timer did not stop after the respondent stopped interacting with the survey. For these cases, it was impossible to distinguish the time actively spent answering survey questions. NPSAS staff included "true break-offs," or interviews for which the timer stopped appropriately when the respondent exited the survey, in the timing analysis if the respondent returned later to complete the survey. This required imputing the time spent on the first form the respondent saw when he or she began the survey again. For this, staff used the median time other respondents spent on the same form. To avoid introducing excessive imputation and uncertainty into the timing estimates, staff excluded cases that required more than two form imputations—capping the number of true break-offs at two per respondent.

[19] For example, CATI tended to take longer than self-administered web nonmobile or mobile interviews because telephone interviewers read each question aloud.

[20] Excluded as an outlier if transformed (using Box-Cox) total time > 75th percentile + (1.5 * interquartile range), or if total time < 25th percentile - (1.5 * interquartile range).

completed interviews (excluding partials), about 89 percent of the cases were in the analyses.

Table 26. Number and percentage of NPSAS interviews included and excluded from the timing report, by interview type: 2015–16

	Number of cases	Percent
Total interviews (including partials)	**79,270**	**100.00**
Interviews included in timing report	68,610	86.6
Completed full interview	58,080	73.3
Prong 1	6,650	8.4
Prong 2	3,890	4.9
Interviews excluded from timing report	10,650	13.4
Pseudo break-offs	6,150	7.8
Greater than 2 forms missing time	1,020	1.3
Prong time outliers	180	0.2
Local (mode) time outliers	880	1.1
Global time outliers	70	0.1
Partial interviews	2,360	3.0

NOTE: Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

Overall, the NPSAS:16 full interview took an average of 29.3 minutes to complete. Each interview was either a computer-assisted telephone interview or a self-administered web interview. For analysis, NPSAS staff separated the web completion mode into two categories to set apart interviews completed on a mobile device (e.g., smartphone or tablet) from those completed on a nonmobile device. Mobile interviews took 27.7 minutes on average to complete, which was significantly less time than interviews completed on nonmobile devices (28.1 minutes; $t(22{,}492.3) = 3.38$, $p < .001$) and telephone interviews (37.0 minutes; $t(20{,}121) = 69.72$, $p < .001$).[21] With an average time of 37.0 minutes to complete, telephone interviews were also significantly longer than nonmobile web interviews, which took 28.1 minutes on average to complete ($t(18{,}754.2) = 82.45$, $p < .001$). Telephone interviews necessitated that interviewers read questions aloud, leading to expected longer completion times.

There was also a significant difference in the timing burden between prong 1 and prong 2 respondents. Prong 2 respondents, who were administered only the most critical items necessary for imputation, received an average of 33 forms and had a significantly shorter timing burden of 9.5 minutes on average to complete the

[21] NPSAS staff used Satterthwaite (1946) approximation in tests with unequal variances.

interview, compared to prong 1 respondents, who took an average of 12 minutes to complete and were administered an average of 40 forms ($t(8{,}849.94) = 26.33$, $p > .001$). Table 27 provides the average time to complete the full and prong interviews by mode of interview administration.

Table 27. Average time in minutes to complete the NPSAS interview, by mode of administration and respondent group: 2015–16

				Mode of administration								
	All respondents			Web nonmobile			Web mobile			Telephone		
Respondent group	Number of cases	Avg. time	Mdn. time	Number of cases	Avg. time	Mdn. time	Number of cases	Avg. time	Mdn. time	Number of cases	Avg. time	Mdn. time
Total	**68,610**	**26.45**	**24.41**	**43,960**	**25.67**	**23.19**	**13,760**	**25.43**	**23.53**	**10,890**	**30.92**	**33.00**
Full interview	58,080	29.25	26.84	37,940	28.08	24.95	11,980	27.67	25.16	8,160	37.00	36.07
Prong 1 interview	6,650	11.98	11.22	3,580	11.43	9.78	1,020	11.16	9.86	2,060	13.32	12.95
Prong 2 interview	3,890	9.47	8.44	2,440	9.09	7.71	770	9.40	8.19	670	10.97	10.48

NOTE: The timing analysis excludes partial interviews, pseudo break-offs, and timing outliers. Avg = average. Mdn = median. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

Traditionally, the NPSAS student interview consists of consecutive sections that contain questions relating to a general content area, such as financial aid or employment. The NPSAS:16 full-scale interview follows this structure of consecutive questions by content area, with a few exceptions in structure of the survey instrument to accommodate the critical items for both the abbreviated prong respondents and full respondents. To illustrate, some background items (e.g., military status and ethnicity) occur at the beginning of the interview, after the enrollment content area, rather than with the other background items that occur at the end of the interview. Therefore, this timing report analyzes timing of the full interview by grouping items in terms of substantive content rather than consecutive sections.

The NPSAS:16 full-scale interview contained distinct respondent subgroups based on student type. Three important subgroups included respondents eligible for the follow-up B&B study, other undergraduates, and graduate students. B&B-eligible respondents, those who completed their bachelor's degree in the 2015–16 academic year, received more items than other respondents. Graduate respondents received fewer items than both the B&B-eligible and other undergraduate student groups. Table 28 provides a breakdown of the average time to complete the full interview by student type and interview section.

Table 28. Average time, in minutes, to complete the full NPSAS interview, by student type and interview section: 2015–16

	All respondents			Interview group								
				B&B eligible			Other undergraduates			Graduate students		
Interview section	Number of case	Avg. time	Mdn. time	Number of case	Avg. time	Mdn. time	Number of case	Avg. time	Mdn. time	Number of case	Avg. time	Mdn. time
Total	**58,080**	**29.25**	**26.84**	**16,540**	**33.11**	**30.55**	**27,840**	**28.96**	**26.82**	**13,700**	**25.19**	**22.61**
Enrollment	58,080	5.61	4.77	16,540	5.14	4.79	27,840	5.94	5.09	13,700	5.50	4.63
Education experiences	58,070	4.20	3.58	16,540	5.55	3.70	27,840	4.34	3.71	13,700	2.29	1.75
Financial aid	58,070	4.54	3.78	16,530	4.45	4.60	27,840	4.35	3.66	13,700	5.06	4.15
Employment	58,040	3.29	2.63	16,530	5.32	4.60	27,820	2.53	3.66	13,690	2.37	1.95
Income and expenses	58,080	6.01	5.26	16,540	5.92	5.20	27,840	6.37	2.12	13,700	5.38	4.60
Background	58,080	2.53	2.16	16,540	2.43	2.07	27,840	2.76	5.64	13,700	2.20	1.81
Locating	57,990	1.86	1.41	16,520	3.26	2.73	27,790	1.33	2.38	13,690	1.25	1.03
Incentives	57,790	1.22	0.82	16,350	1.05	0.70	27,760	1.36	0.92	13,680	1.13	0.78

NOTE: The timing analysis excludes partial interviews, pseudo breakoffs, and timing outliers. Avg = average. Mdn = median. B&B = Baccalaureate and Beyond Longitudinal Study. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

As expected, B&B-eligible respondents took longer than other student types to complete the full interview across modes, averaging 33.1 minutes. This was significantly longer than other undergraduate students, who took 29 minutes on average ($t(32{,}675.3) = 34.20$, $p < .001$). Graduate student respondents, who were not B&B-eligible, had the shortest completion time for the full interview, with an average of 25.2 minutes. This was significantly less time than B&B-eligible respondents (33.1 minutes) and other undergraduate respondents (29 minutes) ($t(30{,}154.3) = 58.02$, $p < .001$; $t(28{,}814.7) = 32.04$, $p < .001$) respectively).

The B&B-eligible respondents received a greater number of employment questions because the 1-year B&B follow-up examines work experiences and employment outcomes of bachelor's degree recipients. In addition, these respondents also received a longer locating section, necessary for participation in the follow-up studies. B&B-eligible respondents spent an average of 5.3 minutes for all employment-related questions in the full interview, significantly longer than the average 2.5 minutes for other undergraduates ($t(24{,}481.4) = 99.13$, $p > .001$) and the average 2.4 minutes for graduate students ($t(27{,}968.2) = 96.62$, $p > .001$). B&B-eligible respondents took 3.3 minutes on average to complete the locating section, significantly longer than the 1.3 minutes on average for other undergraduates ($t(22{,}648.3) = -100.8$, $p > .001$) and the 1.3 minutes on average for graduate students ($t(25{,}9540.2) = 98.28$, $p > .001$).

Table 29 provides a breakdown of the average time to complete for each student subgroup by mode of administration. As anticipated, respondents across all student

subgroups who completed the interview by telephone took the longest to complete. The longest completion time by telephone were B&B-eligible respondents, with an average of 42.9 minutes, which was significantly longer than both other undergraduates (35.5 minutes) and graduate respondents (33.3 minutes) (($t(3,3404.4) = 37.43, p < .001$) and ($t(3,210.8) = 37.80, p < .001$) respectively). This was anticipated given that the B&B-eligible respondents received the most questions in the full interview and questions were read aloud to respondents by telephone interviewers. There were no significant differences between web nonmobile and web mobile completion times for B&B-eligible respondents and other graduate student respondents; however, there was a significant difference between web nonmobile (27.9 minutes) and web mobile (26.9 minutes) completion for other undergraduate respondents ($t(14,172.9) = 6.30, p < .001$).

Table 29. Average time, in minutes, to complete the full NPSAS interview, by mode of administration and respondent group: 2015–16

				Mode of administration								
	All respondents			Web nonmobile			Web mobile			Telephone		
Status	Number of cases	Avg. time	Mdn. time	Number of cases	Avg. time	Mdn. time	Number of cases	Avg. time	Mdn. time	Number of cases	Avg. time	Mdn. time
Total	**58,080**	**29.25**	**26.84**	**37,940**	**28.08**	**24.95**	**11,980**	**27.67**	**25.16**	**8,160**	**37.00**	**36.07**
B&B-eligible	16,540	33.11	30.55	11,210	31.82	28.61	3,280	31.40	29.12	2,050	42.93	41.87
Other undergraduates	27,840	28.96	26.82	16,430	27.93	24.77	6,710	26.88	24.22	4,710	35.54	34.78
Other graduate students	13,700	25.19	22.61	10,310	24.27	21.32	1,980	24.18	21.92	1,410	33.28	32.39

NOTE: This table includes only cases that completed the full interview. Prong respondents, partial interviews, and outliers were excluded. Avg = average. Mdn = median. B&B = Baccalaureate and Beyond Longitudinal Study. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

Analyses included average time to administer each individual form within the interview, excluding the front-end locating and incentives forms. Table 30 lists the forms with the longest administration times, by form type, in descending average time. The first seven forms in the table are coders, followed by the other 10 forms with the highest average times. Coders represent some of the longest form times overall in the interview. These forms were expected to yield the longest times, given that coders require the respondent or telephone interviewer to (1) enter text strings on the form, (2) hit "Enter" to conduct a keyword search on an underlying database, and (3) select a response from the returned list of possible matches. For this reason, NPSAS staff kept coder forms separate from other form types for the form-level timing analysis.

Two coders took, on average, the most time to administer in the interview. *Occupation coder for expected occupation* (N16DEXOCC) had the longest average form time at 86.2 seconds. This form asked B&B-eligible respondents to identify and code the occupation they planned to have upon completion of their bachelor's degree using the 2014 O*NET OnLine database. The second longest form in the interview was *Major/field of study 1 at NPSAS institution* (N16BMAJ1), with an average administration time of 61.3 seconds. This form asked respondents to identify and code their declared major at the NPSAS institution using an underlying database of the 2010 CIP codes.

Two forms designed to collect a respondent's entire student financial aid package (N16CAIDGATE1 and N16CAIDGATE2) also had high administration times relative to other survey items. These forms asked respondents to select all the types of aid they had received for the academic year from comprehensive check-box lists of financial aid types. *Undergraduate financial aid in the 2015–16 academic year* (N16CAIDGATE1) collected financial aid information from undergraduate respondents and took 41.2 seconds on average. The comparable graduate student form, *Graduate financial aid in the 2015–16 academic year* (N16CAIDGATE2), took 49.2 seconds to complete.

Other forms with higher administration times also appear in table 30, including *Cost of required textbooks and materials* (N16CCSTBKS), which took 60.4 seconds to complete on average, and *Amount of graduate teaching assistantship in NPSAS year* (N16CGRTAAMT), which took an average of 44.5 seconds to complete. Another form, *Future expected wages* (N16DFUTWAGES), required B&B-eligible respondents to consider and report their most-likely beginning, highest-possible beginning, and lowest-possible beginning salary in their first job after completing their bachelor's degree. This form took 46.7 seconds on average to complete. All three forms required respondents to report received or anticipated money in dollar amounts.

Table 30. Forms with the highest average interview times, in seconds, by form and item type: 2015–16

Form name	Form description	Type	Number	Average (seconds)	Median (seconds)
Coders					
N16DEXOCC	Occupation coder for expected occupation	Coder	15,610	86.22	63.30
N16BMAJ1	Major/field of study 1 at NPSAS institution	Coder	54,820	61.28	42.72
N16ASCH01	Other postsecondary institution attended in NPSAS academic year	Coder	6,100	52.09	40.25
N16BHSCDR	Last high school attended	Coder	5,160	52.00	39.54
N16BMAJ2	Major/field of study 2 at NPSAS institution	Coder	2,260	37.14	24.55
N16BOMJ1A	Original declared major at NPSAS institution	Coder	4,410	29.81	19.80
N16FDISTNC	ZIP code of permanent residence in NPSAS academic year	Coder	43,560	24.66	20.19
Noncoder forms					
N16CCSTBKS	Cost of required textbooks and other materials during NPSAS academic year	Textbox	57,910	60.40	45.36
N16CAIDGATE2	Graduate financial aid gate for the NPSAS academic year	Checkbox list	14,350	49.21	32.42
N16DFUTWAGES	Point estimate of future expected wages	Textbox	18,320	46.71	33.23
N16CGRTAAMT	Amount of graduate teaching assistantship in NPSAS year	Textbox	930	44.45	21.25
N16CAIDGATE1	Undergraduate financial aid gate for the NPSAS year	Checkbox list	43,630	41.24	26.47
N16ANENRL	Months attended NPSAS institution between July 2015 and June 2016	Months form	64,660	41.08	30.41
N16EFIN1YEAR	Financial aid literacy: Using interest rate and inflation to determine balance in future years	Radio buttons	57,880	40.56	29.53
N16COTGRTAMT	Amount of fellowships, private scholarships, employer assistance, or Veterans benefits	Textbox	14,930	36.09	22.30
N16BIMPACT1	Activities participated in as part of undergraduate education	Likert	16,510	34.51	26.19
N16EFIN5YEAR	Using savings and interest rate to determine balance in future years	Radio buttons	57,880	40.56	22.90

NOTE: For individual form-time calculations, forms from the front-end, locating, and incentive sections were excluded from the analysis. Partial interviews and outliers were also excluded. NPSAS = National Postsecondary Student Aid Study. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

4.4.4 Telephone Interviewer Hours

As of the end of NPSAS:16 data collection, interviewers had logged approximately 46,150 hours, with roughly 11,210 telephone interviews completed. The telephone interviewers spent time on activities such as case management, including locating and contacting sample members, prompting sample members to complete interviews, reviewing case events, scheduling appointments for callbacks, recording events in the case management system, and responding to incoming help desk calls. During NPSAS:16, telephone interviewers responded to an estimated 6,140 inbound calls to the help desk.

4.4.5 Number of Calls to Sample Members

On average, interviewers made almost seven calls per sample member during the interview period. The average call counts for completed cases varied by mode of administration. Table 31 shows the average number of telephone calls by control and level of institution and student type.

Table 31. Average number of calls to sample members, by control and level of institution and student type: 2015–16

Control and level of institution and student type	Eligible sample	Number of calls	Average number of calls
Total	**119,550**	**807,280**	**6.8**
Control of institution			
Public	56,850	413,020	7.3
Private nonprofit	25,170	157,530	6.3
Private for-profit	37,530	236,740	6.3
Level of institution			
Less-than-2-year	3,050	25,330	8.3
2-year	24,520	199,820	8.2
4-year, non-doctorate-granting	42,730	275,880	6.5
4-year, doctorate-granting	49,260	306,250	6.2
Control and level of institution			
Public			
Less-than-2-year	370	3,760	10.2
2-year	17,360	147,580	8.5
4-year, non-doctorate-granting, primarily subbaccalaureate	5,610	45,880	8.2
4-year, non-doctorate-granting, primarily baccalaureate	6,950	42,900	6.2
4-year, doctorate-granting	26,570	172,900	6.5
Private nonprofit			
Less-than-4-year	960	7,210	7.5
4-year, non-doctorate-granting	11,140	71,370	6.4
4-year, doctorate-granting	13,910	83,970	6.0
Private for-profit			
Less-than-2-year	2,520	20,350	8.1
2-year	6,360	46,260	7.3
4-year	27,810	165,120	5.9
Student type			
Total undergraduate	95,020	663,430	7.0
Potential B&B student	33,760	229,650	6.8
Other undergraduate	62,520	442,460	7.1
Graduate	24,530	143,850	5.9

NOTE: Respondent count includes eligible students who met the criteria for qualification as a student interview respondent, which required completing at least a partial interview. Excludes 2,480 cases determined to be ineligible for the study using data obtained from one or more sources. B&B = Baccalaureate and Beyond Longitudinal Study. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

Respondents who completed an interview by telephone required fewer calls (5.8 on average) than cases who completed the interview by web with telephone prompting (7.3 on average). The majority of web interview respondents did not receive any calls because they had completed the interview before telephone efforts began. Table 32 shows the call counts by response status and mode of administration.

Table 32. Average number of calls, by response status and mode of administration: 2015–16

Response status and mode of administration	Eligible sample	Number of calls	Average number of calls
Total	**119,550**	**807,280**	**6.8**
Response status			
Full interview	65,490	123,780	1.9
Abbreviated 10-minute interview	4,280	67,000	15.6
Abbreviated 15-minute interview	7,250	68,160	9.4
Partial interview	1,830	25,580	14.0
Nonrespondent or exclusion	40,690	522,760	12.8
Mode of administration[1]			
Web interviews	65,460	192,270	2.9
Web, with telephone calls prompting[2]	26,450	192,270	7.3
Telephone interviews	11,570	66,670	5.8

[1] Count for mode of administration excludes the 1,830 partial interview respondents because mode is not determined until the full interview is completed.
[2] 'Web, with telephone calls prompting' is a subset of 'web interviews.'
NOTE: Excludes 2,480 cases determined to be ineligible for the study using data obtained from one or more sources. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

4.4.6 Refusal Conversion

NPSAS staff integrated refusal conversion techniques into telephone interviewer training, revisiting them throughout data collection in quality circle meetings. Project staff sorted sample members who ever refused to be interviewed, or had a gatekeeper refuse on their behalf, into a separate queue managed by a subset of interviewers who had received specialized refusal conversion training. Overall, about 11 percent of eligible sample members ever refused or had someone refuse on their behalf; of those that refused, approximately 19 percent subsequently completed the interview (table 33).

Table 33. Refusal and refusal conversion rates, by control and level of institution and student type: 2015–16

Control and level of institution and student type	Eligible sample	Ever any refusal		Interviewed, refusal conversion		
		Number	Percent of eligible	Number	Percent of refused	Percent of eligible
Total	**119,550**	**12,620**	**10.6**	**2,360**	**18.7**	**2.0**
Control of institution						
Public	56,850	6,440	11.3	1,200	18.7	2.1
Private nonprofit	25,170	2,610	10.4	530	20.1	2.1
Private for-profit	37,530	3,580	9.5	630	17.7	1.7
Level of institution						
Less-than-2-year	3,050	300	9.9	50	15.9	1.6
2-year	24,520	2,780	11.3	470	16.8	1.9
4-year, non-doctorate-granting	42,730	4,280	10.0	790	18.4	1.8
4-year, doctorate-granting	49,260	5,270	10.7	1,060	20.1	2.1
Control and level of institution						
Public						
Less-than-2-year	370	40	11.6	10	11.6	1.4
2-year	17,360	2,150	12.4	370	17.4	2.2
4-year, non-doctorate-granting, primarily subbaccalaureate	5,610	650	11.5	140	20.9	2.4
4-year, non-doctorate-granting, primarily baccalaureate	6,950	730	10.5	130	18.0	1.9
4-year, doctorate-granting	26,570	2,880	10.8	560	19.4	2.1
Private nonprofit						
Less-than-4-year	960	80	7.9	20	23.7	1.9
4-year, non-doctorate-granting	11,140	1,130	10.1	220	19.3	1.9
4-year, doctorate-granting	13,910	1,470	10.6	300	20.4	2.2
Private for-profit						
Less-than-2-year	2,520	250	10.0	40	16.7	1.7
2-year	6,360	560	8.8	80	13.6	1.2
4-year	27,810	2,700	9.7	510	18.7	1.8
Student type						
Total undergraduate	95,020	9,900	10.4	1,810	18.3	1.9
Potential B&B student	33,760	3,790	11.2	680	17.8	2.0
Other undergraduate	62,520	6,290	10.1	1,150	18.4	1.8
Graduate	24,530	2,720	11.1	550	20.4	2.3

NOTE: Respondent count includes eligible students who met the criteria for qualification as a student interview respondent, which required completing at least a partial interview. Excludes 2,480 cases determined to be ineligible for the study using data obtained from one or more sources. B&B = Baccalaureate and Beyond Longitudinal Study. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

4.4.7 Potential B&B-eligible Sample Member Identification

NPSAS staff requested that all institutions awarding bachelor's degrees identify baccalaureate recipients. Instead of waiting until the typical graduation month of May for institutions to positively identify baccalaureate recipients, NPSAS staff requested

that the enrollment lists for 4-year institutions include indicators of students who have received or will potentially receive a bachelor's during at any time during the NPSAS year, through June 30, 2016. Because most enrollment lists were received well before June 30, some sample students identified by the institution as baccalaureate candidates were determined during the interview not to be baccalaureate recipients (false positives). Likewise, some sample students not identified by the institution as baccalaureate candidates were determined during the student interview to have received baccalaureate degrees (false negatives) during the specified timeframe.

The percentage of students who were confirmed as baccalaureate recipients by the student interview is shown in table 34. Of the 26,350 students sampled as potential baccalaureate recipients and interviewed, 5,190 were not baccalaureate recipients, a false-positive rate of about 20 percent. Conversely, of the 35,960 students sampled as other undergraduates (initially not identified as B&B-eligible) and interviewed, about 1,090 were baccalaureate recipients, a false-negative rate of about 3 percent. Of the 16,550 students sampled as graduate students and interviewed, about 280 were determined to be baccalaureate recipients have completed baccalaureate requirements during the NPSAS year, a false-negative rate of about 2 percent. Overall, the false-negative rate was about 3 percent.

Table 34. Baccalaureate determination, by student type: 2015–16

Student type	Students interviewed[1]	Confirmed B&B eligibility	
		Number	Percent
Total	**78,860**	**22,540**	**28.6**
Total undergraduate	62,310	22,260	35.7
Potential B&B student	26,350	21,160	80.3
Other undergraduate	35,960	1,090	3.0
Graduate	16,550	280	1.7

[1] Includes all eligible sample members who completed the eligibility section of the student interview because confirmation of baccalaureate receipt status required contact with the sample members.
NOTE: B&B = Baccalaureate and Beyond Longitudinal Study. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

4.5 Study Members

As in the previous three NPSAS administrations, NPSAS:16 staff identified key variables across the various data sources to define a minimum set of required student-level data elements necessary to support the analytic objectives of the study. Staff classified sample members for whom data for the key variables were available

as study members, the NPSAS:16 unit of analysis. Specifically, a study member was any sample member NPSAS staff determined to be eligible for the study, per the criteria delineated in chapter 2, and who had, at a minimum, valid data from any combination of student records, student interviews, and administrative, federal, and private databases such as CPS, NSLDS, NSC, ACT files, and SAT files for the following variables;

- student type (undergraduate or graduate);
- date of birth (or age);
- sex; and
- at least 8 of the following 15 variables
  - dependency status;
  - marital status;
  - any dependents;
  - income;
  - expected family contribution (EFC);
  - degree program;
  - class level;
  - baccalaureate status;
  - months enrolled;
  - tuition;
  - received federal aid;
  - received nonfederal aid;
  - student budget;
  - race; and
  - parent education.

The final sample numbered 122,030 students (table 34), approximately 98 percent of whom (n = 119,550) were eligible for NPSAS:16. On completion of data collection, NPSAS staff determined that 94.4 percent of the eligible sample had sufficient data to satisfy the study-member definition criteria. The unweighted study membership rates (among eligible students) varied by control and level of institution, ranging from a low of 90 percent for students from public 2-year institutions to 97 percent for students from public less-than-2-year institutions. NPSAS statisticians calculated weighted study membership rates based on the institution weights and students' probabilities of selection. The weighted rate of study membership was 93 percent across all control and level of institutions.

Table 35. Number of sampled and eligible students and rates of study membership, by control and level of institution: 2015–16

Control and level of institution[2]	Sampled students	Eligible students[3]	Study members[1]	
			Unweighted percent	Weighted percent[4]
Total	**122,030**	**119,550**	**94.4**	**93.1**
Control of institution				
Public	58,370	56,850	92.7	92.4
Private nonprofit	25,510	25,170	96.6	95.4
Private for-profit	38,150	37,530	95.4	92.8
Level of institution				
Less-than-2-year	3,170	3,050	95.9	96.4
2-year	25,570	24,510	92.1	91.7
4-year, non-doctorate-granting	43,500	42,730	95.1	94.2
4-year, doctorate-granting	49,790	49,260	94.8	93.3
Control and level of institution				
Public less-than-2-year	400	370	97.0	97.5
Public 2-year	18,210	17,350	90.4	91.3
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	5,850	5,610	92.1	91.8
Public 4-year, non-doctorate-granting, primarily baccalaureate	7,090	6,950	93.6	94.0
Public 4-year, doctorate-granting	26,830	26,570	94.2	93.2
Private nonprofit, 2-year or less	990	960	96.7	99.0
Private nonprofit, 4-year, non-doctorate-granting	11,300	11,140	96.5	96.5
Private nonprofit, 4-year, doctorate-granting	14,080	13,910	96.6	94.7
Private for-profit, less-than-2-year	2,610	2,520	96.3	96.4
Private for-profit, 2-year	6,540	6,360	96.0	96.7
Private for-profit, 4-year	28,140	27,810	95.1	90.9

[1] A study member is defined as any eligible sample member for whom sufficient key data were obtained from one or more sources, including the student interview, student records, and the U.S. Department of Education's Central Processing System.
[2] Institution characteristics are based on data from the sampling frame, which was formed from the 2013–14 Integrated Postsecondary Education Data System.
[3] Sample member eligibility was determined during the student interview or from student records in the absence of a student interview.
[4] The weight described in this column is a base weight.
NOTE: Percentages are based on the unrounded count of eligible institutions within the row under consideration. Sample sizes rounded to the nearest 10. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

4.5.1 Responsive Design

Two-pronged approach. The NPSAS data collection required data from both institutions and students and the collection took place over almost 10 months. Historically, the overall NPSAS student interview response rate is approximately 70 percent (table 22), and the study member rate, combining data across multiple data sources, is about 90 percent (table 35). Both rates were lower for some institution sectors, particularly the private for-profit sector. While institution and student response rates were the primary indicators of data collection success, to also consider

variance reduction, NPSAS staff implemented a two-pronged responsive design approach for NPSAS:16.

The first prong of the responsive design strategy focused on increasing the number of study members. As described in section 4.5, a respondent is considered a study member if data for three key variables, plus at least 8 of 15 additional variables, were collected either from the student interview, institution records, or administrative data. Prong 1 cases were administered an abbreviated interview collecting the data needed to qualify a sample member as a study member. In contrast, the second prong of the responsive design focused on improving the overall quality of data for study members who had institution records data, but were the interview nonrespondents who were most likely to have variability in their imputed data. Prong 2 cases received an abbreviated interview containing only analytically important interview variables. (Because there would not be sufficient time to evaluate whether cases administered the prong 1 abbreviated interview would qualify to receive the prong 2 abbreviated interview, all those who qualified for prong 1 also received the prong 2 interview items.)

Step 1: Increasing the Number of Study Members

To implement prong 1, NPSAS staff tracked accumulation of the specific variables needed for study membership (i.e., the three key variables plus 8 of 15 additional variables) across the data collection sources from which those variables are typically obtained. Sample members with at least 6 of the variables needed for study membership were offered an abbreviated interview containing only the items needed to qualify as a study member. When the prong 1 abbreviated interview was completed, satisfying the requirements for study membership, that respondent case was included on the analysis file.

Staff identified a total of about 27,600 cases to receive the prong 1 abbreviated interview. These cases were identified at four points in time:

- July 21—sampled students in enrollment list collection waves 1–3;
- August 5—sampled students in enrollment list collection waves 4–10;
- September 2—sampled students in enrollment list collection waves 11–14; and
- September 15—sampled students in enrollment list collection waves 15–17.

Step 2: Improving Data for Study Members

To implement prong 2, NPSAS staff first created a superset of variables based on their analytic importance to NPSAS:16, then performed multiple imputation on 14

of those variables that were available from NPSAS:08 data (the last NPSAS to act as a base-year data collection for a B&B cohort). Nine of the imputed variables had relative standard errors greater than 25 percent, indicating that the imputed values would not appropriately represent the target population values. Those nine variables, plus another nine variables newly added to the NPSAS:16 main interview, were selected for imputation "on the fly," during the NPSAS:16 data collection, to determine the extent of the variability that would be present if the sample member were to remain an interview nonrespondent.

The final list of variables that were included in the prong 2 abbreviated interview and multiply imputed during data collection were[22]

- N16AMARR—Current marital status;
- N16CPRVEST—Estimated total amount borrowed in private loans;
- N16EDEPS—Financially supported children;
- N16EINCSP—Spouse's income for 2015;
- N16FMILITA—Military status: Veteran;
- N16CNODEBT—Reason for not applying: did not want to take on debt;
- N16CUGPRVT—Total amount borrowed in private loans for undergraduate education;
- N16CUGPRVEST—Estimated total amount borrowed in private loans for undergraduate education;
- N16CPAYSTIBR—Heard of income-driven repayment plans;
- N16CPAYSTLFP—Heard of loan forgiveness programs;
- N16CUSEIBR—Likelihood of using income-based repayment plan;
- N16CUSELFP—Likelihood of using loan forgiveness programs;
- N16DEXPWAGE—Future expected wages: most likely beginning salary;
- N16DHIWAGE—Future expected wages: highest possible beginning salary; and
- N16DLOWAGE—Future expected wages: lowest possible beginning salary.

Imputations were conducted at three timepoints in data collection:

- August 25—sampled students in enrollment list collection waves 1–10;
- September 9—sampled students in enrollment list collection waves 11–14; and

[22] Three of the nine variables identified through multiple imputation of NPSAS:08 data were included in the prong 2 abbreviated interview but subsequently dropped from the multiple imputation due to complications in multiply imputing the NPSAS:16 data for graduate students. The dropped variables are N16BREMEVER (Taken any remedial courses since high school), N16BREMSY (Took remedial courses in NPSAS year), and N16EPRHSD (Number of people financially supported by parents/guardians in NPSAS year).

- September 16—sampled students in enrollment list collection waves 15–17.

Sample members with high variation on the fifteen imputed items were to receive the prong 2 abbreviated interview which contained the same fifteen items used in the multiple imputation, plus items added to determine eligibility, to provide context for questions being asked, and for locating sample members for incentive payments. Staff identified a total of about 20,100 cases to receive the prong 2 abbreviated interview. All prong 1 cases who completed the abbreviated interview were then asked to continue with the prong 2 abbreviated questionnaire.

Results

Using the prong 1 abbreviated interview to target NPSAS:16 sample members who did not qualify for study membership was successful in increasing both study membership and response rates. Of the 27,600 sample members offered the prong 1 abbreviated interview, about 7,300 (approximately 34 percent) completed it. Over the course of interviewing, institution records were received for about 21,700 (79 percent) of the eligible prong 1 cases. Despite receipt of those records, however, over 1,000 sample members still lacked sufficient data to qualify as a study member and, therefore, had to complete the prong 1 abbreviated interview to be included on the analysis file (these are included in the prong 1 count).

The prong 2 abbreviated interview was completed by about 4,300 or approximately 22 percent of the 20,100 sample members offered the prong 2 interview. NPSAS statisticians evaluated 10 of the 15 variables listed above that were multiply imputed for prong 2 during data collection. N16DEXPWAGE, N16DHIWAGE, and N16DLOWAGE were not evaluated because these were only administered to potential and confirmed baccalaureate recipients. N16CUGPRVT was not evaluated because it is a point estimate of private loan borrowing for undergraduate education, and N16CUGPRVEST is evaluated instead. N16FMILITA was not evaluated because staff were able to obtain data on veterans from the VBA instead of needing to use interview response for this analysis. For the 10 variables evaluated, the statisticians used the versions of the variables that are on the analysis file, rather than the raw interview items listed above.

The ten variables that were multiply imputed during data collection were evaluated by setting their values to missing and rerunning the imputations for them, simulating how well the imputations would perform if prong 2 cases did not have the abbreviated interview. Multiple imputation was also performed for both the

originally imputed variables and the reimputed variables to capture the variance due to imputation.[23]

Table 36 compares the distribution of the ten variables between the original estimates and the reimputed estimates using multiply imputed estimates and standard errors. The estimates changed significantly for six of the ten variables, and the standard errors were lower for four of the ten variables for the original estimates containing the prong data. (Note that this analysis focuses on all values of each categorical variable, including the legitimate skips [-3s]).

[23] The analysis variable SPSINC (interview variable N16EINCSP) was reimputed, but multiple imputation was not performed due to the complexity and manual intervention required for each imputation run. Therefore, the variance for SPSINC does not account for the imputation.

Table 36. Comparison of originally imputed data and reimputed data, using multiple imputation, for selected variables

				Originally imputed data		Reimputed data	
Interview variable	Analysis variable	Analysis variable label	Level	Percent	Standard error	Percent	Standard error
Categorical variables							
N16CPAYSTIBR	AWAREIDR	Aware of income-driven student loan repayment plans	No	65.0	0.25	64.4*	0.29
			Yes	35.0	0.25	35.6*	0.29
N16CPAYSTLFP	AWARELFP	Aware of student loan forgiveness programs	No	60.4	0.28	60.7	0.27
			Yes	39.6	0.28	39.3	0.27
N16CUSEIBR	USEIDR	Likelihood of using income-driven student loan repayment plans	Not applicable	89.5	0.15	89.0*	0.16
			Very unlikely	2.2	0.07	2.4	0.09
			Somewhat unlikely	0.9	0.04	0.9	0.05
			Neither unlikely nor likely	1.7	0.06	1.8	0.07
			Somewhat likely	2.2	0.08	2.4	0.07
			Very likely	3.4	0.08	3.6*	0.09
N16CUSELFP	USELFP	Likelihood of using loan forgiveness program	Not applicable	89.5	0.15	89.7	0.15
			Very unlikely	3.2	0.08	3.2	0.09
			Somewhat unlikely	1.2	0.05	1.2	0.05
			Neither unlikely nor likely	2.0	0.06	2.0	0.07
			Somewhat likely	1.9	0.06	1.9	0.06
			Very likely	2.2	0.07	2.1	0.08
N16CNODEBT	REANOAPA	Reason for not applying: did not want to take on debt	Not applicable	80.5	0.20	80.5	0.20
			No	13.1	0.18	12.7*	0.17
			Yes	6.4	0.15	6.8*	0.15
N16AMARR	SMARITAL	Student's marital status	Single, divorced, or widowed	79.9	0.23	79.9	0.23
			Married	18.7	0.22	18.6	0.22
			Separated	1.4	0.06	1.4	0.06
N16EDEPS	DEPCHILD	Dependents: Has dependent children	No	76.6	0.25	76.6	0.26
			Yes	23.4	0.25	23.4	0.26
Continuous variables				Mean	Standard error	Mean	Standard error
N16CUGPRVT	NFEDCUM1	Cumulative nonfederal loan amount for undergraduate students	†	$1,655	$47.9	$1,877*	$44.5
N16EINCSP	SPSINC	Independent students: spouse's income	†	$39,146	$640.5	$37,646*	$459.8
N16CPRVEST	PRIVLOAN	Private (alternative) loans	†	$697	$23.6	$768*	$22.1

† Not applicable.
* The difference between the originally imputed data and the reimputed data is significant at the 0.05 level.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16)

Because there was some improvement to estimates and standard errors by using the multiple imputation approach to target students who receive an abbreviated interview, these results suggest that the two-pronged approach was modestly successful. This approach may be most appropriate for continuous income and financial aid variables, which have complex relationships with other variables,

making imputation more complicated. Further, the results highlight that the NPSAS imputation methodology works well when imputing data for a large number of cases.

Practically, the responsive design implemented for NPSAS:16 was challenging for NPSAS staff. While performing the multiple imputations during data collection was manageable, there was not enough time to perform the prong 2 post–data collection imputations at the same time as those required for NPSAS:16 data processing. However, performing imputations to evaluate the prong 2 design after the NPSAS:16 main study data were imputed was both labor and time intensive.

4.6 Evaluation of Student Interview Items

NPSAS:16 staff evaluated the student interview for data quality and response patterns. The evaluation activities included recoding and upcoding of data collected by instrument coders. For further insight into which items may have proved problematic for respondents, staff analyzed help text access rates, item nonresponse rates, and conversion text success rates.

4.6.1 Instrument Coders

To reduce respondent burden and improve data quality, the NPSAS:16 full-scale student survey made use of assisted coding systems. These coders, as described in section 4.1.2, created standardized codes from text-string responses for several interview items: additional postsecondary institutions attended during the 2015–16 academic year, majors or fields of study at the NPSAS institution, prospective future occupations, and last high school attended.

The following coder analyses are limited to a sample of respondents who either completed the full interview, completed an abbreviated interview, or completed enough of the full interview to be deemed a final partial complete (combined n = 78,860). Instances of major included in the analyses were limited to responses from the student survey, not from student records.

Recoding. Ten percent of the major and occupation codes were randomly selected from the student interview for recoding, a process in which staff reviewed the codes chosen in the survey alongside the original text strings and determined whether the coder selection most accurately described the text string provided by the respondent. The recoding process results in one of three recoding scenarios by expert staff: 1) assign the same code as the original selected in the survey, 2) recode to a different code than selected in the survey, or 3) determine that the original text strings provided by the respondent are too vague to code. This review was conducted for

majors and occupations due to variability in names of programs of study across institutions and in occupation titles given to the same or similar jobs across employers.

Overall, for the major code review, expert coding staff agreed with the respondent's choice from the survey 94 percent of the time. For the occupation code review, coding staff agreed with 90 percent of responses chosen in the interview. Given that respondents must scroll through a list of returned results on a smaller sized-screen during mobile mode administration, agreement rates between modes of administration were also compared to assess the impact that mode of administration may have on data quality.

All comparisons for major recodes (same as original, different, or text strings too vague to code) were significantly different across modes, except for the "text strings too vague to code" category between web nonmobile and telephone respondents. However, due to the administration of the major coder across multiple instances in the survey, it should be expected that smaller differences in rates between modes of administration would still yield statistically significant results given the large sample size. Rates at which majors were recoded "same as original" were significantly different between web nonmobile (94 percent) and web mobile (91 percent) respondents ($\chi 2$ (1, n = 6,274) = 12.98, $p < .001$), web nonmobile and telephone (97 percent) respondents ($\chi 2$ (1, n = 5,538) = 11.10, $p < .001$), and web mobile and telephone respondents ($\chi 2$ (1, n = 2,614) = 27.18, $p < .001$). Rates of majors recoded to a different code were significantly different between web nonmobile (5 percent) and web mobile (6 percent) respondents ($\chi 2$ (1, n = 6,274) = 7.90, $p < .01$).

Expert staff agreed with occupation codes chosen by telephone interviewers 94 percent of the time, significantly higher than the rate of agreement for web mobile respondents (89 percent) ($\chi 2$ (1, n = 616) = 3.99, $p < .05$). The rate at which telephone respondents' occupation text strings were too vague for expert coders to code (2 percent) was significantly less than it was for web nonmobile respondents (6 percent) ($\chi 2$ (1, n = 1,383) = 3.94, $p < .05$), and significantly less than the rate of vague text strings provided by web mobile respondents (8 percent) ($\chi 2$ (1, n = 616) = 7.49, $p < .01$). These differences may be attributed to the proficiency that telephone interviewers developed over time with coder use. Table 37 shows the rate of recodes for the major and occupation coders in the survey by mode of administration.

Table 37. Percentage of recoding values, by mode of administration and coding system: 2015–16

Coding system	Recoded same as original				Recoded to a different value				Text string too vague to code			
	Overall	Web non-mobile	Web mobile	Tele-phone	Overall	Web non-mobile	Web mobile	Tele-phone	Overall	Web non-mobile	Web mobile	Tele-phone
Major	93.6	93.8	91.2	96.6	4.8	4.7	6.4	2.7	1.6	1.5	2.3	0.7
Occupation	90.2	89.9	89.3	94.2	3.9	4.3	2.7	3.4	5.9	5.8	8.0	2.4

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

Upcoding. Upcoding occurred when expert coding staff assigned a code to any text string without a corresponding standardized response. Table 38 provides the rate at which additional postsecondary institution, major, last high school attended, and occupation were upcoded overall and by mode of administration.

Overall, text strings from the high school coder and postsecondary institution coder were upcoded most frequently (at 10 percent and 9 percent, respectively). Text strings from the major coder were upcoded at an overall rate of 3 percent, and occupation text strings were upcoded at an overall rate of 1 percent.

Table 38. Summary of upcoding results, by mode of administration and coding system: 2015–16

Coding system	Percent of text strings			
	Overall	Web nonmobile	Web mobile	Telephone
Postsecondary institution	9.4	9.4	7.5	12.6
Major	2.6	3.0	2.7	0.8
High school	10.1	9.7	10.2	12.1
Occupation	1.1	1.1	1.5	0.4

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

4.6.2 Help Text

At any time during the NPSAS:16 full-scale student interview, both web respondents and telephone interviewers had the option of clicking a "help" icon to view question-specific help text. In addition to this feature, some questions provided embedded hyperlinks within the question or response options to access the same help text. Hyperlinks were embedded when comprehension of a specific term was important to the quality of the data and to encourage respondents who were unsure of a term's meaning to look it up. Regardless of method of access, the help text provided definitions of key terms used in question and response option wording, as well as any explanations thought to help clarify and standardize meaning for respondents. The help-text access rate was calculated by dividing the number of times that respondents (or interviewers) accessed help text by the number of times that questions were

administered. The following help-text access rates are for questions administered to at least 25 respondents.

Overall, respondents and interviewers accessed help text less than 1 percent of the time. Web nonmobile and web mobile respondents accessed help text at a rate less than one percent, and telephone interviewers accessed help text at a rate of about 2 percent. Telephone interviewers were trained to access help text when respondents expressed uncertainty about an answer. The question-level rate of help-text access was also analyzed by mode of survey administration to identify whether specific questions proved to be more difficult to answer by mode. This analysis revealed that 11 survey questions had an overall help-text access rate of 2 percent or greater. Table 39 summarizes these results.

Table 39. Help-text access rates, by mode of administration: 2015–16

		Overall		Web nonmobile		Web mobile		Telephone	
Form	Form label	Number administered to	Percent of help text access	Number administered to	Percent of help text access	Number administered to	Percent of help text access	Number administered to	Percent of help text access
N16BIBEXP	Took International Baccalaureate (IB) courses while in high school	36,750	3.6	22,160	2.5	9,700	1.7	4,900	12.1
N16ECARRYBAL	Credit card amount carried over each month	43,860	3.3	28,890	3.7	9,680	2.3	5,290	3.1
N16CCSTBKS	Cost of required textbooks and other required instructional materials	65,490	2.9	41,640	1.8	15,220	0.5	8,640	12.9
N16CGRTAAMT	Amount of graduate teaching assistantship in 2015–16 academic year	1,000	2.6	830	0.8	90	†	70	26.0
N16CPRVLN	Took out private loans in 2015–16 academic year	34,980	2.5	21,750	3.0	8,590	1.9	4,640	1.7
N16BREMMATH	Number of remedial courses taken in 2015–16: math	46,210	2.4	28,050	0.8	11,630	5.7	6,530	3.3
N16CUGPRVT	Total amount borrowed in private loans for undergraduate education	37,890	2.4	23,280	2.5	9,440	1.4	5,180	3.3
N16BCLSDGREE	Took classes at NPSAS institution to transfer credit to degree program	1,940	2.2	1,210	2.5	450	1.1	280	2.5
N16CGRPRVT	Total amount borrowed in private loans for graduate education	8,690	2.2	6,170	2.3	1,670	1.2	840	2.7
N16EFAFDEP	Emancipated minor, ward of the court, in legal guardianship, or in foster care	65,490	2.1	41,640	0.2	15,220	0.4	8,640	14.2
N16ECARE	Number of dependent children in child care	7,560	2.1	4,290	2.4	2,000	0.9	1,280	2.7

† Not applicable
NOTE: This table only includes those items that were administered to at least 25 respondents who completed the full survey. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

The question with the highest overall help-text access rate was *Took International Baccalaureate (IB) courses while in high school* (N16BIBEXP), with a rate of 4 percent. This question asked respondents to indicate if they had taken any IB courses in high school. The question with the second highest rate of help-text access was *Credit card amount carried over each month* (N16ECARRYBAL), with a rate of 3 percent. This question asked respondents whether they usually carried a balance on their credit cards from month to month. Help-text access rates for the remaining nine interview items in the analysis ranged from 2 to 3 percent.

4.6.3 Conversion Text

To minimize item-level nonresponse in the NPSAS:16 full-scale student interview, the survey used conversion text to encourage reluctant respondents to provide an answer. When encountered in the web nonmobile and mobile interview modes, conversion text mimicked the refusal conversion strategy that would have been attempted by a telephone interviewer. If the respondent left the items blank, the survey displayed the items again, with additional text emphasizing the importance of the item and sometimes with the addition of a "don't know" option. Of the items in the NPSAS:16 full-scale student interview, a subset of 39 critical items included conversion text.

To determine a conversion rate for items with conversion text, the total number of valid responses on each of the critical items after the survey displayed conversion text was divided by the total number of cases in which the conversion text was triggered. Overall, conversion text led to an interview response 72 percent of the time after being displayed. Web nonmobile interviews accounted for 56 percent of the total instances in which conversion text was triggered and 62 percent of the total converted instances. Mobile interviews accounted for 23 percent of the total instances in which conversion text was triggered and 25 percent of the total converted instances. The remaining 21 percent of total instances in which conversion text was triggered occurred in telephone interviews, accounting for 13 percent of the total converted instances.

Table 40 displays the rates of conversion by mode of administration. Analysis excludes all items for which conversion text triggered fewer than 300 times. Of the remaining 13 critical items analyzed, conversion rates ranged from 40 percent to 97 percent, with only four items resulting in conversion rates lower than 60 percent. Items with lower conversion rates tended to request more sensitive information, such as *race* (N16FRAC1), *respondent's income in 2015* (N16EINCOM), *spouse's income in 2015* (N16EINCSP), and *parent 1: highest level of education* (N16FPARED1).

Table 40. Conversion rates for critical items, by mode of administration: 2015–16

		Total					Web nonmobile				
Item	Item description	Total percent converted	Percent converted to a valid response	Percent converted to a "don't know"	Number of cases	Number converted	Total percent converted	Percent converted to a valid response	Percent converted to a "don't know"	Number of cases	Number converted
N16BIBEXP	Took International Baccalaureate (IB) courses while in high school	96.5	96.5	†	340	330	97.8	97.8	†	230	220
N16EOTDEPS	Financially supported others in 2015–16 academic year	92.2	92.2	†	640	590	93.5	93.5	†	420	390
N16FMILIT	Military status	91.8	91.8	†	380	350	94.1	94.1	†	270	260
N16EPRHSD	Number of people financially supported by parents in 2015–16	85.1	85.1	†	2,380	2,030	86.4	86.4	†	1,410	1,220
N16CPRVLN	Took out private loans in 2015–16 academic year	84.2	84.2	†	300	260	92.6	92.6	†	200	190
N16ANENRL	NPSAS attendance: July 2015–June 2016	82.8	82.8	†	880	730	90.1	90.1	†	510	460
N16CPAYSTRAT	Heard of income-driven repayment plans or loan forgiveness programs	73.6	73.6	†	880	650	75.7	75.7	†	670	510
N16DFUTWAGES	Estimated future wages: highest, lowest, and expected wages	67.8	67.8	†	1,230	830	74.0	74.0	†	780	570
N16EPARNC	Parents' or guardians' income in 2015	63.6	25.8	37.8	1,310	840	74.0	34.3	39.7	480	360
N16FRAC1	Race	58.8	58.8	†	3,230	1,900	70.4	70.4	†	1,690	1,190
N16EINCOM	Respondent's income in 2015	58.0	35.3	22.7	810	470	68.2	41.2	27.0	370	260
N16EINCSP	Spouse's income in 2015	56.5	29.2	27.3	550	310	66.5	39.6	26.9	210	140
N16FPARED1	Parent 1: highest level of education	40.4	40.4	†	370	150	40.7	40.7	†	170	70

See notes at end of table.

Table 40. Conversion rates for critical items, by mode of administration: 2015–16—Continued

		Web mobile					Telephone				
Item	Item description	Total percent converted	Percent converted to a valid response	Percent converted to a "don't know"	Number of cases	Number converted	Total percent converted	Percent converted to a valid response	Percent converted to a "don't know"	Number of cases	Number converted
N16BIBEXP	Took International Baccalaureate (IB) courses while in high school	95.4	95.4	†	70	60	91.7	91.7	†	50	40
N16EOTDEPS	Financially supported others in 2015–16 academic year	91.0	91.0	†	140	130	88.1	88.1	†	80	70
N16FMILIT	Military status	84.6	84.6	†	50	40	87.5	87.5	†	60	50
N16EPRHSD	Number of people financially supported by parents in 2015–16	87.7	87.6	†	870	770	46.2	46.2	†	110	50
N16CPRVLN	Took out private loans in 2015–16 academic year	80.4	80.4	†	60	50	48.8	48.8	†	40	20
N16ANENRL	NPSAS attendance: July 2015–June 2016	90.2	90.2	†	220	200	47.3	47.3	†	150	70
N16CPAYSTRAT	Heard of income-driven repayment plans or loan forgiveness programs	70.9	70.9	†	180	120	51.2	51.2	†	40	20
N16DFUTWAGES	Estimated future wages: highest, lowest, and expected wages	74.8	74.8	†	310	230	19.4	19.4	†	140	30
N16EPARNC	Parents' or guardians' income in 2015	67.5	30.1	37.4	160	110	55.2	18.7	33.5	670	370
N16FRAC1	Race	70.4	70.4	†	780	550	21.6	21.6	†	770	170
N16EINCOM	Respondent's income in 2015	68.6	39.0	29.7	120	80	41.9	27.0	14.9	320	130
N16EINCSP	Spouse's income in 2015	71.7	47.2	24.5	50	40	46.1	17.9	28.2	280	130
N16FPARED1	Parent 1: highest level of education	44.2	44.2	†	80	30	37.6	37.6	†	130	50

† Not applicable.
NOTE: Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

Four survey items triggered conversion text over 1,000 times: *race* (N16FRAC1), *number of people financially supported by parents in 2015–16 academic year* (N16EPRHSD), *estimated future wages* (N16DFUTWAGES), and *parents' or guardians' income in 2015* (N16EPARNC). Again, three of these items requested information from respondents about demographic information (N16FRAC1) and family status information (N16EPRHSD and N16EPARNC) that could have been considered sensitive. The sensitivity of the information collected in these questions may have contributed to the number of times conversion text was triggered, as well as the relatively low conversion rate in comparison to some of the other critical items. The last item triggered more than 1,000 times, N16DFUTWAGES, asked respondents to predict their expected future wages in the job they intended to have after finishing their bachelor's degrees.

Significance tests were conducted to determine significant differences between modes of administration for rates of total conversion for these 13 critical items. Web nonmobile-mode conversion rates were significantly higher than telephone-mode conversion rates for the following items: N16ANENRL ($\chi 2$ (1, n = 656) = 134.07, $p < .001$), N16BIBEXP ($\chi 2$ (1, n = 277) = 4.77, $p < .05$), N16CPRVLN ($\chi 2$ (1, n = 247) = 53.52, $p < .001$), N16EINCOM ($\chi 2$ (1, n = 689) = 47.96, $p < .001$), N16EPARNC ($\chi 2$ (1, n = 1,149) = 42.33, $p < .001$), N16EPRHSD ($\chi 2$ (1, n = 1,512) = 116.10, $p < .001$), N16FRAC1 ($\chi 2$ (1, n = 2,450) = 506.89, $p < .001$), N16CPAYSTRAT ($\chi 2$ (1, n = 708) = 12.13, $p < .001$), N16DFUTWAGES ($\chi 2$ (1, n = 708) = 159.65, $p < .001$), and N16EINCSP ($\chi 2$ (1, n = 492) = 20.35, $p < .001$).

Web nonmobile-mode conversion rates were significantly higher than web mobile-mode conversion rates for the following items: N16CPRVLN ($\chi 2$ (1, n = 260) = 7.37, $p < .01$) and N16MILIT ($\chi 2$ (1, n = 323) = 5.70, $p < .05$).

Web mobile-mode conversion rates were significantly higher than telephone-mode conversion rates for the following items: N16ANENRL ($\chi 2$ (1, n = 374) = 83.66, $p < .001$), N16CPRVLN ($\chi 2$ (1, n = 99) = 10.87, $p < .001$), N16EINCOM ($\chi 2$ (1, n = 433) = 24.56, $p < .001$), N16EPARNC ($\chi 2$ (1, n = 831) = 8.05, $p < .01$), N16EPRHSD ($\chi 2$ (1, n = 978) = 116.64, $p < .001$), N16FRAC1 ($\chi 2$ (1, n = 1,545) = 370.26, $p < .001$), N16CPAYSTRAT ($\chi 2$ (1, n = 216) = 5.81, $p < .05$), N16DFUTWAGES ($\chi 2$ (1, n = 453) = 122.74, $p < .001$), and N16EINCSP ($\chi 2$ (1, n = 333) = 11.71, $p < .001$).

The most significant differences were shown between telephone mode and web nonmobile mode and between telephone mode and web mobile mode. Telephone mode demonstrated a significantly lower rate of conversion compared to web nonmobile and mobile modes. This low conversion for telephone mode may be

attributed to the presence of a telephone interviewer and the sensitive nature of some items. Web nonmobile and mobile respondents may have felt more comfortable with providing sensitive information after conversion text was triggered, compared to telephone interview respondents, who had to provide the sensitive information to an actual interviewer. Table 40 provides more detail on significant mode differences by item.

4.6.4 Item-level Nonresponse

Analysis of the rate of nonresponse for individual items in the student interview identified potentially burdensome or sensitive survey questions. NPSAS:16 staff calculated item nonresponse rates from the full-scale student survey for all items administered to 10 or more respondents. The following analysis includes all items with an overall nonresponse of 10 percent or greater. Overall, only four items met the threshold of 10 percent or more missing data. Table 41 summarizes these results.

Table 41. Item nonresponse rates for items with more than 10 percent of data missing, by mode of administration: 2015–16

		Overall		Web nonmobile		Telephone		Web mobile	
Item	Item Label	Number administered to	Percent missing	Number administered to	Percent missing	Number administered to	Percent missing	Number administered to	Percent missing
N16CGLNEST	Estimated total amount borrowed for graduate education	260	15.4	170	15.5	40	16.3	50	14.3
N16CGRPRVEST	Estimated total amount borrowed in private loans for graduate education	160	10.5	110	12.0	10	0	40	9.3
N16CULNEST	Estimated total amount borrowed for undergraduate education	1,590	13.7	770	16.7	370	10.7	460	10.9
N16CCSTOTH	Cost of required textbooks/ materials: other materials not mentioned	65,490	11.0	41,640	11.5	8,640	1.5	15,220	15.1

† Not applicable.
NOTE: This table only includes those items that were administered to at least 10 respondents who completed the full survey. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

The item with the highest overall nonresponse rate was *Estimated total amount borrowed for graduate education* (N16CGLNEST). Of the 260 respondents who received this item, about 15 percent did not provide an answer. The item with the second highest nonresponse rate was *Estimated total amount borrowed for undergraduate education* (N16CULNEST). Of the 1,590 respondents who received this item, approximately

14 percent did not provide an answer. Similar to these items, *Estimated total amount borrowed in private loans for graduate education* (N16CGRPRVEST) had a nonresponse rate of about 10 percent. These three items were administered to respondents who, on preceding questions, had not provided exact amounts of student loans borrowed. Therefore, it was likely that the higher nonresponse rates on these items was due to respondents' inability to recall this level of detail about their borrowing behavior. The fourth and final item, *Cost of required textbooks/materials: other materials not mentioned* (N16CCSTOTH), had a nonresponse rate of about 11 percent. This item was one of five items in the interview that asked respondents to provide the amount they spent on instructional materials required for classes in the 2015–16 academic year. The cognitive burden of recalling all instructional materials purchased for all classes taken in the 2015–16 academic year may have contributed to the higher nonresponse rate.

Item-level nonresponse rates were also examined by mode of administration for the four interview items with 10 percent or more missing data. Item-level nonresponse rates did not differ across modes for *Estimated total amount borrowed for graduate education* (N16CGLNEST) or *Estimated total amount borrowed in private loans for graduate education* (N16CGRPRVEST). Higher rates of nonresponse were observed in web nonmobile mode (approximately 17 percent) than in both telephone mode (about 11 percent) (χ^2 (1, n = 1131) = 7.13, $p < .01$) and web mobile mode (about 11 percent) (χ^2(1, n = 1224) = 7.74, $p < .01$) for *Estimated total amount borrowed for undergraduate education* (N16CULNEST). Higher rates of nonresponse were observed in web mobile mode (approximately 15 percent) than in both telephone mode (1 percent) (χ^2 (1, n = 23,857) = 1100.00, $p < .001$) and web nonmobile mode (about 12 percent) (χ^2 (1, n = 56,857) = 132.29, $p < .001$) for *Cost of required textbooks/materials: other materials not mentioned* (N16CCSTOTH). Finally, higher rates of nonresponse were observed in web nonmobile mode (about 12 percent) than in telephone mode (approximately 2 percent) (χ^2 (1, n = 50,274) = 820.13, $p < .001$) for *Cost of required textbooks/materials: other materials not mentioned* (N16CCSTOTH).

Chapter 5. Administrative Records Matching Overview and Outcomes

In addition to the student records collection and student interview, student data for the 2015–16 National Postsecondary Student Aid Study (NPSAS:16) also came from administrative databases, including two from the Department of Education Federal Student Aid (FSA) Office: the Central Processing System (CPS) and the National Student Loan Data System (NSLDS). Additional data sources included the National Student Clearinghouse (NSC), ACT, the College Board, and the Veterans Benefits Administration (VBA). These additional data sources were useful in providing information that could not be collected from institutions or students and when assessing the accuracy of similar information from other sources. This chapter provides detail on administrative data matching processes and outcomes.

5.1 Administrative Records Matching

CPS. To reduce institution and student burden, NPSAS staff obtained federal financial aid data from the CPS. As part of the process of applying for federal student financial aid, students enter information about themselves and their family into the FAFSA form. CPS then processes the FAFSA information and provides it to requesting postsecondary institutions as part of the process for determining student eligibility for federal financial aid.

The initial CPS match for NPSAS:16 began after students were sampled from institution enrollment lists. Because there was some overlap in data elements between CPS and the NPSAS student interview, NPSAS staff conducted this match before data collection to reduce potential student burden. Some interview data elements could be skipped by respondents if those data elements were obtained from FAFSA data. Staff conducted record matching for NPSAS:16 against CPS data for the 2015–16 financial aid year using the sample member's CPS ID—the student's SSN concatenated with the first two letters of the sample member's last name. NPSAS staff matched to CPS data a second time, near the end of data collection, to include any sample members for whom SSNs had been obtained in the interim. Staff did not submit sample members without available SSNs to the CPS for matching.

NSLDS. NPSAS staff obtained student-level data on Pell Grants and federal student loans by matching sample members to the NSLDS database. In a cooperative effort, NPSAS staff and the U.S. Department of Education initiated a record match between NPSAS records and the NSLDS database twice during the data collection period. The first match was to obtain data for preliminary analyses, and the second match occurred after the end of data collection to retrieve the most current NSLDS data. As with the CPS, sample members missing SSNs were not part of the match. The NPSAS study member had to have at least one valid grant or loan record within the NSLDS database to match successfully. The NSLDS Pell Grant and loan files included information on the year of interest and a complete federal grant and loan history for each student. All NSLDS data transfers used a password-protected NCES system transmitting over an encrypted SSL connection.

NSC. NPSAS staff obtained enrollment data for the student sample from the NSC StudentTracker service. This administrative record match provided information on institutions attended, enrollment dates, and degree completions. An individual student record would match with the NSC only if the student's institution was a participant in the NSC.[24] NPSAS staff requested StudentTracker data at the end of data collection to ensure availability of the most up-to-date student identifying data for the match. A match could yield student enrollment information for institutions other than or in addition to the NPSAS institution. All files were encrypted and transmitted over encrypted Secure File Transfer Protocol connections. Personally identifying data used for the match included sample member name, SSN, and date of birth.

ACT. NPSAS staff merged NPSAS sample member data files with ACT datafiles to obtain admissions test data. The ACT files contained survey data and a record of the highest test score registered by each student between the 2009–10 and 2014–15 academic years. NPSAS staff performed this record match after data collection to use the most updated personally identifying data (first and last name, middle initial, date of birth [DOB], and last four digits of the SSN) as matching criteria. All data transfers used a password-protected National Center for Education Statistics (NCES) system transmitting over an encrypted Secure Sockets Layer (SSL) connection.

SAT. To obtain SAT test scores and questionnaire data, NPSAS staff merged sample member datafiles with College Board records spanning high school graduation years from 2010 to 2015. If the file merge yielded multiple test records per student, it returned only the most recent record. As with ACT, staff merged files after the end

[24] For more information on NSC participation, visit www.studentclearinghouse.org.

of data collection using name, DOB, SSN, and sex. The file transfers were secured through an NCES system that required a log-in and password and an encrypted SSL connection.

VBA. In a procedure new to NPSAS studies, NPSAS staff performed a file match with the VBA to identify veterans, amounts of federal veterans education benefits, and any associated enrollment information. After the end of data collection, NPSAS staff provided a file containing SSN, name, and DOB to the VBA for data matching. The match used SSN as the primary identifier, with the other fields used to identify the proper person in rare cases of multiple matches. As with the ACT and SAT file matching, all data transmission used an NCES secure file transfer system.

5.2 Administrative Records Matching Outcomes

CPS. Table 42 summarizes the results of matching student data to the CPS overall and by institution and student characteristics. The overall matching rate for the 2015–16 academic year was about 70 percent. Match rates varied by sector of institution, ranging from a low of approximately 60 percent for private nonprofit, 4-year, doctorate-granting institutions to a high of about 91 percent at private for-profit, 2-year institutions.

Table 42. Results of Central Processing System matching for 2015–16, by control and level of institution and student type: 2015–16

Control and level of institution and student type	Eligible students	Sent to CPS	Percent	Matched to CPS	Percent[1]
Total	**119,550**	**114,100**	**95.4**	**80,330**	**70.4**
Control of institution					
Public	56,850	54,310	95.5	35,600	65.5
Private nonprofit	25,170	23,950	95.2	15,590	65.1
Private for-profit	37,530	35,830	95.5	29,140	81.3
Level of institution					
Less-than-2-year	3,050	2,980	97.7	2,510	84.2
2-year	24,510	23,800	97.1	17,590	73.9
4-year, non-doctorate-granting	42,730	41,100	96.2	30,510	74.2
4-year, doctorate-granting	49,260	46,220	93.8	29,720	64.3
Control and level of institution					
Public less-than-2-year	370	360	97.8	280	77.1
Public 2-year	17,350	16,720	96.3	11,210	67.0
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	5,610	5,400	96.2	3,670	68.0
Public 4-year, non-doctorate-granting, primarily baccalaureate	6,950	6,770	97.4	4,430	65.5
Public 4-year, doctorate-granting	26,570	25,070	94.4	16,020	63.9
Private nonprofit, less-than-4-year	960	950	98.9	810	85.2
Private nonprofit, 4-year, non-doctorate-granting	11,140	10,800	97.0	7,680	71.1
Private nonprofit, 4-year, doctorate-granting	13,910	13,040	93.8	7,840	60.1
Private for-profit, less-than-2-year	2,520	2,460	97.5	2,100	85.6
Private for-profit, 2-year	6,360	6,300	98.9	5,710	90.7
Private for-profit, 4-year	27,810	26,240	94.4	20,590	78.5
Student type					
Total undergraduate	95,020	91,470	96.3	69,190	75.6
Potential B&B student	32,500	31,500	96.9	22,580	71.7
Other undergraduate	62,510	59,970	95.9	46,610	77.7
Graduate	24,530	22,630	92.3	11,140	49.2

[1] Percentage of cases sent to CPS.
NOTE: CPS = Central Processing System. B&B = Baccalaureate and Beyond Longitudinal Study. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

Approximately 76 percent of all undergraduate students matched to the 2015–16 CPS, including 72 percent of students potentially receiving baccalaureates in 2015–16 and 78 percent of other undergraduates, while only about 49 percent of graduate students matched to the CPS. This discrepancy is understandable because nearly all institutions require undergraduate aid applicants to file a FAFSA to determine eligibility for federal Pell Grants, federal loans, and federal campus-based aid. Graduate students, however, are not usually required to file a FAFSA unless they are specifically applying for federal student loans. Graduate students often apply for financial aid directly through their institution or department. Fellowship and

assistantship amounts, which are usually not need-based and do not require the completion of the federal financial aid forms, are not available as part of the CPS data.

NSLDS. As stated previously, NSLDS matching only returned records of sample members who, at some point in time, had received Pell Grant or federal student loan funding. The NSLDS database is historical and includes information not only for the 2015–16 academic year but also for years prior. Table 43 shows the overall NSLDS match rates for study members. In this table, a match indicates that a student had at least one loan or Pell Grant, though not necessarily during 2015–16.

Table 43. Results of National Student Loan Data System loan and Pell Grant matching, by control and level of institution and student type: 2015–16

Control and level of institution and student type	Eligible students	Sent to NSLDS		Matched to NSLDS loan		Matched to NSLDS Pell Grant	
		Number	Percent	Number	Percent	Number	Percent
Total	**119,550**	**114,100**	**95.4**	**78,500**	**68.8**	**67,680**	**59.3**
Control of institution							
Public	56,850	54,310	95.5	31,060	57.2	29,180	53.7
Private nonprofit	25,170	23,950	95.2	16,630	69.4	11,150	46.6
Private for-profit	37,530	35,830	95.5	30,810	86.0	27,350	76.3
Level of institution							
Less-than-2-year	3,050	2,980	97.7	2,370	79.4	2,490	83.5
2-year	24,510	23,800	97.1	13,540	56.9	16,230	68.2
4-year, non-doctorate-granting	42,730	41,100	96.2	31,070	75.6	27,270	66.4
4-year, doctorate-granting	49,260	46,220	93.8	31,530	68.2	21,690	46.9
Control and level of institution							
Public less-than-2-year	370	360	97.8	190	53.0	270	75.4
Public 2-year	17,350	16,720	96.3	7,430	44.4	10,200	61.0
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	5,610	5,400	96.2	3,130	58.1	3,410	63.3
Public 4-year, non-doctorate-granting, primarily baccalaureate	6,950	6,770	97.4	4,370	64.6	3,780	55.9
Public 4-year, doctorate-granting	26,570	25,070	94.4	15,930	63.6	11,510	45.9
Private nonprofit, less-than-4-year	960	950	98.9	830	88.1	750	78.9
Private nonprofit, 4-year, non-doctorate-granting	11,140	10,800	97.0	7,850	72.7	5,840	54.1
Private nonprofit, 4-year, doctorate-granting	13,910	13,040	93.8	8,740	67.0	5,330	40.8
Private for-profit, less-than-2-year	2,520	2,460	97.5	2,020	82.2	2,060	84.0
Private for-profit, 2-year	6,360	6,300	98.9	5,430	86.3	5,430	86.2
Private for-profit, 4-year	27,810	26,240	94.4	22,570	86.0	19,090	72.8
Student type							
Total undergraduate	95,020	91,470	96.3	62,540	68.4	58,560	64.0
Potential B&B student	32,500	31,500	96.9	23,440	74.4	18,950	60.1
Other undergraduate	62,510	59,970	95.9	39,100	65.2	39,610	66.0
Graduate	24,530	22,630	92.3	15,950	70.5	9,120	40.3

NOTE: Both institution and student classifications were verified to correct classification errors on the sampling frame. Matching was completed on historical files that include awards made in 2015–16 and prior years. Percentage is of the number sent to NSLDS. B&B = Baccalaureate and Beyond Longitudinal Study. NSLDS = National Student Loan Data System. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

In total, NPSAS staff matched almost 78,500 eligible sample members, approximately 69 percent of total sample members with a recorded SSN. The NSLDS match rates for institution sectors ranged from a low of about 44 percent for public 2-year institutions to a high of 88 percent for private nonprofit, less-than-4-year institutions. Match rates by institution control ranged from 57 percent of public institutions to 86 percent of private for-profit institutions. The match rates by institution level ranged from 57 percent for 2-year institutions to 79 percent for less-than-2-year institutions. Approximately 68 percent of undergraduate students

matched to the loan database, while about 71 percent of the graduate students had a match.

NSLDS match yielded Pell Grant matches for 67,680 eligible sample members (about 59 percent of those sent for matching). Approximately 64 percent of undergraduate students matched to the Pell Grant database, while about 40 percent of graduate students had a match. Match rates by control and level of institution ranged from 41 percent for private nonprofit, 4-year, doctorate-granting institutions to 86 percent for private for-profit, 2-year institutions.

NSC. NSC match used enrollment and degree records for the 2015–16 academic year. An individual student record match was possible only if an institution the student attended was a participant in the NSC. NSC matches for sample members included their NPSAS-sampled institution and any other participating institutions they attended during the 2015–16 academic year.

Of the total eligible sample members, about 89,400 (about 75 percent) matched to the NSC for their NPSAS-sampled institution. By institution sector, the match rate ranged from zero for private for-profit, less-than-2-year institutions to 94 percent for public 4-year, doctorate-granting institutions. Institution level match rates ranged from 6 percent for less-than-2-year institutions to 90 percent for 4-year, doctorate-granting institutions. Institution-control match rates ranged from 42 percent for private for-profit institutions to 91 percent for public institutions. Matches to institutions other than the sample members' NPSAS institutions yielded results for about 20,610 eligible sample members (about 17 percent). Because sample members could match to multiple institutions or to a single institution other than that reported in NPSAS, these subsets are not mutually exclusive. Table 44 shows NSC match rates by control and level of institution and student type.

Table 44. Results of National Student Clearinghouse NPSAS institution matching, by control and level of institution and student type: 2015–16

Control and level of institution and student type	Eligible students	Matched for the NPSAS institution		Matched for another institution	
		Number	Percent	Number	Percent
Total	**119,550**	**89,400**	**74.8**	**20,610**	**17.2**
Control of institution					
Public	56,850	51,500	90.6	9,180	16.1
Private nonprofit	25,170	22,180	88.1	2,750	10.9
Private for-profit	37,530	15,710	41.9	8,690	23.1
Level of institution					
Less-than-2-year	3,050	190	6.1	240	7.7
2-year	24,510	17,640	71.9	3,800	15.5
4-non-doctorate-granting	42,730	27,480	64.3	11,120	26
4-year, doctorate-granting	49,260	44,100	89.5	5,450	11.1
Control and level of institution					
Public less-than-2-year	370	40	11.9	30	9.2
Public 2-year	17,350	15,790	91	3,170	18.3
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	5,610	4,270	76.1	1,480	26.5
Public 4-year, non-doctorate-granting, primarily baccalaureate	6,950	6,470	93.1	980	14.1
Public 4-year, doctorate-granting	26,570	24,930	93.8	3,510	13.2
Private nonprofit, less-than-4-year	960	520	54.5	150	15.5
Private nonprofit, 4-year, non-doctorate-granting	11,140	9,260	83.1	1,770	15.9
Private nonprofit, 4-year, doctorate-granting	13,910	12,830	92.2	1,190	8.6
Private for-profit, less-than-2-year	2,520	#	#	190	7.5
Private for-profit, 2-year	6,360	1,460	23	500	7.8
Private for-profit, 4-year	27,810	13,820	49.7	7,640	27.5
Student type					
Total undergraduate	95,020	68,940	72.6	18,380	19.3
Potential B&B student	32,500	24,720	76	7,050	21.7
Other undergraduate	62,510	44,220	70.7	11,330	18.1
Graduate	24,530	20,460	83.4	2,240	9.1

Rounds to zero.
NOTE: Sample members matched to only the NPSAS year enrollment period (July1, 2015–June 30, 2016). B&B = Baccalaureate and Beyond Longitudinal Study. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

ACT and SAT. As stated previously, ACT survey data and scores came from the matched record with the highest test score recorded for each sample member between the 2009–10 and 2014–15 academic years. In total, 28,000 eligible sample members (23 percent) matched to the ACT database (see table 45 for detail). The

institution-sector match rate ranged from 8 percent for students sampled from private for-profit, 4-year institutions to 37 percent for students sampled from public 4-year, doctorate-granting institutions. Match rates also varied by student type, with about 28 percent of undergraduate students having an ACT record on file for the matched years, and only 5 percent of the graduate students having records in the database.

NPSAS staff obtained the most recent student records of SAT, and questionnaire data were obtained for high school graduation years 2010–15. As shown in table 45, staff matched SAT data records for 25,390 eligible sample members (21 percent). Match rates by institution sector ranged from about 6 percent of students from private for-profit, 4-year institutions to 33 percent of students from public 4-year, doctorate-granting institutions.

Table 45. Results of ACT and SAT matching, by control and level of institution and student type: 2015–16

Control and level of institution and student type	Eligible students	Matched to ACT		Matched to SAT	
		Number	Percent	Number	Percent
Total	**119,550**	**28,000**	**23.4**	**25,390**	**21.2**
Control of institution					
Public	56,850	17,410	30.6	15,770	27.7
Private nonprofit	25,170	7,030	27.9	7,040	28.0
Private for-profit	37,530	3,560	9.5	2,580	6.9
Level of institution					
Less-than-2-year	3,050	430	14.0	310	10.1
2-year	24,510	5,070	20.7	4,150	16.9
4-year, non-doctorate-granting	42,730	8,540	20.0	7,840	18.3
4-year, doctorate-granting	49,260	13,960	28.3	13,100	26.6
Control and level of institution					
Public less-than-2-year	370	90	24.6	30	8.4
Public 2-year	17,350	4,100	23.6	3,460	19.9
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	5,610	1,670	29.8	1,370	24.4
Public 4-year, non-doctorate-granting, primarily baccalaureate	6,950	1,670	24.0	2,020	29.1
Public 4-year, doctorate-granting	26,570	9,880	37.2	8,880	33.4
Private nonprofit, less-than-4-year	960	170	17.3	130	13.9
Private nonprofit, 4-year, non-doctorate-granting	11,140	3,410	30.6	2,950	26.5
Private nonprofit, 4-year, doctorate-granting	13,910	3,550	25.5	4,010	28.8
Private for-profit, less-than-2-year	2,520	320	12.7	260	10.2
Private for-profit, 2-year	6,360	820	12.9	570	9.0
Private for-profit, 4-year	27,810	2,320	8.3	1,700	6.1
Student type					
Total undergraduate	95,020	26,830	28.2	24,820	26.1
Potential B&B student	32,500	9,610	29.6	9,310	28.6
Other undergraduate	62,510	17,220	27.5	15,520	24.8
Graduate	24,530	1,180	4.8	570	2.3

NOTE: Sample members were matched to the 2009–10 through 2014–15 academic years for ACT scores and to high school graduation years 2010–15 for SAT scores. B&B = Baccalaureate and Beyond Longitudinal Study. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

VBA. Veterans education benefits information was obtained for 12,660 of the eligible sample members (about 11 percent), as shown in table 46. Match rates across institution sectors were relatively consistent, ranging from approximately 7 percent of cases sampled from public 2-year institutions to about 15 percent at private for-profit, 4-year institutions. Students potentially receiving baccalaureates in 2015–16 matched to the VBA data at a rate of about 17 percent, and graduate students had a match rate of approximately 15 percent. As indicated in chapter 2, veterans who

were potential baccalaureate recipients were one of the groups oversampled in NPSAS:16.

Table 46. Results of Veterans Benefits Administration matching, by control and level of institution and student type: 2015–16

Control and level of institution and student type	Eligible students	Matched to VBA	Percent
Total	**119,550**	**12,660**	**10.6**
Control of institution			
Public	56,850	4,710	8.3
Private nonprofit	25,170	3,060	12.2
Private for-profit	37,530	4,890	13.0
Level of institution			
Less-than-2-year	3,050	280	9.1
2-year	24,510	1,860	7.6
4-year, non-doctorate-granting	42,730	5,830	13.6
4-year, doctorate-granting	49,260	4,690	9.5
Control and level of institution			
Public less-than-2-year	370	40	9.5
Public 2-year	17,350	1,260	7.3
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	5,610	490	8.8
Public 4-year, non-doctorate-granting, primarily baccalaureate	6,950	750	10.8
Public 4-year, doctorate-granting	26,570	2,170	8.2
Private nonprofit, less-than-4-year	960	80	8.4
Private nonprofit, 4-year, non-doctorate-granting	11,140	1,580	14.2
Private nonprofit, 4-year, doctorate-granting	13,910	1,480	10.6
Private for-profit, less-than-2-year	2,520	230	9.3
Private for-profit, 2-year	6,360	530	8.3
Private for-profit, 4-year	27,810	4,050	14.6
Student type			
Total undergraduate	95,020	9,030	9.5
Potential B&B student	32,500	5,570	17.1
Other undergraduate	62,510	3,460	5.5
Graduate	24,530	3,630	14.8

NOTE: VBA = Veterans Benefits Administration. B&B = Baccalaureate and Beyond Longitudinal Study. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

Chapter 6. Data File Processing and Preparation

NPSAS:16 student-level and institution-level data are compiled from institution student records, student interviews, and matches to governmental and administrative databases. These files are fully documented and available to researchers as a set of restricted-use, micro-level data files. The public may generate tables of estimates and simple regressions based upon restricted-use data via PowerStats and other publicly facing web tools available on the NCES website, https://nces.ed.gov/datalab/. This chapter provides detail on data file processing, editing, and documentation of those data files.

6.1 Overview of the NPSAS:16 Study Files

The primary analysis file (or derived file) for NPSAS:16 contains data for approximately 112,820 study members and includes more than 500 variables. Data were subject to quality checks during editing of student interview and institution record data after the start of data collection.

Complete data for NPSAS:16 are located in restricted-use files and are documented in detail in the associated codebooks. The restricted files are available to researchers who have applied for and received authorization from NCES to access restricted-use files. Researchers may obtain authorization by contacting the Institute of Education Sciences (IES) Data Security Office.[25] The restricted-use NPSAS:16 files are listed in table 47 below. NPSAS staff also used SAT, ACT, NSC, and VBA data to create derived variables, in combination with data from other sources (interview, institution record data, CPS, NSLDS). The SAT, ACT, NSC, and VBA data files are not available as source files.

[25] More information on obtaining restricted-use data files can be found at https://nces.ed.gov/statprog/rudman/.

Table 47. Restricted-use NPSAS:16 files, by filename, description, and file path: 2015–16

Filename	Description	File path
NPSAS undergraduate analysis (derived) file	Contains analytic variables derived from all NPSAS:16 data sources, as well as selected direct student-interview variables, for the 89,220 undergraduate study members.	/DATA/DERIVED/UNDERGRADUATE/N16DERIVEDUG_DATAFILE.CSV
NPSAS graduate analysis (derived) file	Contains analytic variables derived from all NPSAS:16 data sources, as well as selected direct student-interview variables, for the 23,600 graduate study members.	/DATA/DERIVED/GRADUATE/N16DERIVEDGR_DATAFILE.CSV
Student interview data file	Contains data collected from the student interviews of the 112,820 study members.	/DATA/SOURCE/N16INTERVIEW/N16INTERVIEW_DATAFILE.CSV
Student records data file	Contains data collected from the institution records of the 112,820 study members.	/DATA/SOURCE/N16STUDREC/N16STUDREC_DATAFILE.CSV
CPS 2015–16 data file	Contains data received from CPS for the 80,040 study members who matched to the 2015–16 financial aid application files.	/DATA/SOURCE/CPS16/N16CPS16_DATAFILE.CSV
CPS 2016–17 data file	Contains data received from CPS for the 43,370 study members who matched to the 2016–17 financial aid application files.	/DATA/SOURCE/CPS17/N16CPS17_DATAFILE.CSV
Institution file	Contains selected institution-level variables for the 1,750 sampled institutions that can be linked to the student interview and student records data files by the IPEDS UNITID number.	/DATA/SOURCE/N16INSTITUTION/N16INSTITUTION_DATAFILE.CSV
NSLDS loan file	Contains loan-level data received from NSLDS for approximately 76,410 matched study members who had received federal loans as of January 2017. This file includes one record for each federal loan borrowed by these respondents and provides the most recent information for that loan.	/DATA/SOURCE/NSLDS_LOAN/N16NSLDSLOAN_DATAFILE.CSV
NSLDS loan disbursement file	Contains loan disbursement-level data received from NSLDS for approximately 1,274,270 federal loans borrowed by 76,180 matched study members as of January 2017. This file includes one record for each disbursement made for a federal student loan and includes the amount and date on which each disbursement occurred.	/DATA/SOURCE/NSLDS_LOANDIS/N16NSLDSLOANDIS_DATAFILE.CSV

See notes at end of table.

Table 47. Restricted-use NPSAS:16 files, by filename, description, and file path: 2015–16—Continued

Filename	Description	File path
NSLDS FAFSA history file	Contains student award year-level data from the FAFSA, received from NSLDS, for approximately 80,670 matched study members as of January 2017. This file includes one record for each year in which a study member filed a FAFSA between 1995 and 2015. Each record includes income, expected family contribution, and select demographic information reported on the application.	/DATA/SOURCE/NSLDS_FAFSA/N16NSLDSFAFSA_DATAFILE.CSV
NSLDS loan origination file	Contains student award year-level data on federal Direct Loans awarded to 69,540 study members as of January 2017. This file includes one record for each student and year during which the student was awarded a federal Direct Loan between 2012 and 2017. The file is an aggregation of loan-level data reported by institutions to the U.S. Department of Education's COD system and provides information on loan amount eligibility and the academic year periods associated with loans disbursed during the award year.	/DATA/SOURCE/NSLDS_AWARD/N16NSLDSAWARD_DATAFILE.CSV
NSLDS enrollment file	Contains enrollment status data from NSLDS for 88,540 matched study members as of January 2017. This file includes one record for each enrollment status and effective date at an institution as it was reported to NSLDS.	/DATA/SOURCE/NSLDS_ENROLL/N16NSLDSENROLL_DATAFILE.CSV
NSLDS Pell file	Contains federal grant data received from NSLDS for approximately 66,000 study members who received a federal Pell Grant, TEACH Grant, SMART Grant, or Academic Competitiveness Grant as of January 2017. This file includes one record for each federal grant awarded and provides the amount and other information associated with the grant.	/DATA/SOURCE/NSLDS_PELL/N16NSLDSPELL_DATAFILE.CSV
Weights file	Contains the final NPSAS:16weight and variance estimation variables as a separate record for each of the 112,820 study members.	/DATA/SOURCE/N16WEIGHTS/N16WEIGHTS_DATAFILE.CSV
Weight history file	Contains all intermediate weight adjustment factors, as well as the final institution and student weights created for NPSAS:16. This file includes a separate record for each of the 112,820 study members.	/DATA/SOURCE/N16WEIGHTH/N16WEIGHTH_DATAFILE.CSV

NOTE: BPS = Beginning Postsecondary Students Longitudinal Study. COD = Common Origination and Disbursement. CPS = Central Processing System. FAFSA = Free Application for Federal Student Aid. IPEDS = Integrated Postsecondary Education Data System. NPSAS = National Postsecondary Student Aid Study. NSLDS = National Student Loan Data System.

6.2 Post-Data Collection Editing

During data collection, NPSAS staff performed quality control checks on all information collected from the student interview and institution student records to ensure the quality and accuracy of data. As one of the checks, staff examined all missing data from the interview and student records to assign specific values indicating as to why the data were missing (see table 48). For example, NPSAS staff examined skip-pattern relationships in the interview database by methodically cross-tabulating gate items with their associated nested items. An example of a gate item would be a question asking whether the respondent was employed in a specific year. Items nested within the gate would be specific questions about the employer. In many instances, gate-nest relationships spanned multiple levels within the instrument. Items nested within a gate question may also have been gate items for additional items. Consequently, validating the complex series of gate-nest relationships often required several iterations and a series of multiway cross-tabulations to ensure that the final data adhered to the item routing that study members experienced when navigating the interview.

Table 48. Description of missing data codes: 2015–16

Missing data code	Description
−1	Don't know
−3	Skipped
–5	Implied no
–7	Not administered - abbreviated
–8	Instrument error
−9	Missing

NOTE: In the institution file, IPEDS data use a value of -2 to indicate "not applicable."
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

The data cleaning and editing process for the NPSAS:16 data files was a multistage process that consisted of the following:

1. As a first step, NPSAS staff replaced all blank values in the student interview and student record data files with the appropriate initial data code (-9). They reviewed a one-way frequency distribution of every variable to confirm that no missing or blank values remained. Assigning labels to the expected values revealed any categorical outliers. Staff provided descriptive statistics for all continuous variables. They temporarily recoded all values that were less than zero to missing and examined the minimum, median, maximum, and mean values to assess reasonableness of responses. Staff also investigated anomalous data patterns and corrected them as necessary.

2. NPSAS staff identified legitimate skips in the interview data by using instrument source codes and flowcharts that documented the internal survey logic. Staff defined gate-nest question relationships and examined data for adherence to logic established in the survey design. When an item was skipped in the interview, staff replaced -9s (data missing, reason unknown) with -3s (not applicable). Staff evaluated cross-tabulations of each gate-nest combination and investigated high numbers of nonreplaced -9 codes to ensure conditional response integrity. They further checked nested values to find instances in which a legitimate skip code overwrote valid data, which typically occurred if a respondent answered a gate question and the appropriate nested items but then "backed up" within the survey to change the value of the gate, leading to an alternate path of nested items. Because responses to the first nested items remained in the database, they required further examination and editing. For student records, staff set nonapplicable items to -3 codes. For example, if a student was enrolled in a bachelor's degree program, then staff entered a value of -3 for the doctoral degree type variable.

3. Expert coders reviewed IPEDS, high school, occupation, and major codes (including the strings that interviewers or respondents could not resolve during the interview) and assigned new codes when necessary. Staff reviewed string data collected in occupation title and duty variables, as well as major variables, and sanitized strings by removing any information that could be used to identify respondents. See section 4.6.1 for more information on coder forms.

4. NPSAS staff performed logical recodes of the interview data when the value of missing items could be determined from answers to previous questions. If respondents broke-off or quit an interview prior to completion but previous responses allowed for logical recodes of post-breakoff items, logical recodes were used. For example, if the respondent's children were all over 6 years of age, the instrument skipped the question about receipt of Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) benefits. During logical recodes, this sequence would generate a value of "no" instead of a "not applicable" value.

5. For student records, NPSAS staff reviewed student level data by institution to identify any anomalous data, such as extreme values in continuous variables (e.g., SAT section scores greater than 800) or consistently missing items. Often, staff discussed unusual or consistently missing values with the institutions. In unresolvable cases, staff assigned reserve codes.

While cleaning data, staff documented question wording, response options, logical recoding, and the "applies to" text for each delivered variable from both the student records and interview data collections. For student records and interview documentation, see appendices C and G, respectively.

6.3 Composite and Derived Variable Construction

NPSAS staff derived the analytic variables by examining student-level data available from the various data sources, prioritizing data sources specific to each item, and reconciling discrepancies both within and between sources. In some cases, staff created derived or composite variables by assigning the value from the available source with the highest priority. In other cases, they recoded or combined interview items to create a derived variable (for a listing of the analysis variables derived for NPSAS:16, see appendix N). Further detail on variable derivation is available in PowerStats on the "Get more info" tab for each variable and in the restricted-use file codebooks.

Chapter 7. Weighting and Variance Estimation

This chapter provides information about the weighting procedures and variance estimation for NPSAS:16. The development of statistical analysis weights for the NPSAS:16 sample is discussed, as are procedures that can be used to produce unbiased estimates of sampling variances, including a description of how the Taylor series strata and primary sampling unit (PSU) variables and the bootstrap replicate weights were constructed. The accuracy of NPSAS:16 estimates for precision and the potential for nonresponse bias are examined. The concluding sections describe the rationale behind and the process of imputing missing data and the measures taken during data processing to protect respondent confidentiality.

The use of weights is essential to produce estimates that are representative of the NPSAS:16 target population of students. An analysis weight should be used to produce survey estimates. When testing hypotheses (e.g., conducting *t* tests, regression analyses, etc.) using weighted data from a study such as NPSAS that has a complex design, analysts also should use methods to properly estimate variances. Two such methods are the Taylor series linearization method and bootstrap replication. PSU and stratum identifiers are provided in the data file for use with the Taylor series method with or without the correction for assuming a finite population, and bootstrap replicate weights are provided for use with the bootstrap replication procedure.

7.1 Weighting

NPSAS statisticians computed statistical analysis weights for study members (defined in section 4.5) so that study members would represent the target population described in chapter 2. The statistical analysis weights compensate for the unequal probability of selection of institutions and students into the NPSAS:16 sample. The weights also adjust for multiplicity at the student level, unknown student eligibility, and nonresponse and poststratification at both the institution and the student levels.

NPSAS staff first computed the institution weight and then used it as a component of the student weight. The computed statistical analysis weights for study members were the product of the following 11 weight components:

1. institution field-test sampling adjustments (WT1);
2. institution sampling weight (WT2);
3. institution nonresponse adjustment (WT3);
4. institution poststratification adjustment (WT4);
5. student sampling weight (WT5);
6. student multiplicity adjustment (WT6);
7. student unknown eligibility adjustment (WT7);
8. student not located adjustment (WT8);
9. student refusal adjustment (WT9);
10. student other nonresponse adjustment (WT10); and
11. student poststratification adjustment (WT11).

Each weight component, described in the following sections, represents either a probability of selection or a weight adjustment. NPSAS staff computed all nonresponse and poststratification adjustments using the procedure WTADJUST in SUDAAN (RTI International 2012). The WTADJUST procedure uses a constrained logistic model to predict response. A key feature of this procedure is that the weight adjustments and weight trimming and smoothing are all accomplished in one step.

For the student poststratification adjustment, NPSAS staff set upper and lower bounds on the weights before the weight adjustment procedure. This adjustment trimmed extremely large and/or extremely small weights before the poststratification adjustment only. NPSAS staff set these bounds equal to median ± 3 times the interquartile range, where the median and interquartile range were defined for each combination of institution sector and student sampling strata. This allowed NPSAS staff to set different bounds associated with trimming extreme weights within each combination of institution sector and student sampling strata.

For both the nonresponse adjustments and the poststratification adjustment, NPSAS staff set upper and lower bounds on the weight adjustment factors. For the nonresponse adjustments, they initially set the lower bound at 1; for the poststratification adjustment, they initially set the lower bound at 0.01. During model refinement, which involves collapsing categories of candidate predictor variables

and/or excluding candidate predictor variables, NPSAS staff ran the WTADJUST procedure with no upper limit. Once they achieved convergence of the model, they tightened weight adjustment bounds to reduce the magnitude of the weight adjustment factors and the unequal weighting effects (UWEs).

In this way, NPSAS staff controlled the extreme weights and reduced the design effect due to unequal weighting. The WTADJUST procedure is designed so that the sum of the unadjusted weights for all eligible units equals the sum of the adjusted weights for the respondents. The exact formula for the weight adjustment factors calculated by the SUDAAN WTADJUST procedure is in the *SUDAAN User's Manual* (RTI International 2012).

7.1.1 Initial Institution Weights

NPSAS staff calculated the institution weights in four steps. First, they calculated the adjustment associated with the exclusion of some institutions from the full-scale frame because they were selected for the field test sample. Second, they calculated the weight associated with the probability of selection into the full-scale sample. Then NPSAS staff performed weight adjustments, including nonresponse and poststratification adjustments, that would be incorporated into the final institution weight. Finally, NPSAS staff computed the final institution weight and used it as a component of the final student weight.

Institution field-test sampling adjustment (WT1). As noted in chapter 2, most institutions that were sampled for the field test were excluded from the full-scale institution sampling frame to ensure they would not have to participate in both the field test and the full-scale study. This exclusion was done in such a way as to not compromise population coverage. Each institution on the full-scale sampling frame received a first-stage sampling weight based on the probability that it was *not* selected for the field test.

The institutions in stratum r on the institution sampling frame were partitioned as follows:

- Let $j = 1, 2, \ldots, J_1(r)$ represent those institutions not on the frame from which the field-test sample was selected (near certainty and new Integrated Postsecondary Education Data System [IPEDS] 2014–15 institutions).
- Let $j = J_1(r) + 1, J_1(r)+2, \ldots, J_2(r)$ represent those that were on the frame for the field test but were not selected.
- Let $j = J_2(r) + 1, J_2(r)+2, \ldots, J(r)$ represent the institutions in the simple random sample of n_f (r) institutions selected for the field test.

The first sampling-weight component for the full-scale study is the reciprocal of the probability of *not* being selected for the field test. That is, for the *j*th institution in stratum *r*. it is

$$W_{1r}(j)=\begin{cases}1 & \text{for } j=1,\ldots,J_1(r)\\ \dfrac{J(r)-J_1(r)}{J(r)-J_1(r)-n_f(r)} & \text{for } j=J_1(r)+1,\ldots,J_2(r)\end{cases}$$

WT1 is then equal to $W_{1r}(j)$ for each institution. Summary statistics of the subsampling weight adjustment factors were

- minimum: 1.00;
- median: 1.00; and
- maximum: 1.71.

Institution sampling weight (WT2). The sampling weight for each sample institution is the reciprocal of its probability of selection. As described in appendix B, the probability of selection for institution *i* was

$$\pi_r(i)=\begin{cases}\dfrac{n_r S_r(i)}{S_r(+)} & \text{for noncertainty selections}\\ 1 & \text{for certainty selections,}\end{cases}$$

where

n_r = the sample size in stratum *r*,

$S_r(i)$ = the measure of size for the *i*th institution in stratum *r*, and

$S_r(+)$ = the total measure of size of all institutions in stratum *r*.

Therefore, NPSAS staff assigned the institution sampling weight as follows: $WT2 = 1/\pi_r(i)$.

7.1.2 Adjusting Institution Weights

There were two additional institution weight components.

Institution nonresponse adjustment (WT3). An *institution respondent* is an institution that provided a student enrollment list from which a student sample was selected. NPSAS staff performed a weighting adjustment using the SUDAAN WTADJUST procedure, which uses a constrained logistic model to predict response, to compensate for nonresponding institutions and significantly reduce or eliminate nonresponse bias for variables included in the model. NPSAS staff selected predictor variables that were thought to predict response status, based on knowledge of the NPSAS data, and for which data were nonmissing for most respondents and

nonrespondents. The weight used in the model was the product of WT1 and WT2 multiplied by the institution's enrollment size obtained from the sampling frame. The candidate predictor variables selected by NPSAS staff to predict response status were[26]

- control and level of institution;
- 2015 Carnegie Basic classification;
- institution region;
- percentage of full-time, first-time degree/certificate-seeking undergraduate students receiving federal grant aid;
- percentage of full-time, first-time degree/certificate-seeking undergraduate students receiving state/local grant aid;
- percentage of full-time, first-time degree/certificate-seeking undergraduate students receiving institution grant aid;
- percentage of full-time, first-time degree/certificate-seeking undergraduate students receiving student loan aid;
- percentage of students enrolled who were Hispanic;
- percentage of students enrolled who were Asian or Pacific Islander, non-Hispanic;[27]
- percentage of students enrolled who were Black, non-Hispanic;
- total undergraduate enrollment;
- male undergraduate enrollment;
- female undergraduate enrollment;
- total graduate enrollment;
- male graduate enrollment;
- female graduate enrollment;
- average net price among full-time, first-time degree/certificate-seeking students receiving grant or scholarship aid;
- degree of urbanization;[28]

[26] These predictor variables come from the 2014–15 and 2015–16 IPEDS. For the continuous variables, categories were formed based on quartiles or logical breaks. For the categorical variables, categories were collapsed if there were small cells.
[27] Asian or Pacific Islander, non-Hispanic includes Native Hawaiian.
[28] Degree of urbanization is an IPEDS variable representing the urbanicity (city/suburb/rural) by population size of the institution's location.

- HBCUs status;
- HSIs status;[29] and
- number of bachelor's degrees awarded.

These characteristics were known for 95 percent or more of the respondents and nonrespondents, and any missing data were recoded into a "missing" category.

Predictors used in the nonresponse modeling included all the candidate predictor variables identified above, as well as potentially important two-way and three-way interactions. To identify these interactions, NPSAS staff used the chi-square automatic interaction detection (CHAID) algorithm (Kass 1980). CHAID is a hierarchical clustering algorithm that successively partitions individuals according to categorical predictors for a categorical dependent variable. The algorithm begins with all study individuals as a whole and cycles over each predictor, finding for each predictor an optimal partition of the individuals according to its levels. NPSAS staff retained the most significant optimal partition and applied the CHAID algorithm to the members of that partition to find further partitions using the remaining predictors. NPSAS staff stopped the algorithm after a specified number of partitioning steps or if the algorithm failed to find statistical significance among any of the partitions at a given step. The CHAID algorithm identified 9 two- or three-way interactions (later referred to as CHAID segments). Candidate predictor variables that impeded the creation of a convergent model were dropped from the final model.

NPSAS staff used the β-parameters of the logistic model, the lower and upper bounds set on the factors, and the centering constant to determine the institution nonresponse adjustment (WT3) and all other weight adjustment factors computed by the SUDAAN WTADJUST procedure (the institution poststratification adjustment (WT4), the three student nonresponse adjustments (WT8, WT9, and WT10) and the student poststratification adjustment (WT11). The exact formula for the weight adjustment factors calculated by the SUDAAN WTADJUST procedure is in the *SUDAAN User's Manual* (RTI International 2012). The final lower bound was 1.0, and the final upper bound was 2.0 for this weight adjustment. Table 49 shows the final predictor variables used in the model to determine weight adjustments and the average weight adjustment factors resulting from these variables. Summary statistics of the weight adjustment factors were

[29] Of the listed variables, only the HSI indicator no longer exists in IPEDS. An HSI proxy was created following the definition of HSI as provided by the U.S. Department of Education (https://www2.ed.gov/programs/idueshsi/definition.html) and using IPEDS Hispanic enrollment data.

- minimum: 1.00;
- median: 1.07; and
- maximum: 1.98.

Table 49. Weight adjustment factors for institution nonresponse adjustment: 2015–16

Model predictor variables	Number of respondents	Weighted response rate[1]	Average weight adjustment factor (WT3)[2]
Total	**1,750**	**89.6**	**1.14**
Control and level of institution			
Public less-than-2-year	20	83.6	1.21
Public 2-year	330	88.6	1.13
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	60	93.5	1.08
Public 4-year, non-doctorate-granting, primarily baccalaureate	80	89.7	1.12
Public 4-year, doctorate-granting	330	91.9	1.10
Private nonprofit, less-than-4-year	20	95.3	1.06
Private nonprofit, 4-year, non-doctorate-granting	280	88.1	1.16
Private nonprofit, 4-year. doctorate-granting	250	88.2	1.13
Private for-profit, less-than-2-year	50	74.3	1.35
Private for-profit, 2-year	100	82.1	1.22
Private for-profit, 4-year	240	92.2	1.13
Carnegie classification code			
Associate's	490	89.3	1.12
Research and doctoral	250	90.4	1.12
Master's	440	92.9	1.08
Baccalaureate	240	87.1	1.17
Special focus and other	150	85.1	1.19
Unavailable or unknown	190	81.6	1.28
Bureau of Economic Analysis Code (Office of Business Economics [OBE]) region[3]			
New England	100	80.4	1.27
Mideast	290	85.3	1.18
Great Lakes	260	91.2	1.10
Plains	150	93.7	1.10
Southeast	400	90.8	1.14
Southwest	190	94.9	1.07
Rocky Mountains	70	89.0	1.15
Far West	250	87.0	1.16
Outlying areas	30	90.3	1.13
Percent receiving federal grant aid[4]			
1–33	420	88.8	1.16
34–49	410	89.9	1.12
50–68	440	90.6	1.12
69 or more	400	89.1	1.15
None or unknown	90	86.6	1.21
Percent receiving state/local grant aid[4]			
1–6 or none/unknown	490	91.3	1.15
7–25	410	86.1	1.20
26–47	430	90.4	1.12
48 or more	420	91.1	1.09

See notes at end of table.

Table 49. Weight adjustment factors for institution nonresponse adjustment: 2015–16—Continued

Model predictor variables	Number of respondents	Weighted response rate[1]	Average weight adjustment factor (WT3)[2]
Percent receiving institutional grant aid[4]			
1–8 or none/unknown	490	87.8	1.17
9–42	420	90.2	1.13
43–81	420	90.2	1.15
82 or more	430	91.1	1.11
Percent receiving student loan aid[4]			
1–33	410	87.7	1.17
34–59 or none/unknown	520	90.7	1.15
60–77	410	90.7	1.12
78 or more	410	90.8	1.12
Percent enrolled: Hispanic[4]			
1–4 or none/unknown	510	90.1	1.14
5–8	430	90.5	1.13
9–19	400	90.2	1.14
20 or more	420	87.9	1.15
Percent enrolled: Asian or Pacific Islander, non-Hispanic[4]			
1	510	89.6	1.14
2–3 or none/unknown	540	90.7	1.13
4–6	320	90.5	1.13
7 or more	380	87.6	1.16
Percent enrolled: Black, non-Hispanic[4]			
1–4 or none/unknown	490	87.2	1.18
5–8	390	88.6	1.14
9–19	460	93.1	1.09
20 or more	410	89.2	1.15
Total undergraduate enrollment[4]			
1–1,344	380	81.6	1.20
1,345–4,148	440	88.4	1.14
4,149–12,680	450	91.0	1.10
12,681 or more	450	90.3	1.12
None or unknown	30	79.1	1.37
Total male undergraduate enrollment[4]			
1–524	380	82.8	1.19
525–1,773	440	87.2	1.15
1,774–5,518	450	91.0	1.10
5,519 or more or none/unknown	480	90.4	1.13
Total female undergraduate enrollment[4]			
1–788	380	79.0	1.23
789–2,329	450	89.8	1.12
2,330–7,046	440	90.4	1.10
7,047 or more or none/unknown	480	90.6	1.13

See notes at end of table.

Table 49. Weight adjustment factors for institution nonresponse adjustment: 2015–16— Continued

Model predictor variables	Number of respondents	Weighted response rate[1]	Average weight adjustment factor (WT3)[2]
Total graduate enrollment[4]			
1–500	240	85.1	1.20
501–1,424	250	90.4	1.13
1,425–3,599	250	90.8	1.11
3,600 or more	250	91.6	1.11
None or unknown	770	88.4	1.15
Total male graduate enrollment[4]			
1–158	240	86.6	1.18
159–503	250	90.5	1.14
504–1,357	250	89.5	1.13
1,358 or more or none/unknown	1,020	89.8	1.14
Total female graduate enrollment[4]			
1–285	240	84.0	1.22
286–882	250	91.3	1.12
883–2,151	250	91.8	1.09
2,152 or more or none/unknown	1,010	89.5	1.14
Average net price among students receiving grant or scholarship aid[4]			
$1–$9,199	410	88.6	1.14
$9,200–$15,754	420	91.5	1.12
$15,755–$22,528	420	91.7	1.12
$22,529 or more	400	85.9	1.17
None/unknown	110	86.7	1.20
Degree of urbanization			
Large city	460	90.8	1.12
Mid-size city	250	91.0	1.13
Small city	250	91.2	1.12
Large suburb	380	86.5	1.17
Mid-size suburb	50	74.5	1.36
Small suburb	30	97.2	1.05
Urban area on fringe of town	40	87.3	1.20
Urban area distant from town	120	91.5	1.12
Urban area remote from town	80	93.6	1.08
Rural area on fringe of town	70	91.7	1.10
Rural area distant or remote from town	20	85.9	1.19
Historically Black College or University			
Yes	40	89.5	1.12
No or unavailable or unknown	1,720	89.6	1.14
Hispanic Serving Institution			
Yes	320	88.1	1.15
No	1,430	90.0	1.14

See notes at end of table.

Table 49. Weight adjustment factors for institution nonresponse adjustment: 2015–16—Continued

Model predictor variables	Number of respondents	Weighted response rate[1]	Average weight adjustment factor (WT3)[2]
Interaction terms (CHAID segments)			
Less than 789 female undergraduate enrollment and Less than 60 percent receiving student loan aid or None/Unknown amount of student loan aid	70	63.5	1.57
Less than 789 female undergraduate enrollment and More than 59 percent receiving student loan aid			
More than 788 female undergraduate enrollment and Region in (Great Lakes, Plains, Southeast, or Southwest) and less than 7 percent receiving state/local aid	130	97.1	1.03
More than 788 female undergraduate enrollment and Region in (Great Lakes, Plains, Southeast, or Southwest) and 7–25 percent receiving state/local aid	220	90.1	1.11
More than 788 female undergraduate enrollment and Region in (Great Lakes, Plains, Southeast, or Southwest) and more than 25 percent receiving state/local aid	740	93.0	1.10
More than 788 female undergraduate enrollment and Region in (New England, Mideast, Rocky Mountains, Far West, or Outlying areas) and less than 286 female graduate enrollment	60	70.4	1.46
More than 788 female undergraduate enrollment and Region in (New England, Mideast, Rocky Mountains, Far West, or Outlying areas) and 286–2,151 female graduate enrollment	210	92.0	1.10
More than 788 female undergraduate enrollment and Region in (Great Lakes, Plains, Southeast, or Southwest) and more than 2,151 female graduate enrollment	330	85.6	1.17

[1] The response rate is expressed as a percentage.
[2] The average weight adjustment is expressed as a number.
[3] New England = Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont; Mideast = Delaware, District of Columbia, Maryland, New Jersey, New York, Pennsylvania; Great Lakes = Illinois, Indiana, Michigan, Ohio, Wisconsin; Plains = Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota; Southeast = Alabama, Arkansas, Florida, Georgia, Kentucky, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, West Virginia; Southwest = Arizona, New Mexico, Oklahoma, Texas; Rocky Mountains = Colorado, Idaho, Montana, Utah, Wyoming; Far West = Alaska, California, Hawaii, Nevada, Oregon, Washington; Outlying areas = Puerto Rico.
[4] Continuous variables were categorized using quartiles for the cut points; missing values were assigned to their own category or collapsed into the reference category.
NOTE: CHAID = chi-square automatic interaction detection. Categories were formed from continuous variables based on quartiles. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

Institution poststratification adjustment (WT4). After adjusting for field-test sampling, the inverse of the probability of selection into the full-scale sample, and nonresponse, the institution weight was further adjusted to meet enrollment totals (control totals) by institution type and size (*small* vs. *large*[30]). This adjustment ensures that the resultant weight adequately represents the student target population. The weight used in the poststratification model was the product of WT1, WT2, and WT3 multiplied by the institution's enrollment from the sampling frame. The enrollment totals came from the 12-month unduplicated head count from the 2015–16 IPEDS Institutional Characteristics Header component, Fall, and 12-month Enrollment file.

[30] Institution size was determined based on the median total enrollment as a cut point within each institution type.

The poststratification was done using student enrollment counts rather than institution counts because all NPSAS inferences are at the student level and not at the institution level. Additionally, the institutions were selected with probability proportional to size (PPS), with the size being counts of students. This method of sampling and weighting do not yield an accurate estimate of institutions.

Table 50 shows the characteristics associated with the control totals and the average weight adjustment factors by these characteristics. Summary statistics of the weight adjustment factors were

- minimum: 0.67;
- median: 1.01; and
- maximum: 1.33.

The final lower bound was 0.5 and the final upper bound was 10 for this weight adjustment.

Table 50. Weight adjustment factors for institution poststratification: 2015–16

Model predictor variables	Control total[1]	Average weight adjustment factor (WT4)
Total	**27,687,610**	**1.01**
Public less-than-2-year	71,577	1.09
Public 2-year, small	1,289,520	1.00
Public 2-year, large	8,295,977	1.00
Public 4-year, non-doctorate-granting, primarily subbaccalaureate, small	168,300	1.24
Public 4-year, non-doctorate-granting, primarily subbaccalaureate, large	1,012,661	0.98
Public 4-year, non-doctorate-granting, primarily baccalaureate, small	301,142	1.33
Public 4-year, non-doctorate-granting, primarily baccalaureate, large	1,227,430	0.96
Public 4-year, doctorate-granting, small	1,564,350	1.01
Public 4-year, doctorate-granting, large	5,651,719	1.00
Private nonprofit, less-than-4-year	119,316	1.10
Private nonprofit, 4-year, non-doctorate-granting, small	204,930	0.83
Private nonprofit, 4-year, non-doctorate-granting, large	1,525,982	1.02
Private nonprofit, 4-year, doctorate-granting, small	297,221	0.83
Private nonprofit, 4-year, doctorate-granting, large	3,038,743	1.02
Private for-profit, less-than-2-year, small	68,848	0.67
Private for-profit, less-than-2-year, large	344,227	1.07
Private for-profit, 2-year, small	73,642	1.23
Private for-profit, 2-year, large	456,103	0.99
Private for-profit, 4-year, small	136,303	0.97
Private for-profit, 4-year, large	1,839,619	1.02

[1] Control totals are the sum of enrollment across institutions based on IPEDS:15 enrollment data.
NOTE: Size for poststratification weighting classes was based on the median enrollment within sector or state for the institutions on the sampling frame. IPEDS = Integrated Postsecondary Education Data System.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

7.1.3 Initial Student Weights

The three initial student weight components are the initial sampling weight and weight adjustment factors for student multiplicity and student unknown eligibility. Each of these components is described in this section. As discussed in appendix B, NPSAS staff designed the institution-specific student sampling rates to obtain the desired sample sizes and achieve nearly equal initial sampling weights within the combined institution and student strata.

Student sampling weight (WT5). NPSAS staff defined the overall student sampling strata by crossing the institution sampling strata with the student strata. (For the overall sampling rates for these sampling strata, see appendix B.) They systematically selected the sample students from the enrollment lists at institution-specific rates that were inversely proportional to the institution's probability of selection. Specifically, the institution-specific sampling rate was the overall stratum student sampling rate divided by the institution's probability of selection, or

$$f_{s|i} = \frac{f_s}{\pi_r(i)},$$

where f_s = the overall student stratum sampling rate and $\pi_r(i)$ = the institution's probability of selection.

Initial student sampling rates were set prior to receiving enrollment lists using IPEDS data. If an institution's enrollment list contained more students than expected based on the IPEDS data, the initial student sampling rates across strata would yield a larger-than-expected sample size for that institution. Likewise, if the enrollment list count was smaller than expected based on the IPEDS data, the initial student sampling rates across strata would yield a smaller-than-expected sample size for that institution. To maintain control over the sample sizes and meet stratum yield targets, NPSAS staff adjusted the sampling rates, when necessary, so that the number of students selected within an institution usually did not exceed 300. NPSAS staff imposed a minimum sample size constraint of 10 students to ensure sufficient yield for variance estimation.

NPSAS staff calculated the student sampling weight as the reciprocal of the adjusted institution-specific student stratum sampling rates, or

$$\text{WT5} = 1/f_{s|i}.$$

Student multiplicity adjustment (WT6). Students who attended more than one eligible institution during the 2015–16 academic year had multiple chances of being

selected for the study; that is, they could have been selected from any of the institutions they attended. These students therefore had a higher probability of being selected than was represented in their sampling weight.

NPSAS staff adjusted for this multiplicity by dividing these students' sampling weight by the number of institutions the student attended that were eligible for sample selection. Specifically, they defined the student multiplicity weight adjustment factor as

$$WT6 = 1/M,$$

where M is the multiplicity, or number of eligible institutions attended. NPSAS staff determined multiplicity using information from the student interview, the Pell Grant payment file, and the National Student Loan Data System.

Summary statistics of the weight adjustment factors were

- minimum: 0.14;
- median: 1.00; and
- maximum: 1.00.

Student unknown eligibility adjustment (WT7). NPSAS staff could not determine final eligibility status for nonresponding students. They treated these students as eligible and adjusted their weights to compensate for the small portion of these students who were actually ineligible (as described below).

NPSAS staff based the unknown eligibility weight adjustment factors on the calculated rate of eligibility among students with known eligibility status within weighting classes. They defined these classes by the intersection of institution type with the students' matching status (matched to NSLDS, matched to CPS only, and no match) to financial aid files (NSLDS and CPS). Table 51 shows the weight adjustment factors applied to the students with unknown eligibility. For the known-eligible students, NPSAS staff set the weight adjustment factor equal to 1.

Table 51. Weight adjustment factors for unknown student eligibility status: 2015–16

Weighting class (institution level, by student type, by matching status to financial aid files)	Number adjusted for unknown eligibility	Weight adjustment factor (WT7)
Public less than 2-year		
Matched NSLDS	#	0.96
Matched CPS file only	#	0.93
No matches	10	0.85
Public 2-year		
Matched NSLDS	20	0.98
Matched CPS file only	10	0.96
No matches	1,620	0.89
Public 4-year, non-doctorate-granting, primarily subbaccalaureate, undergraduate		
Matched NSLDS	10	0.98
Matched CPS file only	#	0.94
No matches	440	0.92
Public 4-year, non-doctorate-granting, primarily subbaccalaureate, graduate		
Matched NSLDS	†	†
Matched CPS file only	†	†
No matches	†	†
Public 4-year, non-doctorate-granting, primarily baccalaureate, undergraduate		
Matched NSLDS	10	0.99
Matched CPS file only	#	0.97
No matches	290	0.95
Public 4-year, non-doctorate-granting, primarily baccalaureate, graduate		
Matched NSLDS	#	1.00
Matched CPS file only	†	†
No matches	120	0.94
Public 4-year, doctorate-granting, undergraduate		
Matched NSLDS	50	1.00
Matched CPS file only	10	0.99
No matches	1,290	0.98
Public 4-year, doctorate-granting, graduate		
Matched NSLDS	10	1.00
Matched CPS file only	10	0.98
No matches	220	0.98
Private nonprofit, less-than-4-year		
Matched NSLDS	†	†
Matched CPS file only	#	0.94
No matches	30	0.91
Private nonprofit, 4-year, non-doctorate-granting, undergraduate		
Matched NSLDS	10	0.99
Matched CPS file only	10	0.99
No matches	230	0.97

See notes at end of table.

Table 51. Weight adjustment factors for unknown student eligibility status: 2015–16—Continued

Weighting class (institution level, by student type, by matching status to financial aid files)	Number adjusted for unknown eligibility	Weight adjustment factor (WT7)
Private nonprofit, 4-year, non-doctorate-granting, graduate		
Matched NSLDS	10	1.00
Matched CPS file only	†	†
No matches	100	0.95
Private nonprofit, 4-year, doctorate-granting, undergraduate		
Matched NSLDS	20	0.99
Matched CPS file only	#	0.99
No matches	310	0.96
Private nonprofit, 4-year, doctorate-granting, graduate		
Matched NSLDS	10	0.99
Matched CPS file only	#	0.98
No matches	180	0.97
Private for-profit, less-than-2-year		
Matched NSLDS	10	0.97
Matched CPS file only	#	0.94
No matches	90	0.95
Private for-profit, 2-year		
Matched NSLDS	10	0.98
Matched CPS file only	10	0.96
No matches	240	0.88
Private for-profit, 4-year, undergraduate		
Matched NSLDS	50	0.98
Matched CPS file only	10	0.96
No matches	1,030	0.95
Private for-profit, 4-year, graduate		
Matched NSLDS	10	0.98
Matched CPS file only	#	0.99
No matches	230	0.99

† Not applicable.
# Rounds to zero.
NOTE: Matched NSLDS indicates that the student received a Pell or Direct Loan in 2015–16 academic year. CPS = Central Processing System. NSLDS = National Student Loan Data System. Sample sizes rounded to the nearest 10. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

Summary statistics of the weight adjustment factors were

- minimum: 0.85;
- median: 1.00; and
- maximum: 1.00.

7.1.4 Adjusted Student Weights

There are four additional student weight components. NPSAS staff adjusted the student weights for nonresponse in three stages: first, inability to locate the student, then interview refusal, and finally, other nonresponse (student located, did not refuse, did not respond)—because the predictors of response propensity were potentially different for each of these nonresponse outcomes. Adjusting for these three types of nonresponse separately achieved greater reduction in nonresponse bias because different characteristics were significant predictors of response propensity at each stage. The fourth additional weight component is poststratification.

Student not located adjustment (WT8). The first type of adjustment for student nonresponse was an adjustment for the inability to locate the student[31] ("not located"). NPSAS staff chose predictor variables that were thought to predict not located status, based on knowledge of past NPSAS data collections, and for which data were nonmissing for both study members and nonstudy members. The candidate predictor variables were

- control and level of institution attended (categorical);
- region of institution attended (categorical);
- enrollment of institution attended from IPEDS 2015–16 file (categorical);
- student type derived from multiple data sources (categorical);
- baccalaureate recipient as known at time of sampling (yes/no);
- age group (categorical);
- student type as known at time of sampling (categorical);
- veteran status (yes/no);
- race/ethnicity (categorical);
- sex (two levels);
- CPS record available (yes/no);
- CPS parents' highest education (categorical);
- CPS has dependents (yes/no);
- CPS marital status (categorical);
- Social Security number obtained from enrollment list (yes/no);

[31] Refer to section 4.2.3 for further details on the process of locating sample members.

- any aid receipt (yes/no);
- federal aid receipt (yes/no);
- Pell Grant receipt (yes/no);
- Pell Grant amount (categorical);
- Direct Loan receipt (yes/no);
- Direct Loan amount (categorical);
- Parent Loan for Undergraduate Students (PLUS) amount (categorical);
- institution aid receipt (yes/no);
- state aid receipt (yes/no);
- telephone number count (categorical);
- e-mail address count (categorical);
- mailing address count (categorical)
- student records for key data elements (yes/no); and
- Mahalanobis distance value (categorical).[32]

Predictors used in nonresponse modeling included all the candidate predictor variables identified, as well as potentially important interactions. NPSAS staff used CHAID to identify these interactions (see the description in section 7.1.2). Application of the CHAID algorithm provided interaction terms for each of the three nonresponse adjustment models. For each model, NPSAS staff stopped the algorithm after a specified number of partitioning steps or if the algorithm failed to find statistical significance among any of the partitions at a given step. The CHAID algorithm resulted in identification of two-way and three-way interactions. The interaction terms (CHAID segments) identified were treated as additional candidate predictor variables. Candidate predictor variables that impeded the creation of a convergent model were dropped from the final model. For example, the final model for WT8 includes the indicator of whether a Pell Grant was received but does not include the categorical Pell Grant amount; whereas, in the final model for WT9, the receipt indicator is not included but the categorical amounts are included.

NPSAS staff computed the weight adjustments using SUDAAN's WTADJUST procedure.[33]

[32] Mahalanobis distance value is a measurement that quantifies the distance between an individual and the average respondent.
[33] See the description of the SUDAAN procedure at the beginning of this chapter.

Table 52 shows the final predictor variables used in the model to determine weight adjustments and the average weight adjustment factors resulting from these variables. Summary statistics of the weight adjustment factors were

- minimum: 1.00
- median: 1.00; and
- maximum: 4.73.

The final lower bound was 1.0 and the final upper bound was unbounded to get convergence for this weight adjustment.

Table 52. Weight adjustment factors for student not located adjustment: 2015–16

Model predictor variables	Number located	Weighted response rate	Average weight adjustment factor (WT8)
Total	**117,460**	**97.8**	**1.02**
Control and level of institution			
Public less-than-2-year	370	99.8	1.00
Public 2-year	16,860	97.6	1.03
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	5,380	97.6	1.02
Public 4-year, non-doctorate-granting, primarily baccalaureate	6,300	97.9	1.02
Public 4-year, doctorate-granting	26,680	97.7	1.02
Private nonprofit, less-than-4-year	1,150	99.7	1.00
Private nonprofit, 4-year, non-doctorate-granting	10,690	99.0	1.01
Private nonprofit, 4-year, doctorate-granting	14,070	98.0	1.01
Private for-profit, less-than-2-year	2,500	98.5	1.02
Private for-profit, 2-year	6,020	98.8	1.01
Private for-profit, 4-year	27,450	96.8	1.01
Bureau of Economic Analysis Code (Office of Business Economics [OBE]) region[1]			
New England	4,620	97.5	1.02
Mideast	18,560	97.9	1.02
Great Lakes	16,140	98.4	1.01
Plains	8,700	97.0	1.03
Southeast	29,860	98.1	1.02
Southwest	14,250	98.5	1.01
Rocky Mountains	5,910	98.0	1.01
Far West	17,660	96.7	1.02
Outlying areas	1,770	99.7	1.00
Institution total enrollment[2]			
2,557 or fewer	29,660	98.8	1.01
2,558–10,368	29,290	98.8	1.01
10,369–26,982	29,320	97.7	1.02
26,983 or more	29,190	97.0	1.03
Derived student type			
Undergraduate	93,190	97.7	1.02
Graduate student (excluding doctoral-professional practice)	21,700	98.6	1.01
Doctoral-professional practice	2,570	98.9	1.01
Baccalaureate recipient status (sampled)			
Yes	37,100	98.7	1.01
No	80,360	97.7	1.02
Age as of Dec 31, 2015			
15–23	51,190	97.8	1.02
24–29	28,020	97.6	1.02
30 or more	38,250	98.0	1.01
Sampled student type			
Undergraduate	95,140	97.7	1.02
Graduate (excluding doctoral-professional practice)	20,220	98.5	1.01
Doctoral-professional practice	2,100	99.1	1.00
Veteran status			
Yes	10,390	97.8	1.02
No	107,070	97.8	1.02

See notes at end of table.

Table 52. Weight adjustment factors for student not located adjustment: 2015–16—Continued

Model predictor variables	Number located	Weighted response rate	Average weight adjustment factor (WT8)
Race/ethnicity			
White, non-Hispanic	59,580	98.8	1.01
Black, non-Hispanic	18,850	99.0	1.01
Hispanic	18,950	97.7	1.02
Asian, non-Hispanic	9,350	97.7	1.02
American Indian or Alaskan Native, non-Hispanic	790	98.9	1.01
Native Hawaiian or other Pacific Islander, non-Hispanic	520	98.8	1.01
More than one race, non-Hispanic	3,320	99.5	1.00
Unknown	6,110	82.0	1.15
Sex			
Male	50,330	97.3	1.02
Female	67,130	98.2	1.01
CPS record available			
Yes	80,300	100.0	1.00
No	37,160	93.6	1.06
CPS—Parents' highest education			
Middle school/junior high	3,010	99.9	1.00
High school	24,230	100.0	1.00
College or beyond	42,220	100.0	1.00
Unknown/missing	10,830	100.0	1.00
CPS—Has dependents			
Yes	22,330	100.0	1.00
No or Unknown	57,970	100.0	1.00
CPS—Marital status			
Single/separated/divorced/widowed/unknown	103,040	97.6	1.02
Married/remarried	14,420	100.0	1.00
Social Security number available			
Yes	112,730	98.6	1.01
No	4,730	84.9	1.16
Any aid status			
Received	88,250	99.6	1.00
Did not receive	9,550	100.0	1.00
Unknown	19,660	91.0	1.09
Federal aid status			
Received	70,050	99.8	1.00
Did not receive	37,960	96.8	1.02
Unknown	9,450	90.4	1.10
Pell Grant status			
Received	40,070	100.0	1.00
Did not receive	53,120	96.2	1.03
Direct Loan status			
Received	49,990	100.0	1.00
Did not receive	67,470	96.5	1.03
Institutional aid status			
Received	33,130	99.3	1.01
Did not receive	64,660	98.1	1.01
Unknown	19,680	94.5	1.05

See notes at end of table.

Table 52. Weight adjustment factors for student not located adjustment: 2015–16—Continued

Model predictor variables	Number located	Weighted response rate	Average weight adjustment factor (WT8)
State aid status			
Received	14,940	99.8	1.00
Did not receive	82,480	98.1	1.01
Unknown	20,040	94.3	1.05
Telephone number count			
0	1,930	77.1	1.10
1	39,980	96.8	1.03
2	43,480	99.0	1.01
3 or more	32,070	99.4	1.00
E-mail address count			
0	1,310	81.3	1.22
1	30,710	94.3	1.04
2	62,090	98.9	1.01
3 or more	23,350	99.8	1.00
Mailing address count			
0	1,050	75.9	1.33
1	54,100	97.6	1.02
2	32,600	98.4	1.01
3 or more	29,720	98.7	1.01
Student record complete indicator			
Complete data	109,300	98.1	1.01
Partial/No data	8,160	93.4	1.06
Mahalanobis Distance value[2 3]			
1.17–2.65	29,660	99.0	1.01
2.66–3.17	29,360	97.7	1.02
3.18–3.72	29,200	97.3	1.02
3.73 or more	29,240	97.3	1.02

[1] New England = Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont. Mideast = Delaware, District of Columbia, Maryland, New Jersey, New York, Pennsylvania. Great Lakes = Illinois, Indiana, Michigan, Ohio, Wisconsin. Plains = Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota. Southeast = Alabama, Arkansas, Florida, Georgia, Kentucky, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, West Virginia. Southwest = Arizona, New Mexico, Oklahoma, Texas. Rocky Mountains = Colorado, Idaho, Montana, Utah, Wyoming. Far West = Alaska, California, Hawaii, Nevada, Oregon, Washington. Outlying areas = Puerto Rico.
[2] Categories were defined by quartiles.
[3] Mahalanobis distance value is a measurement that quantifies the distance between an individual and the average respondent.
NOTE: CPS = Central Processing System. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

Student refusal adjustment (WT9). The second stage of the student nonresponse adjustment was an adjustment for refusal, given that the student was located. This additional type of nonresponse adjustment was made to compensate further for the potential student nonresponse bias. The same SUDAAN procedure was used in this adjustment as was used in the adjustment for not located students (WT8). The same candidate predictor variables were used to predict refusal, and the same type of CHAID analysis was used to detect important interactions. Table 53 shows the final predictor variables used in the model to determine weight adjustments and the

average weight adjustment factors resulting from these variables. Summary statistics of the weight adjustment factors were

- minimum: 1.00;
- median: 1.00; and
- maximum: 2.23.

The final lower bound was 1.0 and the final upper bound was 5 for this weight adjustment.

Table 53. Weight adjustment factors for student refusal adjustment: 2015–16

Model predictor variables	Number did not refuse	Weighted response rate	Average weight adjustment factor (WT9)
Total	**115,580**	**97.8**	**1.01**
Control and level of institution			
Public less-than-2-year	360	98.7	1.01
Public 2-year	16,380	97.1	1.03
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	5,270	97.8	1.01
Public 4-year, non-doctorate-granting, primarily baccalaureate	6,180	98.1	1.02
Public 4-year, doctorate-granting	26,240	97.9	1.02
Private nonprofit, less-than-4-year	1,150	99.8	1.00
Private nonprofit, 4-year, non-doctorate-granting	10,560	98.5	1.01
Private nonprofit, 4-year, doctorate-granting	13,900	98.4	1.01
Private for-profit, less-than-2-year	2,480	99.2	1.01
Private for-profit, 2-year	5,960	99.1	1.01
Private for-profit, 4-year	27,090	97.8	1.01
Bureau of Economic Analysis Code (Office of Business Economics [OBE]) region[1]			
New England	4,540	97.7	1.02
Mideast	18,270	98.1	1.01
Great Lakes	15,910	98.2	1.01
Plains	8,510	97.8	1.02
Southeast	29,430	98.1	1.01
Southwest	14,050	98.2	1.01
Rocky Mountains	5,820	98.0	1.01
Far West	17,270	96.7	1.02
Outlying areas	1,770	99.4	1.00
Institution total enrollment[2]			
2,557 or fewer	29,380	99.0	1.01
2,558–10,368	28,860	98.6	1.01
10,369–26,982	28,750	97.7	1.02
26,983 or more	28,580	97.1	1.02
Student type			
Undergraduate	91,650	97.7	1.02
Graduate (excluding doctoral-professional practice)	21,380	98.4	1.01
Doctoral-professional practice	2,540	99.0	1.01
Baccalaureate status (sampled)			
Yes	36,700	98.8	1.01
No	78,880	97.8	1.02
Age as of Dec 31, 2015			
15–23	50,480	98.2	1.01
24–29	27,610	98.0	1.01
30 or more	37,480	97.0	1.02
Sampled student type			
Undergraduate	93,590	97.7	1.01
Graduate (excluding doctoral-professional practice)	19,900	98.4	1.01
Doctoral-professional practice	2,080	98.9	1.01

See notes at end of table.

Table 53. Weight adjustment factors for student refusal adjustment: 2015–16—Continued

Model predictor variables	Number did not refuse	Weighted response rate	Average weight adjustment factor (WT9)
Veteran status			
Yes	10,120	97.0	1.02
No	105,450	97.9	1.01
Race/ethnicity			
White, non-Hispanic	58,670	98.0	1.01
Black, non-Hispanic	18,750	99.2	1.01
Hispanic	18,730	98.4	1.01
Asian, non-Hispanic	9,240	98.5	1.01
American Indian or Alaskan Native, non-Hispanic	770	98.3	1.01
Native Hawaiian or other Pacific Islander, non-Hispanic	500	98.0	1.01
More than one race, non-Hispanic	3,300	99.1	1.01
Unknown	5,620	88.0	1.07
Sex			
Male	49,290	97.2	1.02
Female	66,290	98.4	1.01
CPS record available			
Yes	80,080	99.7	1.00
No	35,500	94.2	1.04
CPS—Parents' highest education			
Middle school/junior high	3,010	99.8	1.00
High school	24,170	99.8	1.00
College or beyond	42,100	99.7	1.00
Unknown/missing	10,800	99.6	1.00
CPS—Has dependents			
Yes	22,270	99.7	1.00
No or Unknown	57,800	99.7	1.00
CPS—Marital status			
Single or Unknown	96,020	97.6	1.02
Married/remarried	14,370	99.7	1.00
Separated	1,600	99.9	1.00
Divorced/widowed	3,600	99.5	1.00
Social Security number available			
Yes	111,090	98.1	1.01
No	4,480	92.9	1.05
Any aid status			
Received	87,730	99.4	1.00
Did not receive	9,550	100.0	1.01
Unknown	18,300	91.7	1.07
Federal aid status			
Received	69,720	99.6	1.00
Did not receive	37,090	96.8	1.02
Unknown	8,770	91.7	1.07

See notes at end of table.

Table 53. Weight adjustment factors for student refusal adjustment: 2015–16—Continued

Model predictor variables	Number did not refuse	Weighted response rate	Average weight adjustment factor (WT9)
Pell Grant status[3]			
Did not receive	51,660	96.4	1.03
$1–$2,888	13,310	99.8	1.00
$2,889–$5,775	15,750	99.8	1.00
$5,775 or more	10,930	99.9	1.00
Direct Loan status[2]			
Did not receive	65,720	96.7	1.02
$1–$4,750	12,640	99.8	1.00
$4,751–$7,500	16,620	99.8	1.00
$7,501–$11,000	8,180	99.6	1.00
$11,001 or more	12,420	99.8	1.00
PLUS Loan status[2]			
Did not receive	87,090	97.6	1.02
$1–$6,000	1,180	100.0	1.00
$6,001–$11,174	1,100	99.5	1.00
$11,175–$18,000	1,150	99.8	1.00
$18,001 or more	1,140	99.8	1.00
Institutional aid status			
Received	32,900	99.2	1.01
Did not receive	63,720	98.0	1.01
Unknown	18,960	95.2	1.03
State aid status			
Received	14,900	99.6	1.00
Did not receive	81,370	98.1	1.01
Unknown	19,310	95.0	1.03
Telephone number count			
0	1,910	98.1	1.01
1	39,100	96.9	1.02
2	42,780	97.9	1.02
3 or more	31,790	99.0	1.01
E-mail address count			
0	1,210	90.0	1.07
1	29,680	94.8	1.03
2	61,410	98.6	1.01
3 or more	23,280	99.7	1.00
Mailing address count			
0	960	90.4	1.07
1	52,800	97.0	1.02
2	32,300	98.7	1.01
3 or more	29,510	99.1	1.01

See notes at end of table.

Table 53. Weight adjustment factors for student refusal adjustment: 2015–16—Continued

Model predictor variables	Number did not refuse	Weighted response rate	Average weight adjustment factor (WT9)
Student record complete indicator			
Complete data	107,770	98.1	1.01
Partial/No data	7,810	94.2	1.03
Mahalanobis Distance value[2]			
1.17–2.65	29,310	98.5	1.01
2.66–3.17	28,930	97.9	1.01
3.18–3.72	28,630	97.4	1.02
3.73 or more	28,700	97.5	1.02

[1] New England = Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont. Mideast = Delaware, District of Columbia, Maryland, New Jersey, New York, Pennsylvania. Great Lakes = Illinois, Indiana, Michigan, Ohio, Wisconsin. Plains = Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota. Southeast = Alabama, Arkansas, Florida, Georgia, Kentucky, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, West Virginia. Southwest = Arizona, New Mexico, Oklahoma, Texas. Rocky Mountains = Colorado, Idaho, Montana, Utah, Wyoming. Far West = Alaska, California, Hawaii, Nevada, Oregon, Washington. Outlying areas = Puerto Rico.
[2] Categories were defined by quartiles.
[3] Pell Grant categories for students receiving less than $5,775 in Pell Grants were defined by computing the median of all students receiving Pell Grants of less than $5,775, then all students receiving Pell Grants of $5,775 or more are in a single category.
NOTE: CPS = Central Processing System. PLUS = Parent Loan for Undergraduate Students. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

Student other nonresponse adjustment (WT10). The third, and final, stage of adjustment for student nonresponse was an adjustment for other nonresponse (e.g. contacted, but not interviewed before the end of the data collection period), given that the student was located and did not explicitly refuse to participate. NPSAS staff made this additional type of student nonresponse adjustment to compensate further for potential student nonresponse bias. As with WT8 and WT9, the same WTADJUST SUDAAN procedure and candidate predictor variables were used, and a CHAID analysis on the predictor variables was run to detect important interactions.

Table 54 shows the final predictor variables used in the model to determine weight adjustments and the average weight adjustment factor resulting from these variables. Summary statistics of the weight adjustment factors were

- minimum: 1.00;
- median: 1.00; and
- maximum: 8.18.

The final lower bound was 1.0 and the final upper bound was 50 for this weight adjustment.

Table 54. Weight adjustment factors for student other nonresponse adjustment: 2015–16

Model predictor variables	Number respondents	Weighted response rate	Average weight adjustment factor (WT10)
Total	**112,820**	**95.7**	**1.03**
Control and level of institution			
Public less-than-2-year	360	99.0	1.01
Public 2-year	15,650	94.8	1.05
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	5,090	94.7	1.04
Public 4-year, non-doctorate-granting, primarily baccalaureate	6,000	96.6	1.03
Public 4-year, doctorate-granting	25,620	95.6	1.04
Private nonprofit, less-than-4-year	1,130	99.5	1.00
Private nonprofit, 4-year, non-doctorate-granting	10,430	98.1	1.02
Private nonprofit, 4-year, doctorate-granting	13,720	97.4	1.02
Private for-profit, less-than-2-year	2,440	97.9	1.02
Private for-profit, 2-year	5,860	98.1	1.02
Private for-profit, 4-year	26,520	92.5	1.04
Bureau of Economic Analysis Code (Office of Business Economics [OBE]) region[1]			
New England	4,400	95.3	1.03
Mideast	17,870	96.5	1.03
Great Lakes	15,600	96.8	1.02
Plains	8,270	94.8	1.05
Southeast	28,740	95.6	1.03
Southwest	13,740	96.5	1.02
Rocky Mountains	5,730	96.9	1.02
Far West	16,720	93.8	1.04
Outlying areas	1,760	99.3	1.01
Institution total enrollment[2]			
2,557 or fewer	28,950	98.2	1.02
2,558–10,368	28,270	97.4	1.02
10,369–26,982	28,000	95.8	1.04
26,983 or more	27,610	93.7	1.05
Student type			
Undergraduate	89,220	95.4	1.03
Graduate (excluding doctoral-professional practice)	21,080	97.5	1.02
Doctoral-professional practice	2,520	98.6	1.01
Baccalaureate status (sampled)			
Yes	36,170	97.8	1.02
No	76,660	95.5	1.04
Age as of Dec 31, 2015			
15–23	49,280	96.0	1.03
24–29	27,030	95.8	1.03
30 or more	36,510	95.0	1.04

See notes at end of table.

Table 54. Weight adjustment factors for student other nonresponse adjustment: 2015–16—Continued

Model predictor variables	Number respondents	Weighted response rate	Average weight adjustment factor (WT10)
Sampled student type			
Undergraduate	91,140	95.4	1.03
Graduate (excluding doctoral-professional practice)	19,610	97.1	1.03
Doctoral-professional practice	2,070	99.0	1.01
Veteran status			
Yes	9,880	95.0	1.04
No	102,940	95.8	1.03
Sex			
Male	47,820	94.7	1.04
Female	65,000	96.5	1.03
CPS—Parents' highest education			
Middle school/junior high	3,010	100.0	1.01
High school	24,150	99.9	1.00
College or beyond	42,090	100.0	1.01
Unknown/missing	10,790	99.9	1.00
CPS—Has dependents			
Yes	22,260	100.0	1.00
No or Unknown	57,780	99.9	1.01
CPS—Marital status			
Single or Unknown	60,490	99.9	1.01
Married/remarried	14,360	100.0	1.00
Separated	1,600	100.0	1.00
Divorced/widowed	3,600	99.9	1.00
Social Security number available			
Yes	108,850	96.8	1.02
No	3,970	78.4	1.25
Federal aid status			
Received	69,540	99.6	1.00
Did not receive	35,800	94.3	1.05
Unknown	7,490	78.9	1.22
Pell Grant status			
Received	39,980	99.9	1.00
Did not receive	49,240	92.5	1.06
Direct Loan status			
Received	49,840	99.9	1.00
Did not receive	62,980	93.2	1.06
Institutional aid status			
Received	32,660	98.8	1.01
Did not receive	62,500	96.6	1.03
Unknown	17,660	87.7	1.09

See notes at end of table.

Table 54. Weight adjustment factors for student other nonresponse adjustment: 2015–16—Continued

Model Predictor Variables	Number respondents	Weighted response rate	Average weight adjustment factor (WT10)
State aid status			
Received	14,860	99.6	1.00
Did not receive	79,950	96.7	1.02
Unknown	18,020	87.6	1.09
Telephone number count			
0	1,790	89.1	1.09
1	37,670	93.1	1.05
2	41,930	96.8	1.02
3 or more	31,440	98.2	1.01
Mailing address count			
0	780	68.2	1.38
1	50,990	94.4	1.05
2	31,800	97.3	1.02
3 or more	29,240	98.3	1.01
Student record complete indicator			
Complete data	105,640	96.5	1.03
Partial/No data	7,180	83.8	1.12
Mahalanobis Distance value[2]			
1.17–2.65	28,900	97.8	1.02
2.66–3.17	28,290	95.8	1.03
3.18–3.72	27,750	94.6	1.04
3.73 or more	27,880	94.5	1.04

[1] New England = Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont. Mideast = Delaware, District of Columbia, Maryland, New Jersey, New York, Pennsylvania. Great Lakes = Illinois, Indiana, Michigan, Ohio, Wisconsin. Plains = Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota. Southeast = Alabama, Arkansas, Florida, Georgia, Kentucky, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, West Virginia. Southwest = Arizona, New Mexico, Oklahoma, Texas. Rocky Mountains = Colorado, Idaho, Montana, Utah, Wyoming. Far West = Alaska, California, Hawaii, Nevada, Oregon, Washington. Outlying areas = Puerto Rico.
[2] Categories were defined by quartiles.
NOTE: CPS = Central Processing System. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

Student poststratification adjustment (WT11). To ensure that the weighted student sample adequately represents the student population, NPSAS staff used SUDAAN WTADJUST to further adjust the student weights so that they sum to known population totals for key characteristics. This adjustment also helps increase the precision of the estimates for these key characteristics and any related characteristics. Control totals were established for the following:

- number of Direct Loan undergraduate student recipients, by subsidized/unsubsidized loan type by institution type;

- number of Direct Loan graduate student recipients, by unsubsidized loan type by institution type;
- total amount of Direct Loans disbursed to undergraduate students, by subsidized/unsubsidized loan type by institution type;
- total amount of Direct Loans disbursed to graduate students, by unsubsidized loan type by institution type;
- number of Pell Grant recipients;
- total amount of Pell Grants awarded to undergraduate students, by institution type;
- total amount of Parent PLUS loans disbursed to undergraduate students, by institution type;
- PLUS amounts disbursed to graduate students, by institution type;
- total amount of Federal Supplemental Educational Opportunity Grants (SEOG) disbursed, by institution type;
- fall undergraduate student enrollment, by institution type;
- fall graduate student enrollment, by institution type;
- full-year undergraduate student enrollment, by institution type;
- full-year graduate student enrollment, by institution type; and
- full-year student enrollment, by sex, within institution type.

Direct Loan, Pell Grant, PLUS, and SEOG control totals were obtained from the U.S. Department of Education's Federal Student Aid (FSA) Office. Direct Loans, for which there are several control totals, is the largest single student loan program—in terms of the number of students affected as well as the dollars involved. Therefore, having accurate control total data on Direct Loans by loan type (subsidized or unsubsidized), institution type, and level (undergraduate or graduate) is crucial for poststratification. Prior to NPSAS:08, the amounts used for poststratifying student weights were gross loan commitments—the amounts that schools and lenders expected to award to students based on their loan applications—collected by FSA. Since 2008, NPSAS staff have used net disbursements—the amounts that the students actually receive—for poststratification because they more accurately reflect the amounts that students are actually borrowing.

Fall and full-year enrollment counts come from the IPEDS 2015 Fall and 2015–16 12-month Enrollment Components (Preliminary) for the 2015–16 academic year.

Student enrollment control totals were determined using IPEDS data, which can be downloaded from the online IPEDS data center at http://nces.ed.gov/ipeds/datacenter/DataFiles.aspx. The IPEDS data files used to construct the student enrollment control totals included the following five files, as named by IPEDS:

- EF2015A: 2015 Fall Enrollments—Includes data on race/ethnicity, sex, attendance status, and level of student for Fall 2015.
- EFFY2016: 2015–16 12-month Enrollments—Includes 12-month unduplicated head count for 2015–16.
- HD2015: 2015–16 Institutional Characteristics Header—Includes directory information for 2015–16.
- IC2015: 2015–16 Institutional Characteristics—Includes data on educational offerings, organization, admissions, services, and athletic associations for 2105–16.
- IC2015_PY: 2015–16 Institutional Characteristics—Includes data on student charges by program (vocational programs) for 2015–16.

The HD2015, IC2015, and IC2015_PY files were used in determining which schools were in the NPSAS population of institutions and were also used to create the institution type variable. The EF2015A and EFFY2016 files were used to determine the enrollment totals for fall and full year, respectively. Because enrollment counts from IPEDS double count students enrolled in multiple institutions and the NPSAS weight should produce an unduplicated count of enrollment, the control totals used in the student poststratification adjustment were the IPEDS counts adjusted for student multiplicity. The counts were modified using the following general formula:

NPSAS control total = IPEDS enrollment total (mean student multiplicity)

Where *NPSAS control total, IPEDS enrollment total, and mean student multiplicity* are for the given institution control and level and student level of interest.

To determine full-year student enrollment, by sex, within institution type the formula used was

*NPSAS Female control total = NPSAS full-year control total * Proportion Female*

Where *Proportion female = IPEDS Female full-year enrollment total / IPEDS full-year enrollment total*

NPSAS Female control total, NPSAS full-year control total, Proportion Female, IPEDS Female full-year enrollment total, and *IPEDS full-year enrollment total* are all for the sector of interest, and all totals include undergraduate and graduate students.

Using the data on federal financial aid and institution enrollment, NPSAS staff poststratified weights to the control totals, truncating and smoothing high-extreme[34] weights.

Table 55 shows the variables associated with the weight adjustment factors for these variables. Summary statistics of the weight adjustment factors were

- minimum: 0.00;
- median: 1.08; and
- maximum: 156.52.

The final lower bound was 0.125 and the final upper bound was 100 for this weight adjustment.

[34] Extreme weights were identified as those greater than the median weight plus three times the interquartile range for weights.

Table 55. Weight adjustment factors for student poststratification: 2015–16

Poststratification categories	Control total	Average weight adjustment factor (WT11)
Fall undergraduate student enrollment, by control and level of institution		
Public, less-than-2-year	48,079	1.75
Public, 2-year	5,078,143	1.01
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	710,846	2.04
Public 4-year, non-doctorate-granting, primarily baccalaureate	965,940	1.02
Public 4-year, doctorate-granting	4,135,867	0.96
Private nonprofit, less-than-4-year	61,267	0.93
Private nonprofit, 4-year, non-doctorate-granting	1,058,177	0.98
Private nonprofit, 4-year, doctorate-granting	1,428,541	1.01
Private for-profit		
Private for-profit, less-than-2-year	214,287	1.16
Private for-profit, 2-year	274,702	1.02
Private for-profit, 4-year	686,664	2.51
Fall graduate student enrollment, by control and level of institution		
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	557	1.01
Public 4-year, non-doctorate-granting, primarily baccalaureate	102,947	1.22
Public 4-year, doctorate-granting	1,228,973	1.55
Private nonprofit, 4-year, non-doctorate-granting	149,793	1.23
Private nonprofit, 4-year, doctorate-granting	1,044,138	1.66
Private for-profit, 4-year	235,722	3.54
Full-year undergraduate student enrollment, by control and level of institution		
Public, less-than-2-year	66,147	1.87
Public, 2-year	7,639,516	1.25
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	997,100	2.33
Public 4-year, non-doctorate-granting, primarily baccalaureate	1,155,188	1.13
Public 4-year, doctorate-granting	4,686,013	1.01
Private nonprofit, less-than-4-year	98,448	1.32
Private nonprofit, 4-year, non-doctorate-granting	1,273,764	1.06
Private nonprofit, 4-year, doctorate-granting	1,679,005	1.08
Private for-profit, less-than-2-year	365,884	1.34
Private for-profit, 2-year	449,485	1.23
Private for-profit, 4-year	1,121,734	2.99
Full-year graduate student enrollment, by control and level of institution		
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	808	0.90
Public 4-year, non-doctorate-granting, primarily baccalaureate	163,203	1.46
Public 4-year, doctorate-granting	1,504,282	1.63
Private nonprofit, 4-year, non-doctorate-granting	223,225	1.51
Private nonprofit, 4-year, doctorate-granting	1,311,956	1.72
Private for-profit, 4-year	369,389	4.33

See notes at end of table.

Table 55. Weight adjustment factors for student poststratification: 2015–16—Continued

Poststratification categories	Control total	Average weight adjustment factor (WT11)
Full-year student enrollment, by sex, within control and level of institution		
Male		
Public less-than-2-year	30,948	1.97
Public 2-year	3,393,232	1.25
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	425,938	2.30
Public 4-year, non-doctorate-granting, primarily baccalaureate	553,240	1.10
Public 4-year, doctorate-granting	2,840,567	1.15
Private nonprofit, less-than-4-year	28,113	1.22
Private nonprofit, 4-year, non-doctorate-granting	621,335	1.21
Private nonprofit, 4-year, doctorate-granting	1,271,322	1.38
Private for-profit, less-than-2-year	95,436	1.26
Private for-profit, 2-year	163,374	1.12
Private for-profit, 4-year	530,407	3.28
Female		
Public less-than-2-year	35,199	1.79
Public 2-year	4,246,284	1.25
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	571,970	2.32
Public 4-year, non-doctorate-granting, primarily baccalaureate	765,151	1.28
Public 4-year, doctorate-granting	3,349,728	1.20
Private nonprofit, less-than-4-year	70,335	1.38
Private nonprofit, 4-year, non-doctorate-granting	875,654	1.13
Private nonprofit, 4-year, doctorate-granting	1,719,639	1.36
Private for-profit, less-than-2-year	270,448	1.37
Private for-profit, 2-year	286,111	1.29
Private for-profit, 4-year	960,716	3.38
Total Amount of Pell Grants awarded, by control and level of institution		
Public less-than-2-year	$80,926,767	1.20
Public 2-year	$8,526,934,650	1.13
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	$1,204,181,505	2.09
Public 4-year, non-doctorate-granting, primarily baccalaureate	$1,873,107,016	1.17
Public 4-year, doctorate-granting	$7,601,152,662	1.16
Private nonprofit, less-than-4-year	$255,749,623	1.36
Private nonprofit, 4-year, non-doctorate-granting	$2,101,014,759	1.14
Private nonprofit, 4-year, doctorate-granting	$2,222,235,897	1.17
Private for-profit, less-than-2-year	$934,366,406	1.40
Private for-profit, 2-year	$1,104,400,570	1.16
Private for-profit, 4-year	$2,621,119,431	3.14

See notes at end of table.

Table 55. Weight adjustment factors for student poststratification: 2015–16—Continued

Poststratification categories	Control total	Average weight adjustment factor (WT11)
Number of Stafford Loan undergraduate student recipients, by subsidized/unsubsidized loan type within control and level of institution		
Subsidized		
Public less-than-2-year	9,235	1.10
Public 2-year	942,740	0.95
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	148,049	1.91
Public 4-year, non-doctorate-granting, primarily baccalaureate	419,567	1.12
Public 4-year, doctorate-granting	2,032,218	1.18
Private nonprofit, less-than-4-year	56,167	1.10
Private nonprofit, 4-year, non-doctorate-granting	628,642	1.10
Private nonprofit, 4-year, doctorate-granting	783,986	1.18
Private for-profit, less-than-2-year	193,773	1.43
Private for-profit, 2-year	224,027	1.12
Private for-profit, 4-year	705,127	3.00
Unsubsidized		
Public less-than-2-year	9,390	1.12
Public 2-year	768,807	0.94
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	121,347	1.84
Public 4-year, non-doctorate-granting, primarily baccalaureate	409,424	1.15
Public 4-year, doctorate-granting	2,017,823	1.17
Private nonprofit, less-than-4-year	54,039	1.11
Private nonprofit, 4-year, non-doctorate-granting	636,027	1.10
Private nonprofit, 4-year, doctorate-granting	791,385	1.19
Private for-profit, less-than-2-year	190,975	1.41
Private for-profit, 2-year	217,164	1.12
Private for-profit, 4-year	698,484	3.23
Number of Unsubsidized Direct Loan graduate student recipients, by control and level of institution		
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	233	2.20
Public 4-year, non-doctorate-granting, primarily baccalaureate	46,928	1.42
Public 4-year, doctorate-granting	514,558	1.60
Private nonprofit, 4-year, non-doctorate-granting	88,082	1.42
Private nonprofit, 4-year, doctorate-granting	553,838	1.59
Private for-profit, 4-year	223,542	5.40

See notes at end of table.

Table 55. Weight adjustment factors for student poststratification: 2015–16—Continued

Poststratification categories	Control total	Average weight adjustment factor (WT11)
Total amount of total Direct Loans disbursed to undergraduate students, by subsidized/unsubsidized loan type		
Subsidized		
Public less-than-2-year	$28,004,751	1.10
Public 2-year	$2,766,426,741	0.95
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	$466,523,076	1.91
Public 4-year, non-doctorate-granting, primarily baccalaureate	$1,620,586,460	1.12
Public 4-year, doctorate-granting	$8,179,945,278	1.18
Private nonprofit, less-than-4-year	$187,441,312	1.10
Private nonprofit, 4-year, non-doctorate-granting	$2,487,596,713	1.10
Private nonprofit, 4-year, doctorate-granting	$3,172,877,568	1.18
Private for-profit, less-than-2-year	$559,796,377	1.43
Private for-profit, 2-year	$783,973,002	1.12
Private for-profit, 4-year	$2,614,965,286	3.00
Unsubsidized		
Public less-than-2-year	$41,697,980	1.12
Public 2-year	$2,661,362,100	0.94
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	$445,092,938	1.84
Public 4-year, non-doctorate-granting, primarily baccalaureate	$1,658,603,090	1.15
Public 4-year, doctorate-granting	$8,075,679,981	1.17
Private nonprofit, less-than-4-year	$219,458,953	1.11
Private nonprofit, 4-year, non-doctorate-granting	$2,507,542,231	1.10
Private nonprofit, 4-year, doctorate-granting	$3,109,119,495	1.19
Private for-profit, less-than-2-year	$757,645,776	1.41
Private for-profit, 2-year	$985,255,170	1.12
Private for-profit, 4-year	$3,469,753,774	3.23
Total amount of Unsubsidized Direct Loans disbursed to graduate students, by control and level of institution		
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	$2,557,188	2.20
Public 4-year, non-doctorate-granting, primarily baccalaureate	$614,653,384	1.42
Public 4-year, doctorate-granting	$9,420,113,341	1.60
Private nonprofit, 4-year, non-doctorate-granting	$1,235,664,473	1.42
Private nonprofit, 4-year, doctorate-granting	$11,247,435,486	1.59
Private for-profit, 4-year	$3,392,932,800	5.40

See notes at end of table.

Table 55. Weight adjustment factors for student poststratification: 2015–16—Continued

Poststratification categories	Control total	Average weight adjustment factor (WT11)
Total PLUS loan amounts disbursed to graduate students, by control and level of institution		
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	$110,751	1.98
Public 4-year, non-doctorate-granting, primarily baccalaureate	$23,732,526	2.00
Public 4-year, doctorate-granting	$2,049,171,076	1.58
Private nonprofit, 4-year, non-doctorate-granting	$204,521,081	1.76
Private nonprofit, 4-year, doctorate-granting	$5,167,176,796	1.69
Private for-profit, 4-year	$545,434,979	2.83
Total PLUS loan amounts disbursed to undergraduate students, by institution type		
Public less-than-2-year	$2,216,988	1.55
Public 2-year	$83,795,756	0.37
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	$72,663,914	1.86
Public 4-year, non-doctorate-granting, primarily baccalaureate	$667,077,000	1.06
Public 4-year, doctorate-granting	$5,130,760,177	1.11
Private nonprofit, less-than-4-year	$37,089,800	1.43
Private nonprofit, 4-year, non-doctorate-granting	$1,826,570,296	1.12
Private nonprofit, 4-year, doctorate-granting	$3,070,011,777	1.08
Private for-profit, less-than-2-year	$156,923,012	1.19
Private for-profit, 2-year	$243,530,349	1.08
Private for-profit, 4-year	$512,693,376	2.36
Number of SEOGs disbursed to undergraduate students		
Public less-than-2-year	664,251	0.91
Public 2-year	172,920,203	1.11
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	22,398,850	3.56
Public 4-year, non-doctorate-granting, primarily baccalaureate	46,852,925	1.37
Public 4-year, doctorate-granting	270,374,332	1.31
Private nonprofit, less-than-4-year	5,294,492	1.37
Private nonprofit, 4-year, non-doctorate-granting	135,495,931	1.22
Private nonprofit, 4-year, doctorate-granting	219,339,138	1.27
Private for-profit, less-than-2-year	14,735,826	1.57
Private for-profit, 2-year	23,724,330	1.17
Private for-profit, 4-year	81,074,600	2.58
Total number of Pell Grants awarded	7,645,451	1.73

NOTE: PLUS = Parent Loan for Undergraduate Students. SEOG = Federal Supplemental Educational Opportunity Grants.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

After poststratification was performed, NPSAS staff computed the final student weight (WTA000) as the product of the 11 weight components described in this section. NPSAS staff compared weighted estimates for key variables from these NPSAS:16 data with estimates from other sources, such as estimates from NPSAS:12, FSA, and the Veterans Benefits Administration (VBA), and found the NPSAS:16 estimates to be reasonable, taking into account differences in timeframe, population, and other factors that would explain differences.

7.1.5 Weighting Adjustment Summary and Evaluation

Institution weighting adjustment summary and evaluation. Table 56 summarizes the institution weight distributions and the variance inflation caused by unequal weighting (referred to as the unequal weighting effect, or UWE) by control and level of institution. The UWE can be inflated by weight adjustments, in addition to unequal probabilities of selection during sampling. The UWE is defined as follows:

$$UWE = n\sum w^2 / \left(\sum w\right)^2$$

where

n = the sample size
w = the final student weight

The median institution weights range from 1.06 for public 4-year, doctorate-granting institutions to 19.27 for private for-profit, less-than-2-year institutions. The mean institution weight ranges from 1.12 for public 4-year, doctorate-granting institutions to 28.00 for private for-profit, less-than-2-year institutions. The UWE is 6.52 overall and ranges from 1.02 for public 4-year, doctorate-granting institutions to 6.01 for public 2-year institutions, with most UWEs less than 4.00. This means that for most institution types, the inflation on the variance of estimates due to the unequal weighting is relatively small and even for the institution types with higher UWEs, there is little concern on the effects it could have on estimation. The sample design and sample sizes accounted for UWEs in this range to ensure precision of estimates.

Table 56. Institution weight distribution and unequal weighting effects, by control and level of institution: 2015–16

Control and level of Institution	Minimum	First quartile	Median	Third quartile	Maximum	Mean	Unequal weighting effect[1]
Total	**0.97**	**1.09**	**1.45**	**2.95**	**143.01**	**3.90**	**6.52**
Public							
Less-than-2-year	1.19	3.95	6.80	17.77	51.20	12.97	2.30
2-year	1.01	1.21	1.83	2.95	112.36	3.26	6.01
4-year, non-doctorate-granting, primarily subbaccalaureate	0.99	1.04	1.25	2.11	9.87	1.79	1.59
4-year, non-doctorate-granting, primarily baccalaureate	0.97	1.05	1.28	3.50	28.03	2.74	2.62
4-year, doctorate-granting	1.00	1.03	1.06	1.14	3.01	1.12	1.02
Private nonprofit							
Less-than-4-year	1.18	5.49	9.52	13.37	111.32	17.33	3.06
4-year, non-doctorate-granting	1.02	1.29	1.98	3.09	65.92	3.38	3.77
4-year, doctorate-granting	1.02	1.11	1.42	2.32	30.25	2.42	2.77
Private for-profit							
Less-than-2-year	1.33	9.48	19.27	35.85	143.01	28.00	1.94
2-year	1.02	2.93	4.87	8.28	109.67	8.99	3.35
4-year	1.02	1.09	1.51	2.97	33.60	2.68	2.60

[1] Unequal weighting effect calculated as $UWE = n\sum w^2/(\sum w)^2$,., where *n* = the sample size and *w* = the final student weight.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

To assess the overall predictive ability of the model used to adjust for institution nonresponse , NPSAS staff used a Receiver Operating Characteristic (ROC) curve (Hanley and McNeil 1982). The ROC curve provides a measure of how well the model correctly classified individuals of known response type—in other words, how well the model predicts an institution's response propensity.[35] NPSAS staff developed the ROC curve in the following manner. The predicted probabilities of response (c) are derived from the model used to adjust for institution nonresponse. For any specified probability of response, c, two proportions were calculated

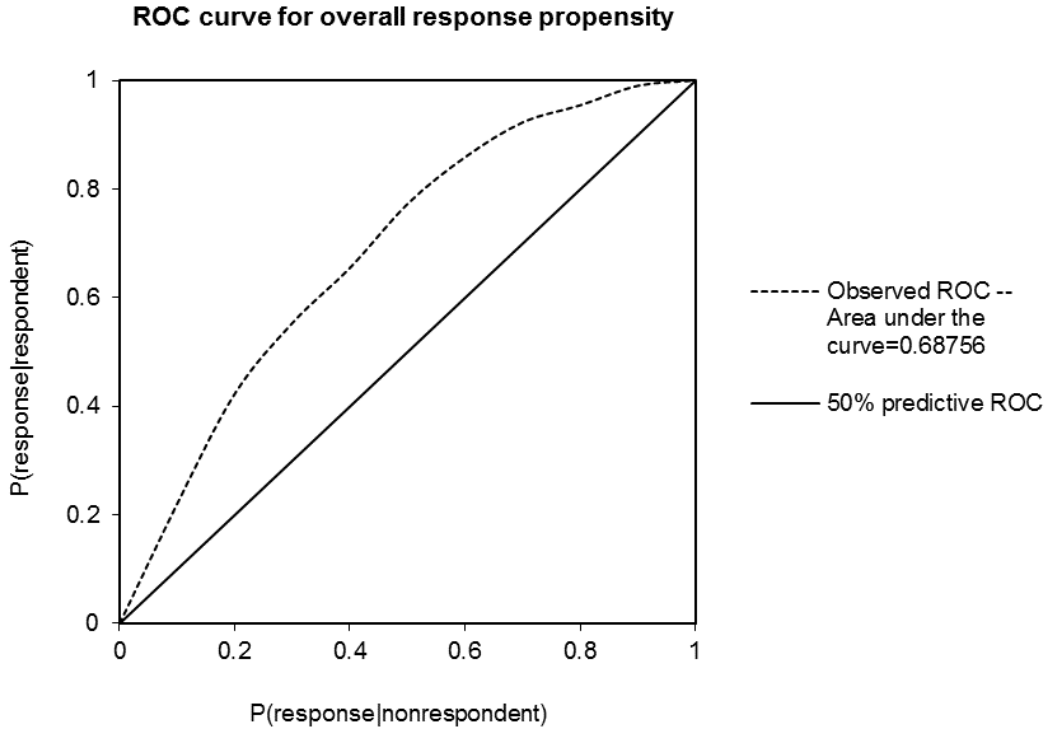
- the proportion of respondents with a predicted probability of response greater than c, and
- the proportion of nonrespondents with a predicted probability of response greater than c.

The plot of the first probability against the second, for c from zero to 1, resulted in the ROC curve shown in figure 9. The area under the curve equals the probability that the fitted model correctly classifies two randomly chosen institutions—one of

[35] For a more detailed example of the ROC curve used in nonresponse modeling, see Iannacchione (2003).

whom is a true respondent, while the other is a true nonrespondent—where the institution with the higher predicted probability of response is classified as the respondent. An area of 0.5 under an ROC curve indicates that a correct classification is made 50 percent of the time, with the model providing no predictive benefit. An area of 1.0 indicates that the true respondent always has the higher predicted probability of response as compared to the true nonrespondent, so the model always classifies the two institutions correctly. In Figure 9, the area under the ROC curve is approximately 0.69, so the predicted probabilities give the correct classification about 69 percent of the time. Researchers can also interpret predictive probabilities from ROC curves in terms of the nonparametric Wilcoxon test statistic, which is used to determine if the level of a quantitative variable, such as predicted probability of response) is different between two samples (respondents and nonrespondents in this case). The ROC area equals the value of the Wilcoxon test statistic. Viewed in this way, the Wilcoxon test rejects the null hypothesis of no predictive ability by showing that the predicted probability of response for the respondents is larger than that for the nonrespondents. Analysts can interpret this result to mean that the variables used in the model are highly informative predictors of a sample institution's overall response propensity.

Figure 9. Receiver Operating Characteristic (ROC) curve for overall institution response propensity: 2015–16



SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

Student weighting adjustment summary and evaluation. Table 57 summarizes the student weight distributions and the variance inflation caused by the UWE, by type of institution and student type. The median student weight ranges from 17 for students in private for-profit, 4-year institutions to 429 for students in public 2-year institutions. The mean student weight ranges from 56 for students in private for-profit, 4-year institutions to 488 for students in public 2-year institutions.

The UWE is 2.8 overall and ranges from 1.27 for students in for-profit less than 2-year institutions to 5.39 for students in private for-profit, 4-year institutions. This means that for students in most institution types, the inflation on the variance of estimates due to the unequal weighting is relatively small and even for the institution types with the higher UWEs there is little concern on the effects it could have on estimation. The sample design and sample sizes accounted for UWEs in this range to ensure precision of estimates.

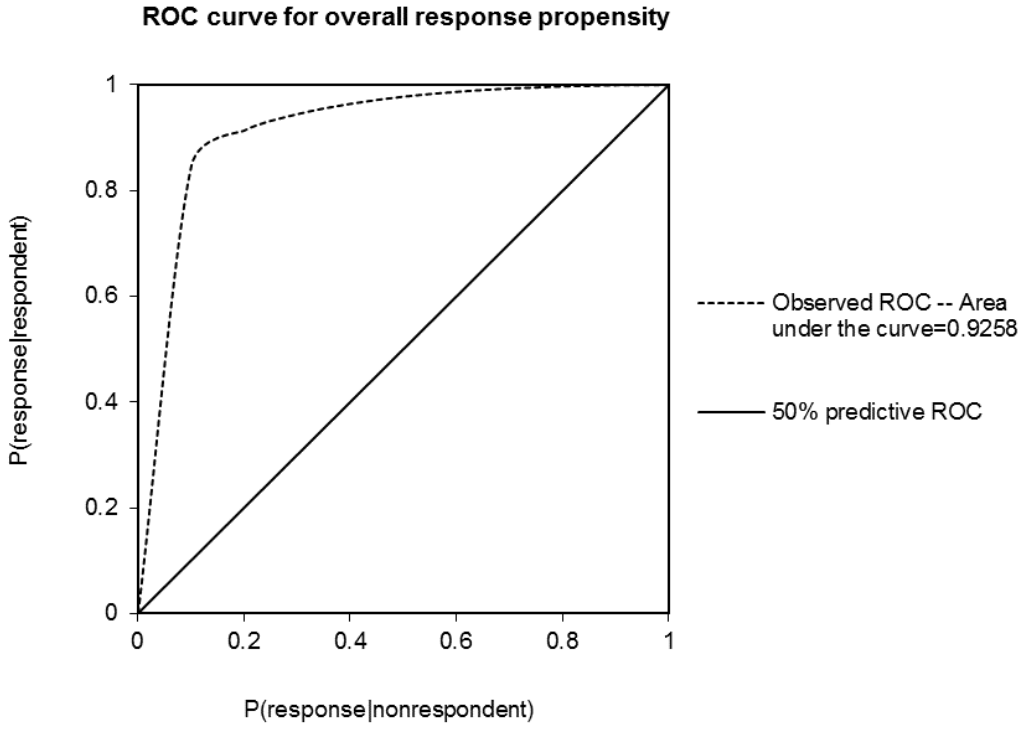
Table 57. Student weight distribution and unequal weighting effects: 2015–16

Analysis domain	Minimum	First quartile	Median	Third quartile	Maximum	Mean	Unequal weighting effect[1]
Total	**1.04**	**23.82**	**91.54**	**295.50**	**3608.34**	**204.79**	**2.80**
Control and level of institution							
Public							
Less-than-2-year	3.86	70.53	139.26	290.21	730.47	184.25	1.56
2-year	6.05	269.81	428.93	555.01	2867.84	488.27	1.49
4-year, non-doctorate-granting, primarily subbaccalaureate	1.04	9.84	34.49	282.71	2481.74	196.09	3.36
4-year, non-doctorate-granting, primarily baccalaureate	5.92	49.45	123.98	301.99	2985.34	219.59	2.38
4-year, doctorate-granting	6.14	68.05	140.94	352.70	2755.62	241.60	2.16
Private nonprofit							
Less-than-4-year	4.27	27.24	69.49	102.31	671.19	86.97	2.10
4-year, non-doctorate-granting	5.75	25.54	61.69	210.93	1676.22	143.57	2.47
4-year, doctorate-granting	6.77	28.61	76.09	302.45	3608.34	217.98	2.98
For-profit							
Less-than-2-year	6.34	94.79	140.25	194.25	601.36	149.95	1.27
2-year	6.24	40.31	65.68	95.74	1142.93	76.72	1.55
4-year	1.46	8.83	16.75	50.09	1854.38	56.22	5.39
Derived student type							
Undergraduate	1.06	31.01	106.76	330.46	2985.34	218.93	2.57
Graduate (excluding doctoral-professional practice)	1.04	14.44	43.09	136.01	3608.34	150.38	4.27
Doctoral-professional practice	1.46	12.15	54.14	208.77	2062.12	159.65	3.25

[1] Unequal weighting effect calculated as sample size multiplied by the sum of the squared weights, divided by the sum of the weights squared.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

To assess the overall predictive ability of the student nonresponse model, an ROC curve was developed as described in the previous section. However, the predicted probabilities of response (c) for the ROC curve associated with the student nonresponse are the product of the predicted response probabilities obtained at each of the three nonresponse adjustment steps. Note that for the second and third nonresponse adjustments (refusal and other nonresponse adjustments) predicted probabilities were calculated for all nonrespondents, but the models were developed excluding those students who had dropped out in the prior nonresponse adjustment. Figure 10 shows that the area under the ROC curve is approximately 0.93, so the predicted probabilities give the correct classification about 93 percent of the time. Predictive probabilities from ROC curves can also be interpreted in terms of the nonparametric Wilcoxon test statistic, where the ROC area is equivalent to the value of the Wilcoxon test statistic. Viewed in this way, the Wilcoxon test rejects the null hypothesis of no predictive ability by showing that the predicted probability of response for the respondents is larger than that for the nonrespondents. This result can be interpreted to mean that the variables used in the model are definitive predictors of a sample student's overall response propensity.

Figure 10. Receiver Operating Characteristic (ROC) curve for overall student response propensity: 2015–16



SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

7.2 Nonresponse Bias Analysis

The accuracy of survey statistics is affected by both random and nonrandom errors. Random errors reduce the precision of survey estimates, while nonrandom errors may result in bias (i.e., estimates that do not converge to the true population parameter as the sample size increases without limit) or loss of precision.

The sources of error in a survey are often dichotomized as sampling and nonsampling errors. Sampling error refers to the error that occurs because the survey is based on a sample of population members rather than the entire population. All other types of errors are nonsampling errors, including survey nonresponse (because of inability to contact sampling members, their refusal to participate in the study, etc.) and measurement errors, such as the errors that occur because the intent of survey questions was not clear to the respondent, because the respondent had insufficient knowledge to answer correctly, or because the data were not captured correctly (e.g., because of recording, editing, or data entry errors).

Sampling errors are primarily random errors for well-designed surveys such as NPSAS:16. However, nonrandom errors can occur if the sampling frame does not provide complete coverage of the target population. The NPSAS:16 survey instrument and data collection procedures were subjected to thorough development and testing to minimize nonsampling errors because these errors are difficult to quantify and are likely to be nonrandom errors.

In this section, nonsampling error is observed by comparing NPSAS:16 nonrespondents and respondents using characteristics known for both groups. Section 7.3 discusses measurement of sampling error by variance estimation.

NCES Statistical Standard 4-4-1 states that "Any survey stage of data collection with a unit or item response rate less than 85 percent must be evaluated for the potential magnitude of nonresponse bias before the data or any analysis using the data may be released. Estimates of survey characteristics for nonrespondents and respondents are required to assess the potential nonresponse bias" (Seastrom 2014).

The bias in an estimated mean based on respondents $\bar{y}_R$ is the difference between the expected value of this mean and the target parameter, π, the population mean. Analysts can estimate the population mean for characteristics that are observed for both respondents and nonrespondents with the full-sample mean, which can be expressed in terms of the respondent mean and nonrespondent mean, $\bar{y}_{NR}$, as follows: $\hat{\pi} = (1 - \eta)\ \bar{y}_R + \eta \bar{y}_{NR}$, where η is the weighted unit (or item) nonresponse rate. For variables that are from the frame rather than from the sample,

analysts can estimate π without sampling error. They can then estimate bias as the difference between the respondent mean and the full sample mean: $\hat{B}(\bar{y}_R) = \bar{y}_R - \hat{\pi}$. Equivalently, bias can be estimated as the difference between the mean for respondents and the mean for nonrespondents, multiplied by the weighted nonresponse rate: $\hat{B}(\bar{y}_R) = \eta(\bar{y}_R - \bar{y}_{NR})$.

Relative bias provides a measure of the magnitude of the bias relative to the sample mean and is estimated as: $\widehat{RB}(\bar{y}_R) = \hat{B}(\bar{y}_R)/\hat{\pi}$. Effect size, as defined by Cohen (1988), is another measure of potential nonresponse bias. For continuous variables, it is computed as the estimated bias divided by the estimated standard deviation: $\hat{B}(\bar{y}_R)/\hat{\sigma}_y$. For categorical variables, it is computed as $\sqrt{\sum_i (p_{0i} - p_{1i})^2/p_{0i}}$, where p_{0i} is the proportion of the full sample in category i, and p_{1i} is the proportion of respondents in category i. Effect sizes can be used in combination with bias and relative bias estimates and significance tests to evaluate the potential for nonresponse bias. Cohen classified an effect size as "small" when it is about 0.10, as "medium" when it is about 0.30, and as "large" when it is about 0.50.

NPSAS staff conducted unit nonresponse bias analyses at the institution, study member, and interview levels, and item nonresponse analyses among study members, for the overall sample and by institutional sector, which was used for institution stratification (described in chapter 2). These analyses are described in the sections below. The unit-level results are summarized in tables 58 through 61, and detailed tables are provided in appendix J. The item-level response rates are shown in table 62, and bias results are summarized in appendix J.

7.2.1 Bias Analysis: Institution Level

An institution respondent is defined as any sampled institution that provided a student enrollment list from which a student sample was selected.

As shown in table 3, about 1,750 of the 1,990 eligible sample institutions were respondents (88 percent unweighted and 90 percent weighted). The weighted response rates, by control and level of institution, range from 74 percent for private for-profit, less-than-2-year institutions to 95 percent for public 4-year, non-doctorate-granting, primarily subbaccalaureate institutions. The institution-weighted response rate is below 85 percent for 3 of the 11 types of institutions: public less-than-2-year; private for-profit, less-than-2-year; and private for-profit, 2-year institutions.

A nonresponse bias analysis was conducted overall and for each institutional sector, regardless of response rate. Because all sectors were included in the nonresponse

weight adjustments (discussed in Section 7.1), the bias analysis can evaluate how well the weight adjustments reduce nonresponse bias for sectors both above and below an 85 percent institution response rate. Nonresponse bias was estimated for characteristics known for most respondents and nonrespondents, including some characteristics not included in the nonresponse weight adjustments, in order to evaluate how well the weight adjustments worked for variables not used in the weight adjustment model. There are extensive data available from IPEDS for all institutions for the 2014–15 and 2015–16 academic years. The following variables[36] were used for the nonresponse bias analysis:

- control and level of institution;[37]
- 2015 Carnegie Basic classification;
- degree of urbanization;[38]
- institution region;
- HBCU status;
- HSI status;[39]
- percentage of full-time, first-time degree/certificate-seeking students receiving federal grant aid;
- percentage of full-time, first-time degree/certificate-seeking students receiving state or local grant aid;
- percentage of full-time, first-time degree/certificate-seeking students receiving institutional grant aid;
- percentage of full-time, first-time degree/certificate-seeking students receiving student loan aid;
- average net price among full-time, first-time degree/certificate-seeking students receiving grant or scholarship aid;
- percentage of students enrolled who were Hispanic;

[36] For the continuous variables, categories were formed based on quartiles.
[37] Institution control and level were only used for the overall sample nonresponse bias analysis.
[38] Degree of urbanization is an IPEDS variable representing the urbanicity (city/suburb/rural) by population size of the institution's location.
[39] Of the listed variables, only the HSI indicator no longer exists in IPEDS. An HSI proxy was created following the definition of HSI as provided by the US Department of Education https://www2.ed.gov/programs/idueshsi/definition.html and using IPEDS Hispanic enrollment data.

- percentage of students enrolled who were Asian or Pacific Islander, non-Hispanic;[40]
- percentage of students enrolled who were Black, non-Hispanic;
- total undergraduate enrollment;
- male undergraduate enrollment;
- female undergraduate enrollment;
- total graduate enrollment;
- male graduate enrollment;
- female graduate enrollment;
- number of bachelor's degrees awarded;
- percentage of full-time, first time degree/certificate-seeking undergraduate students who received any grant aid;
- number of first-time, full-time undergraduate students living on-campus;
- average amount of grant and scholarship aid received; and
- number of full-time, first-time undergraduate students with incomes up to $30,000 who were receiving Title IV aid.

First, for the institution-level variables listed above, the nonresponse bias was estimated for each category as the difference between the weighted means (proportions) of the respondents and the full sample. Each estimated nonresponse bias was tested using a *t*-test to determine if it significantly differed from zero at the 5 percent level. Relative bias was computed as the ratio of the estimated bias to the weighted full-sample mean. Second, nonresponse adjustments were computed to reduce or eliminate nonresponse bias for key variables. Third, using the weights adjusted for nonresponse, the re-estimated nonresponse bias was tested for significance. These tests were complemented by effect size calculations. Finally, to better understand the effect of poststratification on efforts to reduce nonresponse bias, two additional sets of estimates were created. The first set of estimates equals the difference in respondent weighted means before and after poststratification, which corresponds to the effect of poststratification on nonresponse adjustments. The second set of estimates, equal to the difference between base-weighted full-sample means and the poststratified respondent means, corresponds to the cumulative effects of all weighting and adjustment steps.

[40] Asian or Pacific Islander, non-Hispanic includes Native Hawaiian.

As shown in table 58, the institution nonresponse weighting adjustment eliminated some, but not all, significant bias on the observable characteristics for sectors that met reporting requirements (excluding characteristics for sectors with fewer than 5 nonresponding institutions). Before weighting, the median effect size for all institutions was 0.08, ranging from 0.02 for public 4-year, doctorate-granting institutions to 0.27 for public 4-year, non-doctorate-granting, primarily baccalaureate institutions. The percentage of variable categories that were significantly biased was 35 percent overall, and ranged from 0 percent for four institutional sectors to 10 percent for private for-profit, 4-year institutions. When looking at institutional sectors individually, many of the categories for which overall differences are statistically significant do not have enough institutions to detect differences or even meet reporting standards and thus there are fewer significant biases at the sector level. See tables J-1 through J-72 for detailed results by institutional sector. After the nonresponse weight adjustment, the median effect size for all institutions was 0.05, ranging from 0.05 for public 4-year, doctorate-granting institutions to 0.27 for public 4-year, non-doctorate-granting, primarily baccalaureate institutions. The percentage of variable categories that remained significantly biased was 4 percent overall and ranged from 0 percent for three institutional sectors to 8 percent for private nonprofit, 4-year, doctorate granting institutions. The category with the largest absolute bias (18.8) after weight adjustment was percent receiving no or unknown state/local grant aid for private for-profit, 2-year institutions.

As shown in table 59, the mean absolute difference between means for respondents before and after poststratification adjustment ranged from zero for public 2-year institutions to 2.6 for private for-profit, less-than-2-year institutions. Similarly, the median absolute difference between means for respondents before and after poststratification adjustment ranged from zero for public 2-year institutions to 2.2 for private for-profit, less-than-2-year institutions. (Estimates for sectors with fewer than five nonresponding institutions were excluded.) The absolute difference between means for the full sample and respondents after poststratification adjustment ranged from 0.6 for public 4-year, doctorate-granting institutions to 6.6 for public 4-year, non-doctorate-granting, primarily baccalaureate institutions; and the median absolute difference ranged from 0.5 for public 4-year, doctorate-granting institutions to 3.8 for private for-profit, 2-year institutions.

Table 58. Summary of institution-level nonresponse bias analysis, by control and level of institution: 2015–16

Nonresponse bias statistics[1]	Overall	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting primarily subbaccalaureate	Public 4-year non-doctorate-granting primarily baccalaureate	Public 4-year doctorate-granting	Private nonprofit less-than-4-year	Private nonprofit 4-year non-doctorate-granting	Private nonprofit 4-year doctorate-granting	Private for-profit less-than-2-year	Private for-profit 2-year	Private for-profit 4-year
Before nonresponse weight adjustments[2]												
Mean percent relative bias across characteristics	8.61	‡	9.44	1.07	18.46	1.47	‡	8.82	8.25	13.2	13.84	6.88
Median percent relative bias across characteristics	7.14	‡	8.13	0.94	7.71	1.20	‡	7.16	8.45	9.44	12.53	5.69
Percentage of characteristics with significant bias	35.16	‡	2.86	#	#	#	‡	1.01	1.35	#	6.06	9.86
Median effect size	0.08	‡	0.10	0.06	0.27	0.02	‡	0.11	0.14	0.15	0.23	0.12
After nonresponse weight adjustments[3]												
Mean percent relative bias across characteristics	5.18	‡	9.11	1.59	17.60	2.36	‡	7.65	7.66	13.64	11.99	6.78
Median percent relative bias across characteristics	3.36	‡	6.98	1.44	6.68	1.90	‡	4.55	4.71	9.81	15.05	5.50
Percentage of characteristics with significant bias	3.91	‡	1.43	#	#	6.94	‡	1.01	8.11	#	3.03	2.82
Median effect size	0.05	‡	0.12	0.08	0.27	0.05	‡	0.09	0.15	0.15	0.19	0.11

‡ Reporting standards not met (fewer than five unweighted nonrespondents).
# Rounds to zero.
[1] Relative bias and effect size are calculated using the weighted differences between respondent and full-sample means. Relative bias is calculated as 100 times the ratio of estimated bias to the weighted full-sample mean. Effect size is calculated as the square root of the sum over categories of the squared differences over full-sample means.
[2] Respondent and full-sample means are weighted using the institution base weight.
[3] Full-sample means are weighted using the institution base weight and the respondent means are weighted using the institution base weight adjusted for nonresponse.
NOTE: Characteristics that did not meet reporting standards were excluded from calculation of summary statistics. “Base weight” refers to the institution sampling weight adjusted for field-test sampling.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

Table 59. Summary of institution-level differences between means, by control and level of institution: 2015–16

Summary statistics	Overall	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	Private for-profit 4-year
Difference between means for respondents before and after poststratification adjustment[1]												
Mean absolute difference across characteristics	0.56	‡	#	1.70	1.28	0.05	‡	0.95	1.01	2.64	1.12	0.18
Median absolute difference across characteristics	0.35	‡	#	1.98	0.67	0.04	‡	0.76	0.92	2.24	0.97	0.14
Difference between means for full sample and respondents after poststratification adjustment[2]												
Mean absolute difference across characteristics	0.98	‡	2.10	2.57	6.61	0.60	‡	1.76	2.27	4.45	3.85	1.45
Median absolute difference across characteristics	0.66	‡	1.60	3.12	2.32	0.50	‡	1.51	1.70	3.45	3.76	1.44

‡Reporting standards not met (fewer than five unweighted nonrespondents).
Rounds to zero.
[1] Respondent means before poststratification adjustment are weighted using the institution base weight adjusted for nonresponse. Respondent means after poststratification adjustment are weighted using the institution base weight adjusted for nonresponse and poststratification.
[2] Full-sample means are weighted using the institution base weight and respondent means are weighted using the institution base weight adjusted for nonresponse and poststratification.
NOTE: Characteristics that did not meet reporting standards were excluded from calculation of summary statistics. “Base weight” refers to the institution sampling weight adjusted for field-test sampling.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

7.2.2 Bias Analyses: Study Member Level and Student Interview Level

As described in section 4.5, a study member is defined as any student sample member who is determined eligible for the study and has valid data from any source for a selected set of key analytical variables. These were the minimal data requirements, but the vast majority of study members had considerably more complete data.

As discussed in section 4.5 and shown in table 35, of the 119,550 eligible students, the weighted and unweighted study membership rates were 93 and 94 percent, respectively. The weighted study membership rates, by control and level of institution, ranged from 91 to 99 percent.

A study member-level nonresponse bias analysis is not required per the NCES Statistical Standards, given the rates of study membership; however, a nonresponse bias analysis was still conducted to provide context alongside the institution-, student interview-, and item-level nonresponse bias analyses. Using the procedure described above, these analyses were conducted overall and within each institutional sector. The nonresponse bias was estimated for characteristics known for most respondents and nonrespondents, including some characteristics not included in the nonresponse weight adjustments (described in Section 7.1) in order to evaluate how well the weight adjustments worked for variables not used in the weight adjustment model. Bias estimates for characteristics that do not meet reporting requirements because they have fewer than 30 student nonrespondents were excluded from calculations of summary statistics. The following student-level variables[41] were used for the nonresponse bias analysis:

- institutional control and level;[42]
- institutional region;
- student type (undergraduate or graduate);
- sampled baccalaureate recipient status (baccalaureate/not baccalaureate);
- student age as of December 31, 2015;
- major (2-digit CIP code);
- degree program (undergraduates only);

[41] For the continuous variables, categories were formed based on quartiles. Institution-level variables come from 2014–15 and 2015–16 IPEDS and student-level variables from NPSAS:16 institution enrollment lists and CPS for aid applicants.
[42] Institution control and level were only used for the overall sample nonresponse bias analysis.

- parent's education (from CPS for aid applicants);
- marital status (from CPS for aid applicants);
- support children (from CPS for aid applicants);
- income (from CPS for aid applicants);
- federal aid receipt status (yes/no/don't know);[43]
- federal Pell Grant recipient (yes/no/don't know);[43]
- Direct Loan recipient (yes/no/don't know);[43]
- institutional aid recipient (yes/no);
- state aid recipient (yes/no);
- federal Pell Grant amount;
- Direct Loan amount;
- institution enrollment;
- institution percentage of undergraduates who received any grant aid;
- veteran status;
- race;
- ethnicity;
- sex;
- institution percentage of full-time, first-time degree/certificate-seeking undergraduate students who received any grant aid at institution attended;
- institution number of first-time, full-time undergraduate students living on campus of institution attended;
- institution average amount of grant and scholarship aid received at institution attended; and
- institution number of full-time, first-time undergraduate students with incomes up to $30,000 who were receiving Title IV aid at institution attended.

As shown in table 60, the student nonresponse weighting adjustment eliminated some, but not all, significant bias on observable characteristics for sectors that met reporting requirements (have at least 30 nonstudy members). Before weighting, the

[43] The "don't know" category for federal aid captures sample members for whom we do not have a Social Security number (SSN).

median effect size for all institutions was 0.02, ranging from 0.01 for students in private nonprofit, less-than-4-year institutions to 0.06 for students in private for-profit, 4-year institutions. The percentage of characteristics that were significantly biased for study members was 58 percent overall, and ranged from 0 percent for students in private nonprofit, less-than-4-year institutions to 52 percent for students in public 2-year institutions. After the nonresponse weight adjustment, the median effect size for all study members was zero, ranging from zero for students in private nonprofit, less-than-4-year institutions to 0.04 for students in private for-profit 4-year institutions. The percentage of characteristics that remained significantly biased was 19 percent overall and ranged from 1 percent for students in private for-profit, 4-year institutions to 21 percent for students in public 2-year institutions.

Overall, significant bias remained in one category of the variables Age, Support Children, Marital Status, Parent Education, and Sex; two categories of Income; three categories of Undergraduate Degree Program; four categories of Race; and seven categories of Major. Absolute significant bias for these categories ranged from 0.04 for the third quartile of Income to 1.44 for unknown race. Four sectors had only one characteristic with significant bias while public 2-year and public 4-year, doctorate granting institutions each had 14 and 15 characteristics, respectively. Most variables had between one and three significant bias estimates across all sectors and characteristics; however, Major had seven significant bias estimates across four sectors and Race had fifteen across eight sectors.

Table 60. Summary of student-level bias analysis, by control and level of institution: 2015–16

Nonresponse bias statistics[1]	Overall	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	Private for-profit 4-year
Before nonresponse weight adjustments—study member[2]												
Mean percent relative bias across characteristics	3.94	‡	4.78	4.36	3.87	3.43	0.99	3.18	3.04	5.20	2.91	6.54
Median percent relative bias across characteristics	1.64	‡	2.04	2.41	1.89	1.45	0.90	1.40	1.51	2.12	1.12	4.98
Percentage of characteristics with significant bias	58.18	‡	52.24	18.87	38.89	44.87	#	36.84	43.86	31.82	27.08	16.44
Median effect size	0.02	‡	0.04	0.03	0.03	0.02	0.01	0.02	0.03	0.02	0.03	0.06
Before nonresponse weight adjustments—interview[2]												
Mean percent relative bias across characteristics	7.69	7.58	7.86	9.39	6.31	5.76	8.36	6.20	6.33	9.17	7.67	10.53
Median percent relative bias across characteristics	5.01	7.45	3.15	6.38	3.96	2.67	6.65	4.02	3.70	6.98	5.14	7.03
Percentage of characteristics with significant bias	70.34	6.38	46.39	28.26	29.90	47.12	16.67	44.00	33.00	11.39	31.40	37.50
Median effect size	0.07	0.11	0.04	0.06	0.05	0.03	0.09	0.05	0.04	0.08	0.08	0.07
After nonresponse weight adjustments—study member[3]												
Mean percent relative bias across characteristics	1.84	‡	2.89	2.62	1.97	1.79	0.43	1.86	1.56	2.28	1.91	3.88
Median percent relative bias across characteristics	0.15	‡	0.72	1.69	0.94	0.74	0.20	0.59	0.89	0.40	1.05	1.51
Percentage of characteristics with significant bias	19.09	‡	20.90	13.21	12.96	19.23	9.09	5.26	8.77	4.55	2.08	1.37
Median effect size	#	‡	0.01	0.02	0.01	0.01	#	0.02	0.01	0.01	0.02	0.04

‡Reporting standards not met (fewer than 30 unweighted nonrespondents).
Rounds to zero.
[1] Relative bias and effect size are calculated using the weighted differences between respondent and full-sample means. Relative bias is calculated as 100 times the ratio of estimated bias to the weighted full-sample mean. Effect size is calculated as the square root of the sum over categories of the squared differences over full-sample means.
[2] Respondent and full-sample means are weighted using the student base weight.
[3] Full-sample means are weighted using the student base weight and the respondent means are weighted using the student base weight adjusted for nonresponse.
NOTE: Characteristics that did not meet reporting standards were excluded from calculation of summary statistics. “Base weight" refers to the student sampling weight (final institution weight times student sampling adjustment) adjusted for student multiplicity and unknown eligibility.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

As shown in table 61, the absolute differences between means for respondents before and after poststratification adjustment ranged from 0.7 for students in public 2-year institutions and private nonprofit, 4-year, non-doctorate-granting institutions to 3.9 for students in private for-profit, 4-year institutions, while the median difference ranged from 0.2 for students in public 2-year institutions to 3.1 for students in private nonprofit, less-than-4-year institutions. For the absolute differences between means for the full sample and respondents after poststratification adjustment, the mean ranged from 0.7 for students in private nonprofit, 4-year, non-doctorate-granting institutions to 4.0 for students in private for-profit, 4-year institutions, while the median ranged from 0.3 for students in public 2-year institutions to 3.1 for students in private nonprofit, less-than-4-year institutions.

Finally, an additional nonresponse bias analysis was conducted in which interview respondents and interview nonrespondents were compared, following the same procedures outlined above. As shown in table 60, before weighting, the percentage of characteristics for which bias was statistically significant was 70 percent for students overall and ranged from 6 percent for students in public less-than-2-year institutions to 48 percent for students in public 4-year, doctorate-granting institutions. The median effect size for all interview respondents was 0.07 and ranged from 0.03 for students in public 4-year, doctorate-granting institutions to 0.11 for students in public less-than-2-year institutions. Because study members, not interview respondents, are the unit of analysis in NPSAS:16, only a study member weight was created. Thus, nonresponse bias analyses after weight adjustments could not be computed, and it is unknown whether bias was reduced after adjusting the weights.

Table 61. Summary of student-level differences between means, by control and level of institution: 2015–16

Summary statistics	Overall	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	Private for-profit 4-year
Difference between means for respondents before and after poststratification adjustment[1]												
Mean absolute difference across characteristics	0.77	‡	0.72	1.16	0.77	0.97	3.65	0.70	0.81	1.07	1.25	3.94
Median absolute difference across characteristics	0.42	‡	0.23	0.71	0.46	0.48	3.13	0.54	0.43	0.70	0.70	2.82
Difference between means for full sample and respondents after poststratification adjustment[2]												
Mean absolute difference across characteristics	0.76	‡	0.79	1.30	0.80	1.01	3.53	0.72	0.80	1.24	1.44	4.03
Median absolute difference across characteristics	0.41	‡	0.33	0.90	0.46	0.47	3.05	0.58	0.48	0.98	0.76	2.90

‡Reporting standards not met (fewer than 30 unweighted nonrespondents).
[1] Respondent means before poststratification adjustment are weighted using the student base weight adjusted for nonresponse. Respondent means after poststratification adjustment are weighted using the student base weight adjusted for nonresponse and poststratification.
[2] Full-sample means are weighted using the student base weight and respondent means are weighted using the student base weight adjusted for nonresponse and poststratification.
NOTE: Characteristics that did not meet reporting standards were excluded from calculation of summary statistics. “Base weight” refers to the student sampling weight (final institution weight times student sampling adjustment) adjusted for student multiplicity and unknown eligibility.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

7.2.3 Bias Analysis: Item Level

NCES Statistical Standard 4-4-3A states: "For an item with a low total response rate, respondents and nonrespondents can be compared on sampling frame and/or questionnaire variables for which data on respondents and nonrespondents are available. Base weights must be used in such analysis. Comparison items should have very high response rates. A full range of available items should be used for these comparisons. This approach may be limited to the extent that items available for respondents and nonrespondents may not be related to the low response rate item being analyzed" (Seastrom 2014).

Moreover, NCES Statistical Standard 1-3-5 states: "Item response rates (RRI) are calculated as the ratio of the number of respondents for whom an in-scope response was obtained (I^x for item x) to the number of respondents who are asked to answer that item. The number asked to answer an item is the number of unit level respondents (I) minus the number of respondents with a valid skip for item x (V^x). When an abbreviated questionnaire is used to convert refusals, the eliminated questions are treated as item nonresponse. In the case of constructed variables, the numerator includes cases that have available data for the full set of items required to construct the variable, and the denominator includes all respondents eligible to respond to all items in the constructed variable" (Seastrom 2014). The item response rate is calculated as

$$RRI^x = I^x / (I - V^x).$$

A nonresponse bias analysis was conducted for all imputed items[44] and select analysis variables with a weighted response rate less than 85 percent for students overall or for students in a particular sector. Student final (analysis) weights are used for computing item response rates, shown in table 62. The procedures and variables used for the item-level nonresponse bias analysis are the same as those used for the student-level nonresponse bias analysis presented above. A student was defined to be an item respondent for a variable if that student had data for that variable from any source, including logical imputation.

As shown in table 62, the weighted item response rates for imputed and select analysis variables, for all students, ranged from 28 percent for *Dependents: monthly cost of supporting dependents other than children* (DEPOTCST) to 100 percent. The weighted

[44] Variables with only logical imputations (defined in section 7.5) are not included. Some of the imputed items were used to derive analysis variables but are not analysis variables themselves. For a full list of analysis variables, see appendix N. All nonimputed variables either have no missing data or are derived from variables that are imputed or have no missing data.

item response rates by type of institution ranged from 10 percent, for *Study abroad region* (STABREG) for students in private for-profit, less-than-2-year institutions, to 100 percent for several items. Some of the low item response rates can be explained by the interview response rate. While NPSAS staff derived values for many variables from multiple sources, including the student interview, student record data, and administrative data sources, they obtained data for some variables from only one source. Because the unweighted response rate for the student interview was about 66 percent, items obtained solely from that source have 34 percent nonresponse, even when all interview respondents provided an answer.

Out of 289 imputed items and select derived analysis variables, 277 had a sufficient number of eligible students (at least 30) to estimate response rates for students overall. Of these, 197 had an overall weighted response rate below 85 percent. An additional 34 items had an overall response rate greater than 85 percent but had a weighted response rate below 85 percent for at least one institution type, yielding a total of 231 items for which nonresponse bias was conducted. The response rates for 146 items were below 85 percent for all students and for all applicable institution types that met reporting requirements. The results of the nonresponse bias analyses varied across items. Table J-73 provides the results of the detailed bias analysis for the variable STABREG as an exemplar of the analysis that was conducted for all variables for which item nonresponse bias was analyzed. Table J-74 provides a summary of the item-level bias analysis for all items analyzed.

Imputation procedures (described in section 7.4) were conducted by NPSAS staff with a goal of reducing or eliminating item nonresponse bias. Although bias after imputation is not directly measurable, it is possible to compare estimates before and after imputation to determine whether the imputation changed the estimates. Changes are generally indicative of a reduction in bias, whereas no change suggests bias was not reduced or was not present.

For continuous variables, the difference between the preimputation mean and postimputation mean was computed; for categorical variables, the difference between the weighted preimputation and postimputation mean (proportion) were computed for each category. Student analysis weights were used for these comparisons. All differences were tested for statistical significance using *t*-tests. For categorical variables, the differences reported in table J-74 are size-weighted means of category-level differences[45] and are labeled as significant if any category-level difference is significant. These tests were complemented by effect size calculations.

[45] The size-weighted means are weighted using the unweighted count of eligible students in each category for the variable.

Effect sizes for categorical variables are calculated as $\sqrt{\sum_i (p_{0i} - p_{1i})^2 / p_{0i}}$, where p_{0i} is the proportion of respondents in category i after imputation, and p_{1i} is the proportion of respondents in category i before imputation. For continuous variables, effect size is the difference in preimputation and postimputation means, divided by the post-imputation standard deviation.

For students overall, statistically significant differences between the pre- and post-imputation means were found for about 48 percent of the variables (excluding those that did not meet reporting standards). Effect sizes for these differences range from 0.00 to 0.22. About 38 percent of the differences reported by sector were found to be statistically significant, with effect sizes for these differences ranging from 0.00 to 0.74.

Table 62. Weighted student-level item response rates for imputed variables, by control and level of institution: 2015–16

Variable	Variable label	Total	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting primarily subbacca-laureate	Public 4-year non-doctorate-granting primarily bacca-laureate	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	Private for-profit 4-year
AGE	Age as of 12/31/15	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
AGEGROUP	Age group as of 12/31/15	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
AIDSECT	Type of institution	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
AIDSECTG	Type of (graduate) institution	100.0	†	†	100.0	100.0	100.0	†	100.0	100.0	†	†	100.0
ALTONLN	Alternative courses: took classes only online in 2015–16	60.0	45.5	55.4	54.5	63.0	64.1	57.4	68.1	63.6	39.9	39.8	64.1
ALTONLN2	Alternative courses: program at NPSAS was entirely online	59.6	45.1	53.8	54.7	63.1	64.3	57.3	68.0	63.5	38.1	39.3	64.1
ASIANTYPE	Asian type	89.6	94.5	88.9	91.7	90.3	90.6	87.0	93.8	89.6	83.2	82.3	86.2
ATTNSTAT	Attendance pattern	93.9	95.8	94.4	94.9	94.0	95.6	91.2	94.3	93.8	81.4	82.6	90.2
AWAREIDR	Aware of income-driven student loan repayment plans	67.6	54.6	63.8	63.7	71.9	72.1	65.5	71.3	71.0	48.3	51.1	67.4
AWARELFP	Aware of student loan forgiveness programs	67.5	54.6	63.7	63.6	71.6	71.9	65.1	71.1	71.0	48.4	51.0	67.3
BANK1	Bank accounts: had checking or savings account	57.7	42.6	54.1	53.4	60.2	62.6	55.8	63.4	62.3	36.7	38.2	52.8
BANK2	Bank accounts: individual or shared	57.4	41.4	53.5	52.3	60.0	62.3	50.3	63.0	62.1	32.5	35.7	53.0
BAYEARM	Year and month received bachelor's degree	89.2	†	†	63.5	89.9	91.4	†	89.6	87.8	†	†	84.6
CAMPAMT	Federal campus-based aid	84.9	87.8	85.4	86.7	73.6	86.1	80.5	92.9	88.3	79.6	77.6	75.6

See notes at end of table.

Table 62. Weighted student-level item response rates for imputed variables, by control and level of institution: 2015–16—Continued

Variable	Variable label	Total	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting primarily subbacca-laureate	Public 4-year non-doctorate-granting primarily bacca-laureate	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	Private for-profit 4-year
CITIZEN2	Citizenship	95.6	94.5	93.5	94.2	96.6	97.3	98.8	96.6	95.2	99.0	97.6	97.5
CRBALCR	Credit cards: balance carried over each month	57.6	40.1	53.6	55.3	58.9	62.7	50.5	62.2	62.3	33.2	35.2	52.8
CRBALDUE	Credit cards: balance due on all credit cards	55.2	50.8	51.1	55.3	59.6	60.7	46.0	62.3	60.0	32.0	32.2	51.5
CRNUMCRD	Credit cards: number of credit cards in own name	57.7	42.9	54.1	53.1	60.5	62.7	55.9	63.5	62.2	36.9	38.2	52.9
CRTUIT	Credit cards: used credit cards to pay tuition and fees in 2015–16	57.5	40.1	53.5	55.3	59.1	62.6	49.2	62.2	62.3	33.2	35.2	52.8
CRTUIT2	Credit cards: only source available to pay tuition and fees in 2015–16	57.1	33.6	53.2	58.3	61.7	64.2	36.2	61.1	62.3	22.1	23.5	52.9
CSTBKSDG	Amount spent on digital textbooks	56.6	40.9	53.2	52.6	59.4	61.7	52.9	62.3	61.2	35.1	36.0	50.6
CSTBKSPR	Amount spent on textbooks (print only)	56.6	40.9	53.2	52.6	59.4	61.7	52.9	62.3	61.2	35.1	36.0	50.6
CSTOTHER	Amount spent on other required course materials	56.6	40.9	53.2	52.6	59.4	61.7	52.9	62.3	61.2	35.1	36.0	50.6
CSTSUPP	Amount spent on required supplies	56.6	40.9	53.2	52.6	59.4	61.7	52.9	62.3	61.2	35.1	36.0	50.6
CSTTECH	Amount spent on required technology	56.6	40.9	53.2	52.6	59.4	61.7	52.9	62.3	61.2	35.1	36.0	50.6
DECMAJ	Formally declared major field of study	97.2	100.0	97.2	89.6	93.3	98.7	94.1	97.9	98.5	100.0	99.8	97.1

See notes at end of table.

Table 62. Weighted student-level item response rates for imputed variables, by control and level of institution: 2015–16—Continued

Variable	Variable label	Total	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting primarily subbacca-laureate	Public 4-year non-doctorate-granting primarily bacca-laureate	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	Private for-profit 4-year
DEGPR	Prior degree earned since high school	69.0	56.7	65.1	64.9	73.5	73.4	66.9	72.9	72.5	49.5	52.5	68.8
DEGPRAA	Prior degree: associate's degree	70.4	66.2	66.1	64.7	75.9	74.1	63.7	72.2	72.3	44.3	54.2	70.6
DEGPRBA	Prior degree: 4-year bachelor's degree	70.4	66.2	66.1	64.7	75.9	74.1	63.7	72.2	72.3	44.3	54.2	70.6
DEGPRCRT	Prior degree: undergraduate certificate/diploma	70.4	66.2	66.1	64.7	75.9	74.1	63.7	72.2	72.3	44.3	54.2	70.6
DEGPRDOC	Prior degree: doctorate or professional degree	70.4	66.2	66.1	64.7	75.9	74.1	63.7	72.2	72.3	44.3	54.2	70.6
DEGPRMS	Prior degree: master's degree	70.4	66.2	66.1	64.7	75.9	74.1	63.7	72.2	72.3	44.3	54.2	70.6
DEGPRPTB	Prior degree: post-BA certificate	70.4	66.2	66.1	64.7	75.9	74.1	63.7	72.2	72.3	44.3	54.2	70.6
DEGPRPTM	Prior degree: post-MA certificate	70.4	66.2	66.1	64.7	75.9	74.1	63.7	72.2	72.3	44.3	54.2	70.6
DEPANY	Dependents: has dependents	88.0	77.1	84.1	86.8	91.7	90.1	94.3	90.7	86.3	92.9	94.3	94.9
DEPCARE	Dependents: children in paid childcare	50.2	41.1	48.4	50.9	55.8	57.7	49.0	56.2	52.1	32.4	34.3	49.3
DEPCHILD	Dependents: has dependent children	89.0	79.0	85.2	87.2	92.3	91.1	94.2	91.7	87.9	93.6	94.7	95.1
DEPCOL	Independent students: number of dependents in college	82.6	76.5	79.2	81.9	84.1	83.4	90.2	86.3	78.8	90.0	89.3	89.2

See notes at end of table.

Table 62. Weighted student-level item response rates for imputed variables, by control and level of institution: 2015–16—Continued

Variable	Variable label	Total	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting primarily subbacca-laureate	Public 4-year non-doctorate-granting primarily bacca-laureate	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	Private for-profit 4-year
DEPCOLCS	Amount contributed to college costs for dependents in 2015–16	44.2	‡	40.2	65.6	55.4	51.4	‡	46.9	47.5	26.2	17.1	38.6
DEPCOST	Dependents: children in paid childcare—monthly costs	49.0	29.7	46.2	54.4	49.3	57.6	39.9	51.0	51.2	29.9	37.6	48.0
DEPEND	Dependency status	92.6	88.6	91.3	90.7	95.2	92.2	97.3	94.5	91.2	96.6	97.3	98.9
DEPINC	Dependent students: parents' income	80.4	77.2	75.3	77.0	86.9	81.3	91.7	83.7	81.8	94.2	95.9	94.7
DEPINCX	Dependent students: parents' income (cat)	92.2	82.9	89.2	90.1	95.6	93.3	96.9	93.2	93.2	97.3	99.2	98.3
DEPNUM	Dependents: total number	87.1	73.7	83.1	85.8	90.3	89.5	93.7	89.8	85.3	92.0	93.6	93.3
DEPNUMCH	Dependents: number of dependent children	87.3	74.6	83.3	85.5	90.4	90.2	92.9	90.1	86.0	91.6	92.2	90.7
DEPNUMOT	Dependents: number of dependents other than children	86.5	73.3	82.5	85.1	90.0	89.3	92.7	89.2	84.5	91.0	92.0	91.3
DEPOTCST	Dependents: monthly cost of supporting dependents other than children	27.8	‡	24.5	30.2	27.0	36.3	39.9	33.1	22.9	27.8	27.8	31.3
DEPOTHER	Dependents: has dependent other than children	87.7	74.3	83.9	86.5	91.4	89.9	93.7	90.2	85.8	92.3	94.1	94.8
DEPYNG	Dependents: children, age of youngest	52.0	46.7	50.2	51.8	55.6	60.1	54.1	56.3	54.2	35.2	33.4	51.1

See notes at end of table.

Table 62. Weighted student-level item response rates for imputed variables, by control and level of institution: 2015–16—Continued

Variable	Variable label	Total	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting primarily subbacca-laureate	Public 4-year non-doctorate-granting primarily bacca-laureate	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	Private for-profit 4-year
DERMAJ	Derived major: 2010 CIP code	97.5	99.5	97.2	89.9	93.5	99.0	95.2	98.2	99.5	92.2	99.8	97.4
DIS16A	Disability: deaf or serious difficulty hearing	57.8	42.7	54.2	53.6	60.4	62.9	56.0	63.4	62.3	36.9	37.9	53.0
DIS16B	Disability: blind or serious difficulty seeing	57.7	42.7	54.2	53.4	60.4	62.8	55.9	63.4	62.2	36.9	38.2	52.8
DIS17A	Disability: serious difficulty concentrating, remembering, making decisions	57.7	42.7	54.2	53.6	60.3	62.8	55.7	63.4	62.2	36.9	38.2	52.9
DIS17B	Disability: serious difficulty walking or climbing stairs	57.8	42.6	54.2	53.6	60.4	62.9	55.9	63.4	62.3	36.9	38.1	53.0
DISTANCE	Distance from student's home (in miles) to NPSAS school	94.3	94.7	93.4	93.4	96.2	94.8	98.7	94.4	92.6	96.5	97.3	96.8
DISTYPES	Disability: main type of condition or impairment	52.5	33.6	49.5	49.9	55.2	59.2	40.6	55.9	60.3	38.5	31.7	46.5
DODAMT	Department of Defense (military) aid	83.8	86.7	83.4	85.9	70.4	85.0	76.3	92.3	87.1	80.3	75.1	80.7
DSTUINC	Dependent students: income	79.9	75.7	73.7	74.0	86.0	81.6	86.5	84.6	83.7	88.1	93.7	94.4
DSTUINCX	Dependent students: income (categorical)	91.7	83.4	88.7	88.2	94.6	92.9	92.0	93.4	94.0	94.7	98.3	98.1
EMPLWAIV	Institutional tuition waivers for staff	82.7	86.7	80.7	85.9	70.5	84.1	72.3	91.4	87.1	80.9	75.0	81.4

See notes at end of table.

Table 62. Weighted student-level item response rates for imputed variables, by control and level of institution: 2015–16—Continued

Variable	Variable label	Total	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting primarily subbacca-laureate	Public 4-year non-doctorate-granting primarily bacca-laureate	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	Private for-profit 4-year
EMPLYAM3	Employer aid (student and parents)	58.6	45.7	54.8	53.7	60.9	63.4	57.0	64.1	64.0	37.6	38.9	54.3
EMPLYAMT	Employer aid	48.1	36.1	42.8	44.7	43.7	53.5	38.9	58.5	56.3	30.1	28.3	43.9
ENR01	Monthly enrollment status 2015/07	97.3	98.1	98.2	98.3	98.4	98.3	93.7	97.5	96.5	88.7	86.8	93.7
ENR02	Monthly enrollment status 2015/08	96.9	99.2	97.0	96.7	95.7	98.1	94.4	98.1	96.9	88.3	89.3	95.3
ENR03	Monthly enrollment status 2015/09	98.0	99.3	97.7	97.3	99.5	99.1	96.2	98.7	98.5	89.3	91.7	96.5
ENR04	Monthly enrollment status 2015/10	98.3	99.3	97.9	97.4	99.5	99.3	96.6	98.9	98.7	90.9	93.2	97.2
ENR05	Monthly enrollment status 2015/11	98.4	99.7	98.0	97.5	99.6	99.4	96.7	99.0	98.7	91.5	94.1	97.5
ENR06	Monthly enrollment status 2015/12	98.6	99.7	98.2	97.8	99.7	99.5	96.7	99.0	98.9	91.8	94.8	97.8
ENR07	Monthly enrollment status 2016/01	98.4	98.5	98.7	98.4	97.7	98.8	97.0	98.0	98.8	93.5	95.2	97.2
ENR08	Monthly enrollment status 2016/02	99.0	99.3	99.2	99.1	98.2	99.5	98.4	99.1	99.3	94.5	96.8	98.3
ENR09	Monthly enrollment status 2016/03	99.2	99.7	99.3	99.0	99.1	99.5	98.3	99.2	99.5	95.4	97.7	98.4
ENR10	Monthly enrollment status 2016/04	99.3	99.7	99.4	99.0	99.5	99.6	98.6	99.4	99.5	96.1	97.6	98.7
ENR11	Monthly enrollment status 2016/05	99.3	100.0	99.3	99.2	99.4	99.6	98.2	99.3	99.6	95.7	98.0	98.8
ENR12	Monthly enrollment status 2016/06	99.3	100.0	99.1	99.2	99.9	99.7	98.2	99.3	99.6	95.1	97.9	99.0
FAMHELP	Other financial support received	57.9	43.8	54.4	53.6	60.4	62.8	56.3	63.5	62.4	36.9	38.3	53.0

See notes at end of table.

Table 62. Weighted student-level item response rates for imputed variables, by control and level of institution: 2015–16—Continued

Variable	Variable label	Total	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting primarily subbacca-laureate	Public 4-year non-doctorate-granting primarily bacca-laureate	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	Private for-profit 4-year
FAMHPAM	Help from family and friends: total amount in 2015–16	57.8	43.8	54.3	53.6	60.3	62.7	56.3	63.2	62.3	36.9	38.2	53.0
FEDBEN	Received federal benefit: any	60.2	47.0	57.1	57.5	62.1	64.1	64.1	64.4	63.0	50.3	47.8	56.9
FEDBENA	Received federal benefit: Food Stamp Benefit	59.4	44.4	56.0	56.1	61.4	63.5	61.3	63.8	62.8	47.6	46.7	55.9
FEDBENB	Received federal benefit: Free or Reduced-Price School Lunch Benefits	58.9	44.2	55.4	55.1	61.4	63.8	57.9	63.8	63.2	40.3	40.5	54.2
FEDBENC	Received federal benefit: Supplemental Security Income Benefits	57.2	43.6	53.3	52.8	59.4	62.3	56.1	62.6	62.2	37.1	37.7	52.1
FEDBEND	Received federal benefit: TANF Benefits	57.0	41.8	53.3	52.4	59.4	62.0	55.0	62.2	61.9	37.9	37.9	51.9
FEDBENE	Received federal benefit: WIC Benefits	58.4	44.8	54.9	54.1	60.3	63.0	57.0	63.5	63.2	39.5	40.2	54.2
FEDDUERATIO1	Ratio of undergraduate federal loans owed (principal + interest) to undergraduate federal amount borrowed	96.1	99.4	94.8	91.9	96.9	96.5	97.8	96.3	95.8	95.9	98.3	98.6

See notes at end of table.

Table 62. Weighted student-level item response rates for imputed variables, by control and level of institution: 2015–16—Continued

Variable	Variable label	Total	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting primarily subbacca-laureate	Public 4-year non-doctorate-granting primarily bacca-laureate	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	Private for-profit 4-year
FEDDUERATIO2	Ratio of graduate federal loans owed (principal + interest) to graduate federal amount borrowed	91.7	†	†	‡	91.8	89.4	†	90.7	92.1	†	†	97.4
FIN2000	Financial security: $2,000 within the next month	57.7	42.6	54.1	53.3	60.3	62.7	56.0	63.3	62.2	36.9	38.2	52.7
FINLIT1	Financial literacy: effect of inflation on purchasing	57.0	42.3	53.2	52.8	59.4	62.1	55.3	62.6	61.7	36.3	37.4	52.0
FINLIT2	Financial literacy: effect of interest on savings	57.3	42.5	53.6	52.8	59.6	62.4	55.5	62.8	61.8	36.4	37.6	52.4
FINLIT3	Financial literacy: effect of diversification on risk	57.4	42.6	53.7	53.1	60.3	62.5	55.5	63.2	61.9	36.7	38.1	52.6
GAINSUR	Graduate assistantship: included health insurance	68.0	†	†	‡	70.2	73.5	†	46.7	51.6	†	†	‡
GPA	Grade point average	88.3	66.5	85.5	90.9	85.8	94.1	76.8	88.9	93.7	53.3	87.6	87.5
GPACAT	Grade point average (categorical)	90.2	78.4	87.9	92.7	87.8	94.5	82.1	91.4	94.5	64.4	89.6	89.5
GPLUSAMT	Graduate PLUS loans	100.0	†	†	100.0	100.0	100.0	†	100.0	100.0	†	†	100.0
GRADDEG	Graduate degree program	99.5	†	†	100.0	100.0	99.5	†	99.3	99.3	†	†	99.8
GRADLVL	Graduate class level	76.0	†	†	90.3	80.0	75.3	†	75.5	74.9	†	†	81.3
GRADPYM	Year and month began graduate school	72.5	†	†	75.2	74.8	74.4	†	71.0	71.4	†	†	68.8

See notes at end of table.

Table 62. Weighted student-level item response rates for imputed variables, by control and level of institution: 2015–16—Continued

Variable	Variable label	Total	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting primarily subbacca-laureate	Public 4-year non-doctorate-granting primarily bacca-laureate	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	Private for-profit 4-year
GRASTAMT	Graduate assistantships	94.9	†	†	100.0	90.7	96.0	†	96.9	94.3	†	†	93.5
GRGRDAMT	Other graduate assistantship amount	95.0	†	†	100.0	90.7	96.1	†	96.9	94.4	†	†	93.5
GRINFEL	Graduate fellowship amount	95.1	†	†	100.0	90.9	95.8	†	96.6	95.1	†	†	93.7
GRJOBHR	Graduate school job: hours worked per week	68.0	†	†	‡	70.1	73.4	†	53.3	52.0	†	†	28.1
GRJOBWK	Graduate school job: proportion of weeks worked	68.1	†	†	‡	70.1	73.6	†	53.3	52.0	†	†	18.8
GRRESAMT	Graduate research assistantship amount	95.0	†	†	100.0	90.8	96.1	†	96.9	94.4	†	†	93.5
GRTEAAMT	Graduate teaching assistantship amount	94.9	†	†	100.0	90.8	96.0	†	96.9	94.3	†	†	93.5
HCHONORS	Number of honors subjects	49.3	27.7	34.2	41.7	59.3	68.0	27.3	60.3	70.1	19.7	18.8	18.0
HCMATHHI	Highest level of math completed or planned	75.3	63.5	68.5	72.7	80.0	85.4	66.5	82.5	83.0	47.1	50.5	59.1
HCSCINUM	Number of science courses taken	52.7	40.7	38.7	49.4	60.6	71.4	29.6	64.1	68.0	21.6	23.4	20.4
HCTKBIOL	Took or planned to take biology	52.4	39.7	38.5	49.4	59.9	71.1	29.6	63.8	67.8	20.9	23.3	20.1
HCTKCHECM	Took or planned to take chemistry	51.8	38.0	37.5	48.5	58.9	70.7	28.6	63.4	67.4	20.5	22.4	19.9

See notes at end of table.

Table 62. Weighted student-level item response rates for imputed variables, by control and level of institution: 2015–16—Continued

Variable	Variable label	Total	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting primarily subbacca-laureate	Public 4-year non-doctorate-granting primarily bacca-laureate	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	Private for-profit 4-year
HCTKPHYS	Took or planned to take physics	49.0	36.1	35.1	45.3	55.7	67.6	27.2	60.5	64.1	19.5	20.9	18.7
HCYSENGL	Years completed or planned English	51.6	36.6	37.6	46.9	60.6	70.2	28.5	62.5	66.2	20.3	22.5	20.3
HCYSLANG	Years completed or planned foreign languages	38.7	26.0	27.4	33.6	46.8	53.3	23.4	46.0	52.3	15.4	15.7	15.0
HCYSMATH	Years completed or planned math	78.1	66.1	71.4	75.9	83.7	88.5	68.3	84.8	85.2	48.4	51.9	60.6
HCYSSCIE	Years completed or planned science	51.1	35.8	37.2	46.5	60.1	69.7	27.9	62.1	65.7	20.1	22.1	19.8
HCYSSOCI	Years completed or planned social studies	51.2	36.6	37.3	46.3	60.5	70.0	28.1	62.2	65.7	20.1	22.1	20.1
HIGHLVEX	Highest level of education ever expected	57.6	44.1	54.3	54.1	60.8	62.4	57.0	64.4	59.3	37.0	38.7	54.3
HISPANIC	Race/ethnicity: Hispanic or Latino origin	96.5	97.5	97.7	96.5	96.7	96.6	95.1	96.2	95.4	97.2	95.6	93.0
HISPTYPE	Race/ethnicity: type of Hispanic origin	89.5	91.1	88.6	86.5	90.4	92.4	88.7	91.3	90.8	76.2	77.4	86.3
HOMELESS	Homeless or at risk of homelessness	71.6	59.8	67.6	69.6	77.8	78.8	66.8	76.7	71.9	60.6	59.4	59.1
HOMESTUD	Student owns home or pays mortgage	55.5	44.1	53.6	52.3	57.1	60.3	55.7	57.2	59.7	36.7	37.6	52.5
HSCRDAP	Took AP courses while in high school	58.0	43.5	55.4	53.1	62.1	62.4	56.6	65.2	61.8	36.9	39.8	50.9
HSCRDCOL	Took college-level courses while in high school	58.0	43.5	55.4	53.1	62.1	62.3	56.6	65.2	61.8	36.9	39.8	50.9

See notes at end of table.

Table 62. Weighted student-level item response rates for imputed variables, by control and level of institution: 2015–16—Continued

Variable	Variable label	Total	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting primarily subbacca-laureate	Public 4-year non-doctorate-granting primarily bacca-laureate	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	Private for-profit 4-year
HSCRDIB	Took IB courses while in high school	58.0	43.5	55.3	53.1	62.1	62.3	56.6	65.2	61.9	36.9	39.8	50.9
HSDEG	High school degree type	88.0	91.8	87.8	91.1	86.6	88.4	95.2	87.6	84.8	88.5	87.8	90.7
HSGPA	Grade point average in high school	73.6	59.7	65.8	69.3	79.6	85.5	63.2	80.4	83.3	44.7	46.6	52.9
HSGRADYM	High school graduation year and month	66.6	55.3	63.9	63.5	71.1	71.6	64.8	70.6	69.4	48.2	51.4	65.3
HSGRADYY	High school graduation year	91.9	81.5	93.7	94.2	91.4	92.8	89.0	88.4	88.7	88.8	82.7	88.3
HSTYPE	Type of high school attended	54.6	42.6	52.1	50.6	58.1	59.3	54.3	60.2	58.5	37.5	36.5	50.1
IMMIAGE	Age when arrived in the United States	51.3	‡	49.5	49.3	52.5	57.0	41.1	55.0	50.1	29.8	27.2	47.9
INATHAMT	Athletic scholarships	83.3	86.7	82.2	85.9	70.1	84.7	76.3	93.3	88.8	80.9	75.0	80.7
INCOME	Dependency and income in 2014	74.3	70.4	66.3	69.2	80.2	77.8	90.6	80.7	76.0	89.5	91.7	89.1
INCOMEG	Total income of graduate students	59.2	†	†	38.9	49.4	55.1	†	58.0	60.6	†	†	75.7
INGRTAMT	Institutional grants	82.5	86.7	80.7	85.8	70.4	83.8	72.3	91.0	86.8	80.2	74.8	81.2
INLNAMT	Institutional loans	94.2	96.5	94.2	96.0	90.7	94.8	91.7	96.9	95.1	88.7	87.0	92.4
INSMILAMT	Institution military/ armed forces grants	83.6	86.7	82.2	85.9	70.5	85.3	76.3	92.9	87.3	81.0	75.1	81.4
INSTAMT	Institutional aid total	82.5	86.7	80.7	85.8	70.4	83.8	72.3	91.0	86.8	80.2	74.8	81.2
INSTCATGRT	Institutional categorical grants	83.3	87.5	82.3	85.9	70.2	84.8	76.3	93.3	88.8	80.9	75.0	81.0

See notes at end of table.

Table 62. Weighted student-level item response rates for imputed variables, by control and level of institution: 2015–16—Continued

Variable	Variable label	Total	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting primarily subbacca-laureate	Public 4-year non-doctorate-granting primarily bacca-laureate	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	Private for-profit 4-year
INSTNEED	Institutional need-based grants	83.6	86.7	82.3	86.1	70.1	85.0	78.9	93.6	90.3	81.1	77.1	80.6
INSTWRK	Institutional work-study	91.1	96.3	92.2	94.7	84.2	91.1	90.4	95.2	90.5	88.0	85.2	89.2
INSVETAMT	Institution Veterans' education benefits	83.6	86.7	82.2	85.9	70.5	85.3	76.3	92.9	87.3	80.9	75.0	81.4
ISTUINC	Independent students: income	66.5	64.6	61.9	64.7	66.0	63.3	91.2	70.7	61.8	87.9	90.0	83.6
ISTUINCX	Independent students: income (categorical)	86.3	78.6	82.1	85.8	89.0	88.2	95.9	89.3	83.5	93.5	93.6	94.3
JOBEARN1	Job 1: earnings rate per hour	56.4	40.6	52.6	52.8	59.6	61.2	50.7	62.2	61.1	32.4	37.4	52.2
JOBEARN2	Job 2: earnings rate per hour	56.1	35.1	52.4	51.2	58.7	59.1	54.2	62.3	61.2	34.1	40.2	51.9
JOBEARN3	Job 3: earnings rate per hour	57.2	‡	56.5	46.4	60.6	57.2	‡	60.6	62.9	31.5	43.5	52.2
JOBEARN4	Job 4: earnings rate per hour	54.4	‡	52.8	22.7	63.4	56.0	‡	68.0	55.7	‡	‡	42.9
JOBEARN5	Job 5: earnings rate per hour	50.7	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡
JOBEARN6	Job 6: earnings rate per hour	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡
JOBEARN7	Job 7: earnings rate per hour	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡
JOBENR1	Job 1: worked while enrolled in school	57.9	41.1	54.1	54.2	61.5	62.7	51.7	64.2	63.1	33.2	38.0	53.7
JOBENR2	Job 2: worked while enrolled in school	58.1	37.8	53.7	51.9	61.7	61.4	54.3	64.5	63.7	34.9	41.6	53.8

See notes at end of table.

Table 62. Weighted student-level item response rates for imputed variables, by control and level of institution: 2015–16—Continued

Variable	Variable label	Total	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting primarily subbacca-laureate	Public 4-year non-doctorate-granting primarily bacca-laureate	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	Private for-profit 4-year
JOBENR3	Job 3: worked while enrolled in school	59.2	‡	58.4	43.4	63.8	59.1	‡	64.8	65.0	32.6	45.2	54.2
JOBENR4	Job 4: worked while enrolled in school	57.9	‡	52.8	32.9	68.9	63.3	‡	68.0	60.9	‡	‡	43.9
JOBENR5	Job 5: worked while enrolled in school	50.7	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡
JOBENR6	Job 6: worked while enrolled in school	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡
JOBERN7	Job 7: worked while enrolled in school	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡
JOBHOUR1	Job 1: hours worked per week	57.5	40.5	53.7	54.1	60.8	62.2	50.8	63.6	62.6	32.9	37.8	53.3
JOBHOUR2	Job 2: hours worked per week	57.5	35.2	53.2	51.8	61.1	60.6	54.0	63.7	63.0	34.6	41.4	53.5
JOBHOUR3	Job 3: hours worked per week	58.3	‡	58.4	41.2	63.5	57.5	‡	63.5	64.1	32.6	45.2	53.7
JOBHOUR4	Job 4: hours worked per week	57.5	‡	52.8	32.3	68.9	62.3	‡	68.0	60.5	‡	‡	43.9
JOBHOUR5	Job 5: hours worked per week	50.7	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡
JOBHOUR6	Job 6: hours worked per week	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡
JOBHOUR7	Job 7: hours worked per week	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡
JOBMAJOR1	Job 1: job related to major or coursework	58.0	41.1	54.1	53.9	61.4	62.8	51.2	64.2	63.2	33.3	37.9	53.6
JOBMAJOR2	Job 2: job related to major or coursework	58.1	37.8	53.7	51.1	61.7	61.4	54.8	64.7	63.7	34.9	41.3	54.0
JOBMAJOR3	Job 3: job related to major or coursework	59.6	‡	58.1	46.5	63.9	60.0	‡	64.6	65.0	32.6	45.7	54.7

See notes at end of table.

Table 62. Weighted student-level item response rates for imputed variables, by control and level of institution: 2015–16—Continued

Variable	Variable label	Total	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting primarily subbacca-laureate	Public 4-year non-doctorate-granting primarily bacca-laureate	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	Private for-profit 4-year
JOBMAJOR4	Job 4: job related to major or coursework	57.9	‡	52.8	32.9	68.9	63.3	‡	68.0	60.9	‡	‡	43.9
JOBMAJOR5	Job 5: job related to major or coursework	50.7	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡
JOBMAJOR6	Job 6: job related to major or coursework	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡
JOBMAJOR7	Job 7: job related to major or coursework	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡
JOBONOFF1	Job 1: job on or off campus	57.5	45.3	53.7	56.8	65.0	62.2	46.2	64.1	62.2	28.2	34.6	53.9
JOBONOFF2	Job 2: job on or off campus	57.6	50.6	52.9	57.3	64.6	61.0	50.6	64.3	61.8	30.1	39.9	55.2
JOBONOFF3	Job 3: job on or off campus	58.0	‡	55.2	46.2	62.1	58.6	‡	66.6	63.0	24.5	51.4	50.9
JOBONOFF4	Job 4: job on or off campus	61.6	‡	54.1	‡	‡	63.6	‡	72.5	71.4	‡	‡	49.3
JOBONOFF5	Job 5: job on or off campus	69.3	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡
JOBONOFF6	Job 6: job on or off campus	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡
JOBONOFF7	Job 7: job on or off campus	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡
JOBROLE	Job: primary role as student or employee	57.5	44.6	53.7	56.5	64.7	62.0	47.6	64.5	62.4	28.8	35.3	53.4
JOBWKST1	Job 1: work-study job	55.9	45.3	53.5	56.2	65.1	60.4	46.1	63.2	60.2	28.1	34.5	52.0
JOBWKST2	Job 2: work-study job	56.6	50.6	52.7	57.3	64.5	59.3	50.1	65.1	61.3	29.7	39.3	55.5
JOBWKST3	Job 3: work-study job	57.7	‡	55.2	46.0	61.3	58.5	‡	65.2	66.2	24.5	51.4	46.7

See notes at end of table.

Table 62. Weighted student-level item response rates for imputed variables, by control and level of institution: 2015–16—Continued

Variable	Variable label	Total	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting primarily subbacca-laureate	Public 4-year non-doctorate-granting primarily bacca-laureate	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	Private for-profit 4-year
JOBWKST4	Job 4: work-study job	61.1	‡	54.1	‡	‡	62.1	‡	73.2	70.2	‡	‡	55.3
JOBWKST5	Job 5: work-study job	68.4	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡
JOBWKST6	Job 6: work-study job	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡
JOBWKST7	Job 7: work-study job	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡
LNREPAY	Expect help with repaying student loans	51.7	35.3	41.1	38.4	54.6	57.9	50.2	57.9	59.8	34.6	35.1	47.8
LOANLIT1	Loan literacy: government can report unpaid debt to credit bureaus	57.9	42.7	54.3	53.6	60.4	62.9	56.0	63.5	62.4	36.8	38.2	53.0
LOANLIT2	Loan literacy: government can garnish wages for unpaid federal loan debt	57.9	42.7	54.3	53.6	60.4	62.9	56.0	63.5	62.4	36.8	38.2	53.0
LOANLIT3	Loan literacy: government can retain tax refunds, SS for loan debt	57.9	42.7	54.3	53.6	60.4	62.9	56.0	63.5	62.4	36.8	38.2	53.0
LOCALEST	Degree of Urbanization of student's permanent address	96.3	95.6	93.9	94.2	98.6	98.0	98.9	97.8	96.8	96.6	97.8	98.0
LOCALRES	Residence while enrolled	80.0	88.6	75.0	71.3	78.3	84.9	88.8	88.9	88.1	82.6	75.8	81.1
MAJCHGFQ	Majors: frequency of formally changed	48.5	‡	40.4	42.2	55.1	58.1	48.3	58.2	55.6	‡	31.7	45.4

See notes at end of table.

Table 62. Weighted student-level item response rates for imputed variables, by control and level of institution: 2015–16—Continued

Variable	Variable label	Total	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting primarily subbaccalaureate	Public 4-year non-doctorate-granting primarily baccalaureate	Public 4-year doctorate-granting	Private nonprofit less-than-4-year	Private nonprofit 4-year non-doctorate-granting	Private nonprofit 4-year doctorate-granting	Private for-profit less-than-2-year	Private for-profit 2-year	Private for-profit 4-year
MERITNOATH	Institutional merit-only grants except for athletic scholarships	83.5	86.7	82.3	85.9	70.1	85.2	76.6	93.8	89.2	81.0	75.3	80.7
MILTYPE	Military service type	73.4	62.9	70.4	69.8	75.8	75.9	73.4	77.1	76.6	55.4	60.4	77.5
MILTYPE2	Military service type (for dependency)	83.6	75.6	80.7	82.6	87.8	87.5	81.4	86.7	83.3	73.9	77.0	81.6
MNTRENT	Monthly mortgage or rent amount	54.3	42.2	52.2	51.0	55.6	59.4	54.4	56.0	58.5	35.6	36.4	51.5
NFEDCUM1	Cumulative non-federal loan amount for undergraduate students	65.6	53.1	62.4	62.3	69.6	70.1	63.5	69.0	68.6	45.1	47.5	64.2
NFEDCUM2	Cumulative non-federal loan amount for graduate students	64.5	†	†	84.4	63.1	66.8	†	65.2	63.8	†	†	57.9
NUMJOB3	Job: number in 2015–16 (including work-study)	58.0	43.8	54.4	53.8	60.4	63.0	56.4	63.6	62.4	37.1	38.4	53.3
ORPHAN	Orphan, ward of court, emancipated minor, or in legal guardianship	72.7	60.0	69.3	72.0	78.6	80.1	66.9	76.9	72.1	61.3	60.9	59.1
OTHFDGRT	Other federal grants (not Title IV)	80.4	86.7	78.3	85.2	68.5	82.5	66.5	89.3	86.1	76.8	72.6	73.5
PAGI	Dependent students: parents adjusted gross income	80.4	76.8	75.2	76.8	86.7	81.3	90.8	83.2	82.5	93.4	96.5	94.8
PAR1	Parent type 1 (for parents' highest education)	63.6	46.8	59.4	60.0	69.0	68.5	57.2	66.9	66.3	42.0	47.0	65.1

See notes at end of table.

Table 62. Weighted student-level item response rates for imputed variables, by control and level of institution: 2015–16—Continued

Variable	Variable label	Total	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting primarily subbacca-laureate	Public 4-year non-doctorate-granting primarily bacca-laureate	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	Private for-profit 4-year
PAR2	Parent type 2 (for parents' highest education)	61.4	45.2	56.8	57.7	65.8	66.8	55.0	65.0	65.0	39.4	44.3	62.8
PARBORN	Born in the U.S. (parents)	59.3	45.8	55.7	54.7	61.9	64.2	58.2	65.0	63.9	38.3	39.7	54.6
PARED1	Parent 1: highest education level	64.1	47.5	60.1	60.7	69.4	68.8	58.7	67.2	66.8	42.5	47.9	66.0
PARED2	Parent 2: highest education level	62.3	46.6	58.1	58.2	67.6	67.3	56.6	65.5	65.2	39.8	45.6	63.5
PARHELP2	Help from parents: housing, tuition, and other expenses (including independent students)	58.1	43.8	54.6	53.7	60.7	63.0	56.3	63.7	62.5	37.1	38.5	53.2
PARHPAMT2	Help from parents: amount parents helped pay for expenses in 2015–16 (including independent students)	57.6	42.6	54.1	53.5	60.1	62.4	56.1	63.3	62.0	36.9	38.1	53.0
PELLAMT	Pell grants	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
PFAMNUM	Dependent students: family size	91.6	83.8	88.2	87.1	94.7	93.1	93.7	93.1	93.6	95.6	98.2	97.6
PFEDTAX	Dependent students: parents federal tax paid	71.7	68.2	64.3	63.9	79.3	73.3	81.4	78.6	77.3	82.1	86.1	87.3
PHSLOAN	Federal health professions loans	90.8	99.5	93.3	93.6	82.8	90.4	82.6	95.3	91.4	84.2	79.8	84.3
PINCOL	Dependent students: number of family members in college	84.7	78.5	79.2	78.6	90.2	86.8	91.7	87.6	88.2	91.0	94.4	94.9

See notes at end of table.

Table 62. Weighted student-level item response rates for imputed variables, by control and level of institution: 2015–16—Continued

Variable	Variable label	Total	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting primarily subbacca-laureate	Public 4-year non-doctorate-granting primarily bacca-laureate	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	Private for-profit 4-year
PLUSAMT	Parent PLUS loans	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
PMARITAL	Dependent students: parents' marital status	92.0	85.2	88.9	87.8	94.8	93.4	93.7	93.3	94.2	96.1	98.5	98.0
PRIMLANG	English primary language spoken	57.8	42.5	54.2	53.4	60.4	62.9	55.6	63.5	62.3	36.9	38.3	53.0
PRIMLGFQ	Frequency of speaking non-English language with primary caregiver	57.6	‡	55.4	58.3	62.3	64.1	48.5	59.4	58.5	27.9	36.0	50.5
PRIVAID	Private source grants	93.5	96.3	93.8	95.0	87.6	93.8	92.7	97.3	94.9	88.0	85.6	91.5
PRIVLOAN	Private (alternative) loans	66.7	54.9	63.3	63.5	70.2	71.4	64.9	71.0	69.3	46.5	49.9	65.2
PROGSTAT	Completed degree program in 2015–16	88.3	96.2	83.3	88.9	88.5	91.2	88.9	93.9	91.6	80.9	82.0	92.4
PSECTYM	Year and month first enrolled in postsecondary education	92.0	76.8	92.6	92.2	93.6	92.9	85.1	94.3	93.3	74.3	76.9	91.2
PSECTYR	Year first enrolled in postsecondary education	93.0	77.9	93.9	93.6	94.3	93.6	86.3	94.8	94.1	75.1	77.8	92.3
PTAXFILE	Dependent students: parents federal tax filed	80.6	77.2	75.7	77.3	87.0	81.4	91.9	83.7	81.8	94.4	96.1	94.9
RAASIAN	Race: Asian	92.9	96.7	92.4	94.9	93.3	94.1	90.2	95.5	94.2	84.5	86.3	88.2
RABLACK	Race: Black or African-American	92.9	96.7	92.4	94.9	93.3	94.1	90.2	95.5	94.2	84.5	86.3	88.2

See notes at end of table.

Table 62. Weighted student-level item response rates for imputed variables, by control and level of institution: 2015–16—Continued

Variable	Variable label	Total	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting primarily subbacca-laureate	Public 4-year non-doctorate-granting primarily bacca-laureate	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	Private for-profit 4-year
RAINDIAN	Race: American Indian or Alaska Native	92.9	96.7	92.4	94.9	93.3	94.1	90.2	95.5	94.2	84.5	86.3	88.2
RAINDTRB	Race: American Indian or Alaska Native recognized tribe	65.2	‡	65.0	48.9	71.8	69.2	88.0	76.3	59.6	66.9	60.5	61.5
RAISLAND	Race: Native Hawaiian/other Pacific Islander	92.9	96.7	92.4	94.9	93.3	94.1	90.2	95.5	94.2	84.5	86.3	88.2
RAWHITE	Race: White	92.9	96.7	92.4	94.9	93.3	94.1	90.2	95.5	94.2	84.5	86.3	88.2
REANOAPA	Reason for not applying: did not want to take on debt	40.1	18.2	40.6	39.4	42.8	42.4	23.3	41.2	39.6	17.5	21.0	22.9
REANOAPB	Reason for not applying: forms were too much work	40.1	18.2	40.6	39.4	42.8	42.4	23.3	41.2	39.6	17.5	21.0	22.9
REANOAPC	Reason for not applying: no information about how to apply	40.1	18.2	40.6	39.4	42.8	42.4	23.3	41.2	39.6	17.5	21.0	22.9
REANOAPD	Reason for not applying: no need	40.1	18.2	40.6	39.4	42.8	42.4	23.3	41.2	39.6	17.5	21.0	22.9
REANOAPE	Reason for not applying: thought ineligible	40.1	18.2	40.6	39.4	42.8	42.4	23.3	41.2	39.6	17.5	21.0	22.9
REAPOAPF	Reason for not applying for aid in NPSAS year: other	40.1	18.2	40.6	39.4	42.8	42.4	23.3	41.2	39.6	17.5	21.0	22.9
REFUND1	Received a refund of scholarships or grants from NPSAS	58.1	43.7	54.5	53.6	60.6	63.3	55.5	63.8	62.6	37.3	38.5	53.5

See notes at end of table.

Table 62. Weighted student-level item response rates for imputed variables, by control and level of institution: 2015–16—Continued

Variable	Variable label	Total	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting primarily subbacca-laureate	Public 4-year non-doctorate-granting primarily bacca-laureate	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	Private for-profit 4-year
REFUND2	Method of receiving refund from NPSAS	58.8	47.8	55.2	52.1	61.2	64.1	55.4	62.5	66.2	15.0	38.1	52.8
REMETOOK	Remedial courses: took in 2015–16	53.2	43.6	49.1	52.4	63.5	63.8	64.5	61.4	56.8	43.4	40.1	51.8
REMEVER	Remedial courses: ever taken	74.3	64.6	76.9	77.0	74.8	74.8	66.9	73.4	71.6	49.0	55.1	72.7
REMMATH	Remedial courses: number taken in math in 2015–16	46.5	‡	44.3	53.2	66.4	54.0	49.5	58.2	50.0	25.0	34.5	39.0
REMRW	Remedial courses: number taken in reading and/or writing in 2015–16	46.1	‡	44.0	52.7	64.6	52.8	49.4	58.5	50.7	24.7	34.5	39.8
SAGI	Independent students: adjusted gross income	62.2	60.8	56.7	60.9	61.6	58.1	88.6	68.4	58.3	84.3	87.9	82.4
SEOGAMT	Federal Supplemental Educational Opportunity Grant (SEOG)	87.4	87.8	87.6	87.3	79.3	88.4	87.7	94.7	92.6	80.9	83.2	77.8
SFEDTAX	Independent students: federal tax paid	54.3	51.7	49.6	50.1	54.0	50.7	74.8	60.7	51.6	69.0	73.5	72.0
SFEDTAXD	Dependent students: federal tax paid	75.1	75.2	68.8	70.5	82.5	75.9	82.4	80.6	78.7	88.4	91.3	91.2
SIBCOLFT	First sibling to go to college	57.8	42.6	54.2	53.5	60.4	62.9	55.8	63.5	62.3	36.9	38.2	52.9
SINCOL	Independent students: number of family members in college	90.2	83.4	87.3	89.3	91.1	92.8	94.2	91.7	88.4	93.9	93.9	94.0

See notes at end of table.

Table 62. Weighted student-level item response rates for imputed variables, by control and level of institution: 2015–16—Continued

Variable	Variable label	Total	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting primarily subbacca-laureate	Public 4-year non-doctorate-granting primarily bacca-laureate	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	Private for-profit 4-year
SMARITAL	Student's marital status	92.6	95.0	88.7	89.4	95.3	94.2	97.7	96.4	93.1	99.4	99.5	97.2
SPINCOL	Independent students: spouse attending college	76.7	71.8	71.9	75.4	74.6	80.5	86.5	80.7	75.8	84.6	80.6	82.5
SPSINC	Independent students: spouse's income	57.1	59.6	52.9	50.2	55.7	51.9	85.6	63.3	54.8	82.1	83.3	75.1
SPSINCX	Independent students: spouse's income (categorical)	83.3	82.2	78.6	80.9	86.3	85.5	90.7	86.0	81.7	88.4	89.9	91.1
STABEVR	Ever study abroad during undergraduate education	58.9	45.2	55.2	54.6	61.6	63.9	57.4	64.7	63.4	37.7	39.3	54.3
STABREG	Study abroad region	59.7	‡	50.9	48.1	63.4	63.8	‡	62.4	63.4	10.3	25.3	57.1
STABTIME	Length of time studied abroad	55.8	‡	43.5	44.8	62.1	60.1	‡	59.8	61.5	11.1	16.1	47.1
STAFFAMT	Direct Loans	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
STAFSUB	Stafford subsidized loans	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
STAFUNSB	Stafford unsubsidized loans	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
STATEAMT	State aid total	83.4	86.7	82.7	85.4	70.3	84.3	76.3	92.3	86.8	80.3	75.1	80.7
STAXFILE	Independent students: federal tax filed	64.7	65.8	60.3	63.6	63.8	59.7	90.9	69.7	59.8	87.9	90.4	84.0
STAXFILED	Dependent students: federal tax filed	76.3	75.2	69.9	71.0	83.6	77.7	87.1	81.1	79.7	89.4	91.7	92.6
STGTAMT	State grants	83.4	86.7	82.9	85.4	70.3	84.3	76.3	92.3	86.8	80.3	75.1	80.7

See notes at end of table.

Table 62. Weighted student-level item response rates for imputed variables, by control and level of institution: 2015–16—Continued

Variable	Variable label	Total	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting primarily subbacca-laureate	Public 4-year non-doctorate-granting primarily bacca-laureate	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	Private for-profit 4-year
STLNAMT	State loans	94.7	96.7	95.4	96.0	90.7	95.0	92.6	97.6	95.1	89.2	87.7	92.4
STMERIT	State merit-only grants	84.9	89.3	84.6	85.9	70.7	85.6	76.9	93.2	87.8	81.9	76.7	85.1
STMILAMT	State military/armed forces grants	84.3	89.3	84.0	85.9	70.5	85.3	76.3	92.8	87.2	81.8	76.6	81.4
STNDMRT	State grants based both on need and merit	84.2	89.3	84.1	85.9	70.7	84.8	76.3	92.9	87.3	81.8	76.6	81.4
STNDONLY	State need-based grants	84.6	89.3	84.3	86.3	71.4	85.8	76.3	93.1	87.3	81.8	77.3	81.1
STNOND1	State non-need grants	84.0	86.7	83.8	85.6	70.7	85.0	76.3	92.3	87.1	80.3	75.1	80.7
STUSTATE	State of legal residence	98.1	98.9	98.0	97.6	97.8	99.0	98.3	98.3	97.0	98.8	99.1	97.3
STVETAMT	State Veterans' education benefits	84.3	89.3	84.0	85.9	70.5	85.3	76.3	92.8	87.2	81.8	76.6	81.4
STWKAMT	State work-study	91.9	97.8	93.7	94.6	84.2	91.6	92.3	96.1	90.6	88.7	86.2	89.2
TEACTDER	ACT derived composite score	72.5	55.0	58.9	65.6	83.9	89.3	47.9	81.4	85.9	32.7	34.5	32.1
TESATDER	SAT derived composite score	72.5	55.0	58.9	65.6	83.9	89.3	47.9	81.4	85.9	32.7	34.5	32.1
TETOOK	Took SAT or ACT exams	83.1	72.1	76.2	80.0	90.2	94.1	71.1	89.0	91.7	52.7	54.7	62.5
TFEDAID	Federal aid total	80.4	86.7	78.3	85.2	68.5	82.5	66.5	89.3	86.0	76.8	72.6	73.5
TFEDGRT	Federal grants	80.4	86.7	78.3	85.2	68.5	82.5	66.5	89.3	86.1	76.8	72.6	73.5
TFEDLN	Federal student loans	90.8	99.5	93.3	93.6	82.8	90.4	82.6	95.3	91.4	84.2	79.8	84.3
TFEDWRK	Federal work-study	90.9	96.3	93.0	94.3	83.2	90.9	86.9	95.3	90.7	87.1	84.4	84.3
TITIVAMT	Federal Title IV aid	84.9	87.8	85.4	86.7	73.6	86.1	80.5	92.9	88.3	79.6	77.6	75.6

See notes at end of table.

Table 62. Weighted student-level item response rates for imputed variables, by control and level of institution: 2015–16—Continued

Variable	Variable label	Total	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting primarily subbacca-laureate	Public 4-year non-doctorate-granting primarily bacca-laureate	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	Private for-profit 4-year
TOTAID	Total aid	45.8	34.5	40.7	44.0	41.7	51.4	36.1	56.3	53.7	27.5	27.4	38.8
TOTGRT	Total grants	46.3	36.0	40.9	44.2	42.3	51.8	36.2	56.9	55.3	27.9	27.6	39.7
TOTLOAN	Total student loans	61.3	53.3	60.1	60.5	60.3	65.1	51.4	67.6	64.1	38.7	40.6	54.5
TOTWKST	Work-study	90.0	96.3	91.2	94.2	83.2	90.3	86.9	94.2	90.1	86.3	83.4	84.3
UGDEG	Undergraduate degree program	99.9	100.0	99.8	99.9	99.9	100.0	100.0	100.0	99.9	100.0	100.0	100.0
UGLVL1	Class level (Undergraduate)	98.6	97.9	98.1	97.7	99.4	99.1	97.7	99.6	99.2	97.8	97.8	98.8
USBORN	Born in the U.S. (student)	60.8	44.9	57.3	56.8	63.4	65.7	58.2	65.3	66.1	40.5	43.9	54.9
USEIDR	Likelihood of using income-driven student loan repayment plans	47.3	33.6	35.6	29.9	50.4	54.1	32.5	52.1	57.6	31.8	29.1	42.8
USELFP	Likelihood of using loan forgiveness program	51.8	42.5	42.1	34.1	54.8	58.2	37.9	56.7	59.1	33.3	33.9	46.2
VETBEN	Veterans' benefits	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
VOCHELP	Vocational rehabilitation and training	83.3	86.7	82.9	86.0	70.1	84.3	76.3	92.6	88.4	80.3	75.2	79.8
WAIVNOEMP	Institutional tuition and fee waivers except for those to staff	82.8	86.7	80.8	85.9	70.7	84.3	72.3	91.3	87.1	80.9	75.0	81.4

† Not applicable.
‡Reporting standards not met (fewer than 30 unweighted eligible students).
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

7.3 Variance Estimation

Every estimate calculated from a probability-based sample survey, such as a mean, a percentage, or a regression coefficient, has a variance associated with it. Hypothesis testing, calculation of confidence intervals, and modelling that use complex survey data all require the calculation of variances using appropriate methods that account for the sampling design. Complex sample designs, like that used for NPSAS:16, result in data that violate the assumptions that are normally required to assess the statistical significance of results. The variances of the estimates from complex surveys may vary from those that would be expected if the sample were a simple random sample and the observations were independent and identically distributed random variables. Two procedures for estimating variances of statistics from complex surveys are the Taylor-series linearization procedure and the bootstrap replication procedure, which are both available for the NPSAS data files. The analysis strata and primary sampling units (PSUs) created for the Taylor-series procedure are discussed in section 7.3.1, and section 7.3.2 contains a discussion of the replicate weights created for the bootstrap procedure. Use of software packages for proper variance estimation is discussed in section 7.3.3.

The survey design effect for a statistic is defined as the ratio of the design-based variance estimate over the variance estimate that would have been obtained from a simple random sample of the same size (if that had been done). It is often used to measure the effects that sample design features have on the precision of survey estimates. For example, stratification tends to decrease the variance, but multistage sampling and unequal sampling rates usually increase the variance. In addition, weight adjustments for nonresponse (performed to reduce nonresponse bias) and poststratification increase the variance by increasing the weight variation. Design effects are discussed in section 7.3.4.

7.3.1 Taylor Series

The Taylor-series variance estimation procedure is a well-known technique used to estimate the variances of nonlinear statistics. The procedure takes the first-order Taylor-series approximation of the nonlinear statistic and then substitutes the linear representation into the variance formula appropriate for the sample design (Woodruff 1971).

For stratified multistage surveys, the Taylor-series procedure uses analysis strata and PSUs as defined from the sampling strata and PSUs used in the first stage of sampling. For NPSAS:16, NPSAS staff defined analysis strata and PSUs for all

students combined, such that analyses can be conducted for undergraduates and graduates separately; these are available for analyses of any domain (e.g., females).

The first step in the process of creating analysis strata and PSUs was to identify the sampling PSUs used at the first stage of sample selection. As discussed in chapter 2, the PSUs were defined as the 1,750 participating institutions. The next steps were to sort the PSUs by the 11 institution strata, then by certainty (institution probability of selection equal to one) versus noncertainty (institution probability of selection less than one), and then by the selection order for the noncertainty institutions and by IPEDS ID for the certainty institutions. Once sorted, NPSAS staff combined some adjacent PSUs/institutions to form analysis PSUs to meet certain criteria for a minimum number of respondents. Specifically, each analysis PSU was created to contain at least four responding students, which ensures stable variance estimates. Additionally, each PSU was created to contain at least two responding undergraduate students and two responding graduate students, when an institution contains both types of students, so that analyses can be conducted correctly on the separate undergraduate and graduate student analysis files. After the PSUs were combined, the resulting analysis PSUs were paired to form analysis strata. This process resulted in 853 analysis strata. Analysis secondary sampling units (SSUs) were then formed by randomly splitting responding students within a PSU into two groups.

The restricted-use data file provides two sets of variables for Taylor series variance estimation. One set of variables is used in software that assumes that the first-stage sampling units (institutions) were sampled with replacement (or with small selection probabilities) and does not account for the finite population correction (FPC) at the institution level of sampling. The other set of variables is used in software that assumes sampling of institutions without replacement in the calculation of variances and does account for the FPC. Both sets of variables are provided because not all survey data analysis packages have the option to incorporate the FPC in the variance calculations. When the first-stage units are sampled with very small probabilities, the estimated variances using the with-replacement variance formulas and the without-replacement variance formulas are the same.

The set of variables used when assuming the first-stage units were sampled with replacement includes the analysis stratum (ANALSTR) and analysis PSU (ANALPSU). The set of variables used when assuming the first-stage units were sampled without replacement and that account for the FPC includes the analysis stratum (FANALSTR), analysis PSU (FANALPSU), the analysis SSU (FANALSSU), and the count of PSUs in an analysis stratum (PSUCOUNT).

7.3.2 Bootstrap Replicate Weights

The replication variance estimation strategy chosen for NPSAS:16 accounts for the following in order to produce accurate variance estimates:

1. stratification at all stages of sampling;
2. unequal weighting;
3. sample clustering;
4. weight adjustments for nonresponse and for poststratification of selected total estimates to known external totals;
5. nonlinear statistics and percentages, as well as for linear statistics;
6. the finite population corrections at the institution stage of sampling; and
7. the ability to test hypotheses about students based on normal distribution theory by ignoring the finite population corrections at the student level of sampling.

Commonly applied bootstrap variance estimation techniques account for 1 through 5 listed above, however to account for 6 and 7 above, NPSAS staff applied a method adapted from Kott (1988) and Flyer (1987). The following notation is used in the steps delineated below:

n_h = the number of institutions selected and responding from stratum *h*.

$\hat{N}_h$ = the frame count of institutions in stratum *h*.

m_{hi} = the number of SSUs or students selected from institution *i* in stratum *h*.

n_h^* = the bootstrap sample size of PSUs in stratum *h* when bootstrap sampling is at the PSU level in stratum *h*.

n_{hi}^* = the number of times PSU *hi* is selected in the bootstrap sample when bootstrap sampling is at the PSU level.

m_{hi}^* = the bootstrap sample size of SSUs in PSU *hi* when bootstrap sampling is at the SSU level in stratum *h*.

m_{hij}^* = the number of times SSU *hij* is selected in the bootstrap sample when bootstrap sampling is at the SSU level.

w_{hijk}^* = the additional weight adjustment factor for student *hijk*, due to bootstrap sampling.

The process of forming replicates and computing replicate weights is as follows:

1. Approximate the stratum-level first-stage finite population correction (FPC) for the selected stratum sample, using Kott's model-based approximation (Kott 1988)

$$\mathrm{FPC}_h = \frac{\hat{N}_h - n_h}{\hat{N}_h}.$$

2. Generate a uniform (0, 1) random number R_h for each stratum *h*.

3. If $R_h \leq \mathrm{FPC}_h$, form a replicate sample in stratum *h* by randomly selecting $n_h^* = n_h - 1$ institutions with equal probability and with replacement after each selection. When n_h^* is greater than 1, a PSU may be selected more than once; in essence, n_{hi}^* may take on values of 0, 1, . . . , n_h^*. Adjust the weights by the factor

$$w_{hijk}^* = n_{hi}^* \frac{n_h}{n_h^*}.$$

4. Otherwise, form a replicate sample in stratum *h* by randomly selecting $m_{hi}^* = m_{hi} - 1$ second-stage units within each institution in stratum *h*. In this case, m_{hij}^* may take on values of 0, 1, . . . , m_{hi}^*. Adjust the weights by the factor

$$w_{hijk}^* = m_{hij}^* \frac{m_{hi}}{m_{hi}^*}.$$

5. Repeat steps 3 and 4 in all strata to form one replicate sample.

6. Steps 1 through 5 should then be repeated 200 times to form 200 replicate samples.

This method uses random switching between PSU bootstrap sampling and SSU bootstrap sampling to represent the proper mix (in expectation) of the first- and second-stage variance components when an FPC is applied at the first stage of sampling. It extends the general method described by Flyer (1987) for half-sample replication to a more general bootstrap.

This method incorporated the FPC factor only at the first stage, where sampling fractions were generally high. At the second stage, where the sampling fractions were generally low, the FPC factor was set to 1.00.

NPSAS staff used the Flyer-Kott methodology to develop a vector of bootstrap sample weights that they added to the analysis file. These weights are zero for units not selected in a particular bootstrap sample; weights for other units are inflated for the bootstrap subsampling.

The final student weight (WTA000) described in section 7.1 is used for computing estimates, such as means, percentages, and regression coefficients, and the vector of replicate weights allows for computation of additional estimates for the sole purpose of estimating variances. Assuming *B* sets of replicate weights, analysts can estimate the variance of any estimate, $\hat{\theta}$, by replicating the estimation procedure for each replicate and computing a simple variance of the replicate estimates, as follows:

$$\text{var}(\hat{\theta}) = \frac{\sum_{b=1}^{B} (\hat{\theta}_b^* - \hat{\theta})^2}{B},$$

where $\hat{\theta}_b^*$ is the estimate based on the *b*th replicate weight (where $b = 1$ to the number of replicates) and *B* is the total number of sets of replicate weights.

NPSAS staff set the number of replicate weights to 200 to ensure stable variance estimates for a variety of estimates. The student weight adjustments described in section 7.1[46] were applied to each replicate to create the 200 replicate weights included on the analysis file (WTA001–WTA200) so that the variances would be estimated to account for these weight adjustments. For some of the replicates, NPSAS staff had to loosen the bounds on the nonresponse and poststratification adjustment factors or collapse model variables because of model convergence problems (i.e., there was no solution to satisfy all model equations simultaneously). However, the model adjustments were not necessary for many replicates, and when it was necessary, the adjustments were minimal. Therefore, this approach works well for NPSAS:16 to achieve model convergence for all replicates and to minimize the effect of different models on the variance estimates.

7.3.3 Software Use for Variance Estimation

Table 63 summarizes the weight and variance estimation variables and how they are used in selected software packages that allow for Taylor series variance estimation with replacement (SUDAAN, Stata, the SAS survey data analysis procedures, IBM

[46] The institution weight adjustments cannot be replicated, due to the bootstrap methodology used.

SPSS Complex Samples, and the R survey package), Taylor series variance estimation without replacement (SUDAAN, Stata, and the R survey package), and bootstrap variance estimation (SUDAAN, Stata, the SAS survey data analysis procedures, WesVar, and the R survey package). The code shown in the table is intended for use within respective program statements or procedures and cannot be used alone as shown in the table. The code may need to be revised to be appropriate for a user's specific data file and coding decisions, and for that reason the provided code may not work for all users and require editing before it is implemented. Additionally, an example of SUDAAN code is provided in appendix K. This example code, along with the code in table 63, can be helpful in writing code in other software packages.

Table 63. Analysis weights, replicate weights, and variance estimation strata, primary sampling unit (PSU), secondary sampling unit (SSU), and PSU count variables available for NPSAS:16

Analysis weight for estimates	WTA000
Taylor series variance estimation (with replacement)	
Variance estimation stratum and PSU variables	ANALSTR and ANALPSU
Software: statements, parameters, and keywords for Taylor series variance estimation (with replacement)	
SUDAAN	DESIGN = WR WEIGHT WTA000; NEST ANALSTR ANALPSU;
Stata	svyset ANALPSU [pweight = WTA000], strata (ANALSTR) vce(LINEARIZED)
SAS survey data analysis procedures	VARMETHOD = TAYLOR WEIGHT WTA000; STRATA ANALSTR; CLUSTER ANALPSU;
IBM SPSS complex samples[1]	CSPLAN ANALYSIS /PLAN FILE='myfile.csaplan' /PLANVARS ANALYSISWEIGHT=WTA000 /DESIGN STRATA=ANALSTR CLUSTER=ANALPSU /ESTIMATOR TYPE=WR
R survey package[2]	mydesign<-svydesign(id=~ANALPSU, strata=~ANALSTR, weights=~WTA000, data=mydata)

See notes at end of table.

Table 63. Analysis weights, replicate weights, and variance estimation strata, primary sampling unit (PSU), secondary sampling unit (SSU), and PSU count variables available for NPSAS:16—Continued

Analysis weight for estimates	WTA000
Taylor series variance estimation (without replacement)	
Variance estimation stratum, PSU, SSU, and count variables	FANALSTR, FANALPSU, FANALSSU, and PSUCOUNT
Software: statements, parameters, and keywords for Taylor series variance estimation (without replacement)	
SUDAAN	DESIGN = WOR WEIGHT WTA000; NEST FANALSTR FANALPSU FANALSSU; TOTCNT PSUCOUNT _minus1_ _zero_;
Stata	svyset FANALPSU [pw=WTA000], strata(FANALSTR) fpc(PSUCOUNT) || FANALSSU, vce(LINEARIZED)
R survey package[2]	mydesign<-svydesign(id=~FANALPSU, strata=~FANALSTR, weights=~WTA000, fpc=~PSUCOUNT, data=mydata)
Bootstrap variance estimation	
Replicate weight variables	WTA001 – WTA200
Software: statements, parameters, and keywords for BRR variance estimation	
SUDAAN	DESIGN = BRR WEIGHT WTA000; REPWGT WTA001 – WTA200;
Stata	svyset [pweight=WTA000], brrweight(WTA001 - WTA200) vce(BRR) mse
SAS survey data analysis procedures	VARMETHOD = BRR WEIGHT WTA000; REPWEIGHTS WTA001 – WTA200;
WesVar	Method: BRR Full sample weight: WTA000 Replicates: WTA001 – WTA200
R survey package[2]	mydesign<-svrepdesign(type="BRR", weights=~WTA000, repweights= "WTA00[1-200]", combined.weights=FALSE, data=mydata)

[1] The name “myfile” should be replaced with the desired file name.
[2] For the R survey package (Lumley, 2014), “mydesign” can be renamed to any name for an R object to hold the specification of the survey design, and “mydata” is the name of the current dataset. For the without replacement design, the R survey package does not account for the second stage of sampling.
NOTE: To correctly estimate the variances of subpopulation estimates, use a subpopulation statement, when it exists in the software, rather than creating a subsetted dataset. The survey data analysis software specifications are given for the following versions of the software packages: SUDAAN 11.0.1, Stata 12 and newer, SAS 9.3 and newer, IBM SPSS complex samples 20, and WesVar 4.3 and newer.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16).

7.3.4 Variance Approximation

The survey design effect for a given estimate, $\hat{\theta}$, is defined as

$$\text{DEFF}(\hat{\theta}) = \frac{Var_{design}(\hat{\theta})}{Var_{srs}(\hat{\theta})}.$$

The square root of the design effect is another measure which analysts can express as the ratio of the standard errors, or

$$\text{DEFT}(\hat{\theta}) = \frac{SE_{design}(\hat{\theta})}{SE_{srs}(\hat{\theta})}.$$

Most complex multistage sampling designs, like NPSAS:16, result in design effects greater than 1.0. That is, the design-based variance is larger than the simple random sample variance. Appendix L presents design effect estimates for important survey domains and estimates for undergraduate and graduate students to summarize the effects of stratification, multistage sampling, unequal probabilities of selection, and the weight adjustments. These design effects were estimated using SUDAAN and the bootstrap variance estimation procedure described above and in appendix K.

Large design effects imply large standard errors and relatively poor precision. Small design effects imply small standard errors and good precision. In general terms, a design effect less than 2.0 is low, from 2.0 to 3.0 is moderate, and greater than 3.0 is high. Moderate and high design effects often occur in complex surveys such as NPSAS. Unequal weighting causes large design effects and is often due to nonresponse and poststratification adjustments; however, in NPSAS, the unequal weighting is also due to the sample design and different sampling rates between institution strata, as well as to the different sampling rates between student strata.

As discussed above, Taylor series linearization and replication techniques can be used to compute more precise standard errors for data from complex surveys. If statistical analyses are conducted using software packages that assume the data were collected using simple random sampling (i.e., adjustments are not made using the Taylor series or bootstrap replication methods), the standard errors will be calculated under this assumption and will be incorrect. They can be adjusted using the average DEFT, although this method is less precise than Taylor series or replication techniques. Those who must perform an analysis of NPSAS:16 data without using one of the software packages for analysis of complex survey data can use the design effect tables in appendix L to make approximate adjustments to the standard errors of survey statistics computed with the standard software packages that assume simple

random sampling designs. (For details about the use of such software packages, see table 63 and appendix K.)

As the first step in the approximation of a standard error, the analyst should normalize the overall sample weights for packages that use the weighted population size (*N*) in the calculation of standard errors (SPSS but not SAS). The normalized weight will sum to the sample size (*n*) and is calculated as

$$\text{normalized weight} = \text{weight} \times n/N$$

where *n* is the sample size (i.e., the number of cases with a valid main sampling weight) and *N* is the sum of weights.

As the second step in the approximation, the standard errors produced by the statistical software, the test statistics, or the sample weight used in analysis can be adjusted to reflect the actual complex design of the study. To adjust the standard error of an estimate, the analyst should multiply the standard error produced by the statistical software by the square root of the *DEFF* or the *DEFT*. The *DEFF* and *DEFT* used to make adjustments can be calculated for specific estimates, can be the median *DEFF* and *DEFT* across a number of variables, or can be the median *DEFF* and *DEFT* for a specific subgroup in the population. Adjusted standard errors can then be used in hypothesis testing, for example, when calculating *t* and *F* statistics.

A second option is to adjust the *t* and *F* statistics produced by statistical software packages using unadjusted (i.e., SRS) standard errors. To do this, first conduct the desired analysis weighted by the normalized weight and then divide a *t* statistic by the *DEFT* or divide an *F* statistic by the *DEFF*. A third alternative is to create a new analytic weight variable in the data file by dividing the normalized analytic weight by the DEFF and using the adjusted weight in the analyses.

7.4 Imputations

NPSAS staff imputed missing data for all variables included in the restricted-use derived file (also used in PowerStats) in accordance with mass imputation procedures described by Krotki, Black, and Creel (2005). After filling in missing data for cases where values could be deduced with certainty based upon logical or mathematical relationships among observed variables (logical imputation),[47] the weighted sequential

[47] An example of logical imputation is if a student has valid values for the total number of dependents and the number of dependent children but not the number of other dependents, the third value may be calculated as the difference of the first value minus the second value. Likewise, if a student has zero total dependents, it may be logically inferred that the student has zero dependent children.

hot deck (WSHD) method was used to replace missing data by imputing plausible values from statistically selected donor cases (stochastic imputation) (Cox 1980; Iannacchione 1982).

The first stage in the imputation procedure was to determine the pattern and level of missingness and produce an initial set of imputations. Depending on patterns of missing data, some variables that were related substantively and required imputation were grouped into blocks (vectors), and the variables within a block were imputed simultaneously (vector imputation). Then, variables and vectors were prioritized for imputation based upon their level of missing data, imputing those variables and vectors with low levels of missingness prior to imputing variables where the rate of missingness was greater. That is, variables with smaller amounts of uncertainty due to imputation were imputed first, and variables with larger amounts of uncertainty due to imputation were imputed next. For each variable and vector, NPSAS staff identified imputation classes from which donor cases for the hot deck procedure would be selected. To develop those classes, nonparametric classification or regression trees were used to identify homogeneous subgroups of item respondents (Breiman et al. 1984) using complete response variables and any previously imputed variables as possible predictor variables. Within these classes, WSHD was used to select donors.

In the second stage of imputation, for each variable or vector in the same sequence as in the first stage, the missingness was reintroduced and the missing items were reimputed. This time all complete response variables and imputed variables on the data set were available to form the imputation classes. To improve imputation quality, this previously described procedure using trees and WSHD was combined and implemented with the cyclic *p*-partition hot deck (Marker, Judkins, and Winglee 2002) technique (cycling), as discussed in Judkins (1997). This imputation approach reinforces existing patterns within the observed data. This is an iterative process, and typically, the result of cycling is a convergence to plausible values and maintenance of relationships that already exist. For NPSAS:16, there were five iterations, which improved quality without significantly slowing down the imputation process.

To reduce error due to imputation, NPSAS staff performed quality checks throughout the imputation process. In particular, NPSAS staff compared the distributions of the observed, imputed, and complete (observed and imputed) data to screen variables for further investigation. For example, the distributions of observed income and imputed income differ because the missing data are primarily for students who don't apply for federal financial aid. Those who do not apply tend to have a higher income than those who do apply. Consequently, the imputed income distribution is higher than the observed income distribution. In addition, NPSAS

staff verified that the distributions within imputation classes were similar for the observed and imputed data and concluded that the complete (observed and imputed) distribution for income was reasonable. Item response rates are shown in section 7.3.2.3, and the observed and imputed distributions for eight key variables are provided in appendix M.

7.5 Disclosure Risk Analysis and Avoidance

In preparing data files for release, NCES takes steps to minimize the likelihood that individual students participating in the study can be identified. The process of preparing the files for release includes a formal disclosure risk analysis. Every effort is made to protect the confidentiality of information about specific individuals, including performing data swapping procedures on NPSAS:16 data to minimize disclosure risk.

In data swapping, the values of the variables being swapped are exchanged between carefully selected pairs of records: a target record and a donor record. All cases were eligible for swapping. Swapping variables were selected from questionnaire, student record, and administrative record items. Perturbation was carried out under specific targeted, but undisclosed, swap rates.

Because perturbation of the NPSAS:16 data could have changed the relationships between data items, an extensive data quality check was carried out to assess and limit the impact of swapping on these relationships. For example, a set of correlations for a variety of variables was evaluated pre- and posttreatment to verify that the swapping did not greatly affect the associations.

Therefore, the modifications used to reduce the likelihood that any respondent could be identified in the data generally did not affect the overall data quality. The swapping procedures, which the IES Disclosure Review Board reviewed and approved, preserved central tendency estimates but may have resulted in slight increases in nonsampling errors.

References

Box, G.E.P., and Cox, D.R. (1964). An Analysis of Transformations. *Journal of the Royal Statistics Society, Series B*, *26:* 211–234.

Breiman, L., Friedman, J.H., Olshen, R.A., and Stone, C.J. (1984). *Classification and Regression Trees.* Toronto: Chapman and Hall.

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.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd edition). Hillsdale, NJ: Erlbaum.

Cox, B.G. (1980). The Weighted Sequential Hot Deck Imputation Procedure. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, (pp. 721–726).

Flyer, P.A. (1987). Finite Population Correction for Replication Estimates of Variance. In *Proceedings of the Section on Survey Research Methods, American Statistical Association,* Alexandria, VA.

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* (pp. 792–796). Alexandria, VA: American Statistical Association.

Hanley, J.A., and McNeil, B.J. (1982). The Meaning and Use of the Area Under a Receiver Operating Characteristic Curve. *Diagnostic Radiology, 143*(1): 29–36.

Iannacchione, V. (1982, February). *Weighted Sequential Hot Deck Imputation Macros.* Paper presented at the Seventh Annual SAS Users Group International Conference, San Francisco.

Iannacchione, V.G. (2003). Sequential Weight Adjustments for Location and Cooperation Propensity for the 1995 National Survey of Family Growth. *Journal of Official Statistics, 19*(1): 31–43.

Judkins, D.R. (1997). Imputing for Swiss Cheese Pattern of Missing Data. In *Proceedings of the Statistics Canada Symposium 97, New Directions in Surveys and Censuses*, (pp. 143–148).

Kass, G.V. (1980). An Exploratory Technique for Investigating Large Quantities of Categorical Data. *Applied Statistics, 29*(2): 119–127.

Kott, P.S. (1988). Model-Based Finite Population Correction for the Horvitz-Thompson Estimator. *Biometrika, 75*(4): 797–799.

Krotki, K., Black, S., and Creel, D. (2005). Mass Imputation. In *Proceedings of the Section on Survey Research Methods, American Statistical Association* [CD-ROM], Alexandria, VA.

Marker, D.A., Judkins, D.R., and Winglee, M. (2002). Large-Scale Imputation for Complex Surveys. In R. Groves, D. Dillman, J. Eltinge, and R. Little (Eds.), *Survey Nonresponse*. New York: John Wiley & Sons, Inc.

Radwin, D., Conzelmann, J.G., Nunnery, A., Lacy, T.A., Wu, J., Lew, S., Wine, J., and Siegel, P. (2018). *2015–16 National Postsecondary Student Aid Study (NPSAS:16): Student Financial Aid Estimates for 2015–16* (NCES 2018-466). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Riccobono, J.A., Cominole, M.B., Siegel, P.H., Gabel, T.J., Link, M.W., and Berkner, L.K. (2002). *National Postsecondary Student Aid Study 1999–2000 (NPSAS:2000) Methodology Report* (NCES 2002-152). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

RTI International. (2012). *SUDAAN User's Manual, Release 11.0*. Research Triangle Park, NC: RTI International.

Satterthwaite, F.E. (1946). An Approximate Distribution of Estimates of Variance Components. *Biometrics Bulletin, 2*(6): 110–114.

Seastrom, M. (2014). *NCES Statistical Standards* (NCES 2014-097). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved January 30, 2018, from https://nces.ed.gov/statprog/2012/.

Tukey, J.W. (1977). *Exploratory Data Analysis*. Reading, MA: Addison-Wesley.

Woodruff, R.S. (1971). A Simple Method for Approximating the Variance of a Complicated Estimate. *Journal of the American Statistical Association*, *66*: 411–414.