2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC)

Data File Documentation

NOVEMBER 2021

Peter Siegel
Nestor Ramirez
Ruby Johnson
RTI International

Tracy Hunt-White
Project Officer
National Center for Education Statistics

NCES 2022-477
U.S. DEPARTMENT OF EDUCATION

IES Institute of Education Sciences

U.S. Department of Education
Miguel Cardona
Secretary

Institute of Education Sciences
Mark Schneider
Director

National Center for Education Statistics
Peggy G. Carr
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

November 2021

The NCES Home Page address is https://nces.ed.gov.
The NCES Publications and Products address is https://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-IES-13-C-0070 with RTI International. Mention of trade names, commercial products, or organizations does not imply endorsement by the U.S. Government.

Suggested Citation
Siegel, P., Ramirez, N., and Johnson, R. (2021). *2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC) Data File Documentation* (NCES 2022-477). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved [date] from https://nces.ed.gov/pubsearch.

Content Contact
National Center for Education Statistics
NCES.Info@ed.gov

Executive Summary

The 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC), conducted by the National Center for Education Statistics (NCES) at the U.S. Department of Education's Institute of Education Sciences, is a nationally representative administrative data collection of a cross-section of postsecondary students enrolled in Title IV eligible institutions in the 2017–18 academic year in all 50 states, the District of Columbia, and Puerto Rico. NPSAS:18-AC differs from previous iterations of the National Postsecondary Student Aid Study (NPSAS) in two important ways: it is designed to be representative of undergraduate students at the state level, and it does not include a student survey as part of data collection.

Sampling Design

The target population includes all students enrolled in Title IV eligible postsecondary institutions during the 2017–18 academic year in each of the 50 states, the District of Columbia, and Puerto Rico. A sample was generated from this target population with the goal to be nationally representative as well as representative at the state level for undergraduate students. The institution eligibility conditions for NPSAS:18-AC were consistent with those in the most recent iterations of NPSAS.

NPSAS:18-AC used a two-stage sampling design. In the first stage of sampling, a total of 3,130 institutions were selected from three strata: public 2-year institutions, public 4-year institutions, and an "all other sectors" stratum. The 3,130 sample institutions included all public 2-year and all public 4-year institutions as well as a sample of 1,390 institutions from the "all other sectors" stratum. For the second stage of sampling, 349,650 students were sampled overall, of which 325,220 were undergraduate students.

Institution Data Collection Design and Outcomes

NPSAS:18-AC sampled institutions were contacted and asked to designate an institution coordinator who served as a primary point of contact for the submission of student enrollment lists and student records. These student

enrollment lists were checked for quality and completeness several times. Of the 3,080 institutions considered eligible for participation in NPSAS:18-AC, 87 percent designated an institution coordinator. A total of 2,210 (72 percent) provided usable student enrollment lists.

A student sample was generated from the provided student enrollment lists, and student records data were collected from institutions using the student records instrument. Like traditional NPSAS student records collections, the student records instrument collected student-level data on institution information, general student information, enrollment, budget, and financial aid. Institutions provided student records data using the Postsecondary Data Portal. These data were used to identify both students and institutions as student records respondents. Institutions that provided enrollment lists and were student records respondents were considered participating institutions.

A total of 1,910 institutions provided student records data, from which student records data were obtained for 78 percent of eligible student sample members. Student respondents were defined as student sample members who were student records respondents and who had at least 1 month of enrollment confirmed in any data source, including administrative data sources. The final count of student respondents was 266,910.

Administrative Records Matching Overview and Outcomes

In addition to student records data gathered from institutions, NPSAS:18-AC collected student data from administrative sources through data matching. As in prior NPSAS data collections, student data were gathered from the Central Processing System (CPS), National Student Loan Data System (NSLDS), National Student Clearinghouse (NSC), ACT, College Board, and Veterans Benefits Administration (VBA). CPS offers information collected on students' Free Application for Federal Student Aid form; this information was gathered for 62 percent of sample members for the 2017–18 academic year. NSLDS offers information on students who, at some point, had received Pell Grant or federal student loan funding. It provides data for the 2017–18 academic year as well as historical data for prior years. Of the entire sample, 193,340 students (55 percent) matched to NSLDS loan data, and 174,880 students (50 percent) matched to NSLDS Pell Grant data. NSC provides enrollment and degree records for the 2017–18 academic year. A total of 289,220 students (83 percent) matched to NSC for their NPSAS-sampled institutions. ACT and College Board provide ACT and SAT test score and questionnaire data, respectively. Of the entire sample, 127,940

(37 percent) matched to ACT while 66,060 (19 percent) matched to SAT. Lastly, VBA offers information on veterans education benefits. Of the entire sample, 23,130 students (7 percent) matched to VBA.

Data File Processing and Preparation

NPSAS:18-AC data were cleaned and edited using a multistage process conducted and developed for prior NCES studies. Quality control checks were performed on all data received from institutions during data collection. Additionally, all missing data were assessed and assigned specific values indicating the cause of missing data. Missing data were imputed for all variables included in the restricted-use derived files, including using a regression-based imputation approach to model student income to account for nonrandom missingness and imputing separately by state for state financial aid variables.

Data received from institutions and administrative sources were used to create derived variables. These variables were generated using multiple sources, prioritizing data sources specific to each item and reconciling discrepancies both within and between sources. Because NPSAS:18-AC did not include a student survey, several changes were made to the derived variables provided as part of the study. Some derived variables traditionally included in NPSAS were dropped if they included the student survey as a significant response source and could not be accounted for through alternative administrative data sources. Several total aid and aid ratio variables were also affected by the lack of a student survey, as private loan data were excluded. In some cases, administrative sources did not provide the same information needed for some derived variables as information traditionally collected from the student survey. Rather than excluding these derived variables, alternative approaches were considered and developed for these variables, and they were renamed, with the suffix "_AC" added.

As part of the data file preparation process, state representativeness was determined for undergraduate students based on several objective and subjective criteria. Out of all 50 states, the District of Columbia, and Puerto Rico, 30 states were determined to be representative for undergraduate students overall, 36 states were representative of undergraduate students in public 2-year institutions, and 45 states were representative of undergraduate students in public 4-year institutions.

Weighting and Variance Estimation

A single statistical analysis weight was computed for each student so that student respondents represent the target population. Additionally, this student weight was computed so that it could be used for analyses at the national level, within each control and level of institution, and for undergraduate-level analysis at the state level for states that were determined to be state-level representative as well as within states for sectors that were determined to be sector-level representative. Nonresponse adjustments were computed to reduce institution and student nonresponse bias. Poststratification adjustments were conducted so that the weighted student sample adequately represented the student population based on control totals obtained from Federal Student Aid data files, the National Association of State Student Grant and Aid Programs, and Integrated Postsecondary Education Data System data.

Nonresponse bias analyses were conducted for institutions, students, and items when response rates were less than 85 percent. Variables are provided in the data file to compute variances using both the Taylor-series and bootstrap replication methods.

Acknowledgments

We gratefully acknowledge the assistance of the staff members at postsecondary institutions who provided data for the student enrollment lists and student records data collections. We also extend special thanks to members of the Technical Review Panel and webinar attendees whose expertise helped shape the 2017–18 National Postsecondary Student Aid Study, Administrative Collection.

Contents

PAGE

PAGE

Appendixes

List of Tables

TABLE PAGE

TABLE PAGE

TABLE PAGE

List of Figures

Chapter 1. Overview of NPSAS:18-AC

This data file documentation report serves as a resource for users of the 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC) data. It describes the design, methods, and results of NPSAS:18-AC and outlines the differences between NPSAS:18-AC and previous iterations of the National Postsecondary Student Aid Study (NPSAS).

Chapter 1 provides an overview of NPSAS:18-AC, including the background and purpose of the study, an introduction to the study design, an explanation of differences between NPSAS:18-AC and other NPSAS, and a schedule of products. Chapter 2 contains a detailed description of the sampling design used to create the NPSAS:18-AC institution and student samples. Chapter 3 focuses on institution data collection and describes the data collection design, data collection outcomes, and institution data evaluation. Chapter 4 consists of information about administrative record matching activities including matching outcomes. Chapter 5 outlines the data processing procedures used for preparing the NPSAS:18-AC study files, including post–data collection editing and derived variable construction as well as determinations of state representativeness. Finally, chapter 6 details weighting and variance estimation procedures, weighting activities and the accuracy of estimates, data imputation activities, and data disclosure risk and avoidance. Appendix A provides a list of acronyms and abbreviations used throughout the report.

NPSAS is a complex, nationally representative cross-sectional study examining students attending postsecondary institutions that are eligible for federal financial aid in the United States. The study focuses on student enrollment in postsecondary education, with special attention to topics related to how students and families finance postsecondary education. NPSAS:18-AC is the first administrative collection of its kind, consisting exclusively of administrative data from institutions in all 50 states, the District of Columbia, and Puerto Rico.[1] Unlike prior NPSAS studies, NPSAS:18-AC does not include a student survey component. Instead, it relies entirely on other data sources including the U.S. Department of Education records on federal financial aid applications and student loan and grant programs. It is also the first nationally representative study of its

[1] From this point forward, the words “state,” “state level,” and “state representative” refer to the 50 states, the District of Columbia, and Puerto Rico unless otherwise specified.

kind designed to be representative of undergraduate students at the state level, offering policymakers, researchers, and analysts state-representative data on student financial aid. For a list of states and institution sectors within state that are representative, see section 5.4.

Tables and figures throughout this report present relevant information and outcomes from data collection and other stages of the study. 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 sample members' data. Unless otherwise stated, a probability level of .05 was used for all tests of significance conducted for NPSAS:18-AC.

1.1 Background and Purpose of NPSAS:18-AC

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 institutions in all 50 states, the District of Columbia, and Puerto Rico, and includes public and private institutions (both for-profit and nonprofit) and less-than-2-year, 2-year, and 4-year colleges and universities.

The NPSAS:18-AC iteration of NPSAS consists entirely of data from administrative sources. Data for NPSAS:18-AC were collected from postsecondary institutions, federal agencies, and other data providers. Examples of administrative data used in NPSAS:18-AC include academic records; financial aid records from postsecondary institutions; financial aid application and student loan data from federal administrative sources; and data from data providers such as National Student Clearinghouse (NSC), ACT, College Board, and Veterans Benefits Administration (VBA). NPSAS:18-AC was developed to accomplish two goals beyond those of traditional NPSAS studies. The first was to provide researchers, policymakers, practitioners, and other data analysts with relevant postsecondary financing and enrollment data between the traditional NPSAS studies. The NPSAS:18-AC student cohort is a cross-section of postsecondary students enrolled in Title IV eligible institutions in the 2017–18 academic year, falling between the 2015–16 National Postsecondary Student Aid Study (NPSAS:16) and the 2019–20 National Postsecondary Student Aid Study (NPSAS:20). Second, NPSAS:18-AC was designed to be not only nationally representative but also representative of undergraduate students at the state level. This allows data users to use

NPSAS:18-AC to address a wide variety of research questions that would otherwise not be feasible with other NPSAS studies.

NPSAS:18-AC, as well as other NPSAS studies, are authorized by the following legislation:

- the Higher Education Opportunity Act (HEOA) of 2008, 20 U.S.C. § 1015(A)(k); and
- the Education Sciences Reform Act (ESRA) of 2002, 20 U.S.C. § 9543.

1.2 Overview of NPSAS:18-AC Design

The data collected for NPSAS:18-AC come from two sources: (1) postsecondary institutions and (2) administrative data records. The target population includes all students enrolled in Title IV eligible postsecondary institutions during the 2017–18 academic year in each of the 50 states, the District of Columbia, and Puerto Rico. A sample was generated from this target population in order to be nationally representative as well as representative at the state level for undergraduate students.

All student-level data were gathered through institutions and administrative data records. Institutions were requested to send student-level data on enrollment patterns and financial aid. Some of the student financial aid data elements that were received from institutions were also available for verification from the National Student Loan Data System (NSLDS) and the Central Processing System (CPS). NSLDS hosts student data on federal student loans and federal grants such as the Pell Grant. CPS contains data from the Free Application for Federal Student Aid (FAFSA), which is used to determine eligibility for federal aid. NPSAS staff used file matching with CPS and NSLDS data to reduce the data collection burden on sampled institutions by minimizing the number of data elements requested from sampled institutions and utilizing matched administrative data to derive variables wherever possible. Student samples were drawn from the student enrollment lists provided by institutions after confirming study eligibility of students in the enrollment lists. Student data provided by institutions for the NPSAS:18-AC student sample were supplemented by administrative sources. In addition to CPS and NSLDS, student data were supplemented using data from NSC, ACT, College Board, and VBA.

1.3 Differences Between NPSAS and NPSAS:18-AC

NPSAS:18-AC differs from traditional NPSAS studies in several ways. Unlike other NPSAS studies, NPSAS:18-AC data were collected from administrative sources only. No student survey was conducted as part of the NPSAS:18-AC data collection activities. Accordingly, NPSAS:18-AC relies only on federal data, institution records, and external data providers like NSC, ACT, and College Board, to compile information about postsecondary students. NPSAS:18-AC also differs from traditional NPSAS studies by diverging from the 4-year schedule other NPSAS studies have followed. As previously mentioned, NPSAS:18-AC data—collected between NPSAS:16 and NPSAS:20—cover postsecondary students during the 2017–18 academic year. Due to the absence of a student survey, many derived variables that are traditionally included in NPSAS studies were either modified or not included as part of NPSAS:18-AC. Further detail on changes to derived variable construction are provided in section 5.3.

NPSAS:18-AC does not have a longitudinal component, which separates it from traditional NPSAS studies. Data collected as part of NPSAS traditionally serve as the base year for longitudinal studies of postsecondary students. The Beginning Postsecondary Students Longitudinal Study (BPS) and the Baccalaureate and Beyond Longitudinal Study (B&B) sample students from NPSAS studies and follow them over time, with each NPSAS collection alternately serving as the basis for either BPS or B&B—NPSAS:16 served as the base-year cohort for the 2016/17 Baccalaureate and Beyond Longitudinal Study (B&B:16/17), and NPSAS:20 will serve as the base-year cohort for the 2020/22 Beginning Postsecondary Students Longitudinal Study (BPS:20/22). NPSAS:18-AC will not be used as a base-year cohort for a National Center for Education Statistics (NCES) postsecondary longitudinal study.

NPSAS:18-AC was also designed to contain undergraduate student data that would be state representative as well as nationally representative. As mentioned earlier, students were sampled from institutions in all 50 states, the District of Columbia, and Puerto Rico. This sample was generated to provide estimates that describe postsecondary undergraduate students in these states. Other NPSAS studies are not designed to provide representative data at the state level, with the exceptions of the 2003–04 National Postsecondary Student Aid Study (NPSAS:04) and the 2007–08 National Postsecondary Student Aid Study (NPSAS:08), which were representative of 12 and 6 states, respectively. NPSAS:18-AC provides researchers, policymakers, and analysts a unique and rich data source for studying state-level undergraduate educational issues that prior NPSAS studies may not have been suited to examine.

1.4 Schedule and Products

Table 1 shows the schedule for the major activities of NPSAS:18-AC, including student enrollment list collection, institution and administrative records collection, and dissemination of reports and data files. Products generated from NPSAS:18-AC include restricted-use data files along with electronic codebooks and PowerStats, included in the DataLab suite of tools developed by NCES for use by education researchers.

Table 1. Schedule of major activities for NPSAS:18-AC: 2018–21

NPSAS:18-AC activity	Start date	End date
Contact institutions to request student enrollment lists	4/19/2018	11/30/2018
Collect student enrollment lists	5/10/2018	1/15/2019
Select student sample	5/10/2018	1/18/2019
Collect student data from institution records	7/10/2018	5/15/2019
Collect student data from administrative records	5/10/2018	9/9/2019
Process data, construct data files	7/30/2018	2/19/2021
Prepare/update reports	3/18/2019	11/30/2021

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

Additionally, data from NPSAS:18-AC were used for a First Look report that provides descriptive information about the NPSAS:18-AC sample both nationally and for states with sufficient data to be representative of undergraduate students at the state level. This report, along with publications for prior NPSAS studies, is available at https://nces.ed.gov/surveys/npsas. NPSAS:18-AC microlevel 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:18-AC 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:18-AC. All documented procedures and methods were developed and refined in consultation with a Technical Review Panel (TRP) composed of nationally recognized experts in higher education, staff from NCES, and representatives of other federal agencies.[2]

2.1 Respondent Universe

NPSAS:18-AC 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:18-AC, NPSAS statisticians used institution data collected from various Integrated Postsecondary Education Data System (IPEDS) surveys. The student sampling frame included all students from the sampled institutions who met eligibility requirements. Student eligibility requirements are discussed in section 2.1.2.

2.1.1 Institution Universe

The NPSAS:18-AC institution (first-stage) sampling frame included all levels (less-than-2-year, 2-year, and 4-year) and control classifications (public, private nonprofit, and private for-profit) of Title IV eligible postsecondary institutions in the United States. To be eligible for NPSAS:18-AC, 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 administered the institution;
- been located in the 50 states, the District of Columbia, or Puerto Rico;
- not been a U.S. service academy institution; and

[2] See appendix B for a complete list of TRP participants.

- 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. The seven U.S. service academies (U.S. Air Force Academy, U.S. Coast Guard Academy, U.S. Naval Academy, U.S. Merchant Marine Academy, U.S. Military Academy, Naval Postgraduate School, and Air Force Institute of Technology-Graduate School) were also excluded because of the academies' unique funding/tuition base.

The institution eligibility conditions for NPSAS:18-AC were consistent with those in 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 the 1999–2000 National Postsecondary Student Aid Study (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 institutions were also eligible to distribute federal Title IV student aid, were first included in NPSAS:04. Finally, although institutions in Puerto Rico were not included in the 1986–87 National Postsecondary Student Aid Study (NPSAS:87) and 2011–12 National Postsecondary Student Aid Study (NPSAS:12), they were included in NPSAS:18-AC and all other administrations of NPSAS.

2.1.2 Student Universe

The student (second-stage) sampling frame is described below. The requirements for NPSAS student eligibility have largely remained constant over time. For NPSAS:18-AC, the target population consisted of all eligible students who were enrolled at any time between July 1, 2017, and June 30, 2018,[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

[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] In prior iterations of NPSAS, enrollment lists covered the period of July 1 through April 30 so as to not delay data collection. Although the data collection schedule for NPSAS:18-AC was later than usual, to be consistent with past iterations of NPSAS, the July 1 through April 30 dates did not change. Any lack of coverage resulting from the truncated enrollment period is accounted for by the poststratification weight adjustment, which is described in chapter 6.

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 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

The NPSAS:18-AC full-scale institution frame was constructed from the IPEDS 2016–17 Institutional Characteristics Header (IC-H), 2016–17 Institutional Characteristics (IC), 2015–16 12-month Enrollment (E12), and 2016–17 Completions (C) files. Because several individual private for-profit institutions and large chains of private for-profit institutions closed or were sold during the sampling design process, web searches were performed and articles were read to identify and exclude institutions that were still in IPEDS but were no longer eligible for NPSAS:18-AC.

The institution strata were based on sectors (combinations of control and level of institution) within each of the 52 states and territories, for a total of 154 sampling strata (52 x 3 - 2):[5]

1. public 2-year;
2. public 4-year; and
3. all other sectors, including public less-than-2-year, private nonprofit (all levels), and private for-profit (all levels).

The sample sizes presented in table 2 allowed for state-representative undergraduate student samples for public 2-year and public 4-year institutions as well as overall. The sample is also nationally representative for both undergraduate and graduate students.

[5] Two states had no public 2-year institutions during sampling. One state, Alaska, had one institution that was sampled as a public 2-year institution but was recoded to a public 4-year institution after sampling due to updated IPEDS institution characterization for this institution. Tables related to student sample selection in the DFD will provide sampling information for Alaska public 2-year institutions, but otherwise Alaska will be described as having no public 2-year institutions.

Table 2. Size of universe and number of institutions sampled, by institution stratum and state: 2017–18

State	All sectors		Institution stratum[1]					
			Public 2-year[2]		Public 4-year		All other sectors	
	Size of universe[3]	Sample size	Size of universe[3]	Sample size	Size of universe[3]	Sample size	Size of universe[3]	Sample size
All states	**6,710**	**3,130**	**990**	**990**	**750**	**750**	**4,980**	**1,390**
Alabama	90	70	30	30	10	10	50	30
Alaska	10	10	#	#	#	#	10	10
Arizona	120	60	20	20	10	10	90	30
Arkansas	80	60	20	20	10	10	50	30
California	660	180	110	110	50	50	510	30
Colorado	110	60	10	10	20	20	80	30
Connecticut	80	50	10	10	10	10	60	30
Delaware	20	20	†	†	#	#	10	10
District of Columbia	20	20	†	†	#	#	20	20
Florida	360	100	40	40	40	40	290	30
Georgia	160	80	20	20	30	30	110	30
Hawaii	30	30	10	10	#	#	20	20
Idaho	40	40	#	#	#	#	30	30
Illinois	270	90	50	50	10	10	210	30
Indiana	120	50	#	#	20	20	100	30
Iowa	90	50	20	20	#	#	70	30
Kansas	80	60	30	30	10	10	50	30
Kentucky	100	50	20	20	10	10	70	30
Louisiana	120	60	20	20	20	20	90	30
Maine	40	40	10	10	10	10	20	20
Maryland	90	60	20	20	10	10	60	30
Massachusetts	180	60	20	20	20	20	150	30
Michigan	170	80	30	30	20	20	120	30
Minnesota	110	70	30	30	10	10	70	30
Mississippi	60	60	20	20	10	10	30	30
Missouri	180	60	20	20	10	10	150	30
Montana	30	30	10	10	10	10	10	10
Nebraska	50	50	10	10	10	10	30	30
Nevada	40	40	#	#	10	10	40	40
New Hampshire	40	40	10	10	10	10	30	30
New Jersey	150	60	20	20	10	10	120	30
New Mexico	50	50	20	20	10	10	20	20
New York	450	110	40	40	40	40	360	30
North Carolina	180	110	60	60	20	20	100	30
North Dakota	30	30	10	10	10	10	20	20
Ohio	310	100	40	40	40	40	240	30
Oklahoma	130	70	30	30	20	20	90	30
Oregon	90	60	20	20	10	10	60	30
Pennsylvania	360	90	20	20	50	50	300	30
Puerto Rico	150	50	10	10	10	10	130	30
Rhode Island	20	20	#	#	#	#	20	20
South Carolina	110	60	20	20	10	10	70	30
South Dakota	30	30	10	10	10	10	20	20
Tennessee	170	80	40	40	10	10	120	30
Texas	410	140	60	60	50	50	300	30
Utah	70	40	#	#	10	10	60	30
Vermont	30	30	#	#	10	10	20	20

See notes at end of table.

Table 2. Size of universe and number of institutions sampled, by institution stratum and state: 2017–18—Continued

State	All sectors		Institution stratum[1]					
			Public 2-year[2]		Public 4-year		All other sectors	
	Size of universe[3]	Sample size	Size of universe[3]	Sample size	Size of universe[3]	Sample size	Size of universe[3]	Sample size
Virginia	160	70	20	20	20	20	120	30
Washington	110	70	10	10	30	30	70	30
West Virginia	80	50	10	10	10	10	50	30
Wisconsin	100	60	20	20	20	20	70	30
Wyoming	10	10	10	10	#	#	#	#

† Not applicable.
# Rounds to zero.
[1] Institution stratum reflects institution categorization of control and level of institution as determined from the 2015–16 Integrated Postsecondary Education Data System (IPEDS) files.
[2] Alaska had one institution that was sampled as a public 2-year institution but was recoded to a public 4-year institution after sampling due to updated IPEDS institution characterization for this institution. There were no public 2-year schools in Delaware or the District of Columbia.
[3] Based on the 2016–17 IPEDS files.
NOTE: Sample sizes rounded to the nearest 10. Detail may not sum to totals because of rounding. For the purposes of this table, Puerto Rico and the District of Columbia are referred to as "states."
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

Overall, 3,130 institutions were selected that included a census of all public 2-year and all public 4-year institutions as well as a sample of 1,390 institutions from the "all other sectors" stratum. Within the "all other sectors" stratum, the following criteria were used to determine institution sample sizes:

- In states with 36 or fewer institutions in the "all other sectors" stratum, a census of all institutions was selected.
- In states with more than 36 institutions in the "all other sectors" stratum, 30 institutions were selected.

Within the "all other sectors" stratum, institutions were selected using stratified random sampling with probabilities proportional to a composite measure of size (Folsom, Potter, and Williams 1987). This is the same methodology used for NPSAS since the 1995–96 National Postsecondary Student Aid Study (NPSAS:96). Institution measures of size were determined using the most recent IPEDS E12 data on the number of students enrolled. Using a composite measure of enrollment size ensured that the target sample sizes were achieved within institution and student sampling strata while also achieving approximately equal student weights across institutions.

Within the "all other sectors" stratum, additional implicit stratification[6] was accomplished by sorting the sampling frame by the following classifications:

[6] *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.

1. control and level of institution;

4. Historically Black Colleges and Universities (HBCUs) status;

5. Hispanic-Serving Institution (HSI) status;[7]

6. Carnegie Classifications of postsecondary institutions; 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 3 shows the counts of sampled and eligible institutions by control and level of institution.[8] Almost all the 3,130 sampled institutions met the eligibility requirements.

[7] An HSI indicator is no longer available from IPEDS, so an HSI proxy was created following the definition of HSI provided by the U.S. Department of Education (https://www2.ed.gov/programs/idueshsi/definition.html) and using IPEDS Hispanic enrollment data.
[8] Control and level of institution are based on information from the sampling frame, which was formed from the IPEDS 2016–17 IC-H, 2016–17 IC, 2016–17 C, and 2015–16 E12 files.

Table 3. Number of sampled and eligible institutions, by control and level of institution: 2017–18

Control and level of institution[1]	Sampled institutions	Eligible institutions
All institutions	**3,130**	**3,080**
Control of institution		
Public	1,770	1,760
Private nonprofit	770	760
Private for-profit	590	560
Level of institution		
Less-than-2-year	230	220
2-year	1,230	1,210
4-year, non-doctorate-granting	920	900
4-year, doctorate-granting	750	750
Control and level of institution		
Public less-than-2-year	40	40
Public 2-year	970	960
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	150	150
Public 4-year, non-doctorate-granting, primarily baccalaureate	230	230
Public 4-year, doctorate-granting	380	380
Private nonprofit 2-year or less	40	40
Private nonprofit 4-year, non-doctorate-granting	390	380
Private nonprofit 4-year, doctorate-granting	350	350
Private for-profit less-than-2-year	190	180
Private for-profit 2-year	210	200
Private for-profit 4-year	190	180

[1] Control and level of institution are based on data from the sampling frame, which was formed from the Integrated Postsecondary Education Data System 2016–17 Institutional Characteristics Header file.
NOTE: Sample sizes rounded to the nearest 10. Percentages are based on the unrounded count of eligible institutions within the row under consideration. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

2.3 Student Sample

Each sampled institution verified as NPSAS eligible was asked to provide a complete list of students who were enrolled between July 1, 2017, and April 30, 2018, and satisfied all NPSAS eligibility conditions. These lists included information needed to identify students for matching to administrative records and for classifying students to create sampling strata. 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);
- degree program (undergraduate certificate, associate's degree, bachelor's degree, master's degree, doctoral–research/scholarship/other, doctoral–professional practice, other graduate);
- class level (first year, second year, etc.);

- date of birth (DOB);
- Classification of Instructional Programs (CIP) code or major;
- high school/completion program completion date (month and year);
- enrollment in high school (or completion program);
- date of first enrollment (at the postsecondary level);
- veteran status;
- race;
- ethnicity;
- sex; and
- contact information (local and permanent street address and phone number and school and home e-mail address).

The 10 student sampling strata were

1. undergraduate students;
2. graduate students who were veterans;
3. master's degree students in science, technology, engineering, and mathematics (STEM) programs;
4. master's degree students in education and business programs;
5. master's degree students in other programs;
6. doctoral–research/scholarship/other students in STEM programs;
7. doctoral–research/scholarship/other students in education and business programs;
8. doctoral–research/scholarship/other students in other programs;
9. doctoral–professional practice students; and
10. other graduate students.

If students fell into multiple strata, such as graduate students who were veterans, 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 (table 4) at a typical rate would have made it difficult to draw inferences about the experiences of master's degree and doctoral students. Undergraduate students who were veterans were not

oversampled within each state because that would have required too large a total sample size.

Table 4. Oversampled and undersampled student groups: 2017–18

Oversampled student groups	Undersampled student groups
Graduate students who were veterans	Master's degree students in education and business programs
Doctoral–research/scholarship/other students in science, technology, engineering, and mathematics (STEM) programs	Doctoral–research/scholarship/other students in education and business programs
Master's degree students in STEM programs	
Master's degree students in other programs	

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

Before sampling, records for students on the enrollment lists were matched to two administrative data sources—VBA and NSLDS. To identify and sample veterans, SSNs were sent from the student enrollment lists to VBA for record matching. Graduate students who were identified as veterans were placed in the graduate students' veterans stratum. Student SSNs from enrollment lists were also matched to NSLDS data. The match results identified federal student financial aid recipients.

After matching to VBA and NSLDS data was completed, students were randomly selected by means of stratified systematic sampling with predetermined sampling rates that varied by student stratum. Within the undergraduate student stratum, individuals were sorted by whether or not they were veterans and whether or not they received federal aid. Within each graduate student stratum, individuals were sorted by whether or not they received federal aid. Students were then systematically sampled so that the number of undergraduate veterans and aided and unaided sampled students approximately matched the population proportions of these students within the institution and student strata. This implicit stratification was done to help produce more accurate undergraduate veterans and financial aid estimates. Greater detail on VBA and NSLDS matching can be found in chapter 4.

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 6) accounted for students who had more than one chance of being selected because they had attended multiple institutions during the 2017–18 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 (appendix C). In certain instances, sampling rates were modified as follows:

- Student sampling rates were increased for each institution to yield at least 60 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 600 students.[9]
- 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 6.3). Table 5 shows the expected and achieved sample sizes by student sampling type and control and level of institution.

[9] Because of their large enrollments, 40 institutions had a student sample size greater than 600.

Table 5. Number of expected and achieved sample students and percentage achieved, by student sampling type and control and level of institution: 2017–18

Control and level of institution	All students			Student sampling type: Undergraduates			Student sampling type: Graduates		
	Number expected	Number achieved	Percent achieved	Number expected	Number achieved	Percent achieved	Number expected	Number achieved	Percent achieved
All institutions	**400,020**	**349,650**	**87.4**	**375,020**	**325,220**	**86.7**	**25,000**	**24,430**	**97.7**
Public less-than-2-year	1,400	1,060	76.2	1,400	1,060	76.2	†	†	†
Public 2-year	126,500	108,660	85.9	126,500	108,590	85.8	†	70	†
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	18,310	17,880	97.6	18,230	17,740	97.3	80	140	173.1
Public 4-year, non-doctorate-granting, primarily baccalaureate	27,160	26,480	97.5	25,650	24,960	97.3	1,510	1,520	100.7
Public 4-year, doctorate-granting	112,800	99,160	87.9	106,540	91,600	86.0	6,270	7,560	120.6
Private nonprofit less-than-2-year	1,360	1,120	82.4	1,360	1,120	82.4	†	†	†
Private nonprofit 4-year, non-doctorate-granting	32,820	32,620	99.4	29,120	29,130	100.0	3,700	3,490	94.4
Private nonprofit 4-year, doctorate-granting	42,090	36,280	86.2	35,900	29,760	82.9	6,190	6,520	105.3
Private for-profit less-than-2-year	5,320	6,380	120.1	5,320	6,380	120.1	†	†	†
Private for-profit 2-year	6,390	5,970	93.5	6,390	5,970	93.5	†	†	†
Private for-profit 4-year	25,880	14,050	54.3	18,620	8,910	47.9	7,260	5,140	70.8

† Not applicable.
NOTE: Some institution classifications of student type on the enrollment lists (e.g., undergraduate, graduate) were updated over the course of student sampling; 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, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

Table 6 shows the number of expected and achieved sample sizes by student stratum. The achieved sample size of 349,650 was lower than the expected 400,020 because many of the enrollment lists, particularly from private for-profit institutions, contained fewer students than expected based on initial IPEDS counts. Additionally, fewer enrollment lists than expected were received compared with prior NPSAS studies. Sampling rates were adjusted, as described above, but the sample size still fell short of the target. Overall, more students in the other graduate stratum were selected into the sample than planned (for further details about sample allocation, see appendix C). Table 7 shows the initial classification of the student sample by student type and control and level of institution.

Table 6. Expected and achieved student sample size and percentage achieved, by student stratum: 2017–18

	Student sample size		
Student stratum[1]	Number expected[2]	Number achieved[3]	Percent achieved[4]
Total	400,020	349,650	87.4
Undergraduate students	375,020	325,220	86.7
Graduate students	25,000	24,430	97.7
Veterans	3,690	3,810	103.2
Master's degree students in STEM programs	2,510	2,020	80.2
Master's degree students in education or business programs	3,120	3,220	103.4
Master's degree students in other programs	4,260	3,410	80.1
Doctor's–research/scholarship/other students in STEM programs	2,550	2,530	99.4
Doctor's–research/scholarship/other students in education or business programs	3,020	3,180	105.5
Doctor's–research/scholarship/other students in other programs	2,190	1,710	78.1
Doctor's–professional practice	2,710	3,120	115.1
Other graduate[5]	950	1,420	149.9

[1] Some institution classifications of student type on the enrollment lists (e.g., undergraduate, 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 Integrated Postsecondary Education Data System 2016–17 12-month Enrollment and Completions counts.
[3] The student sample was drawn from 2,210 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: STEM = science, technology, engineering, and mathematics. 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, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

Table 7. Initial classification of NPSAS:18-AC student sample, by student type and control and level of institution: 2017–18

Control and level of institution	Total sample[1]		Student type			
			Undergraduates		Graduates	
	Number	Percent	Number	Percent	Number	Percent
All institutions	**349,650**	**100.0**	**325,220**	**100.0**	**24,430**	**100.0**
Control of institution						
Public	253,230	72.4	243,950	75.0	9,280	38.0
Private nonprofit	70,010	20.0	60,000	18.5	10,010	41.0
Private for-profit	26,410	7.6	21,270	6.5	5,140	21.0
Level of institution						
Less-than-2-year	7,530	2.2	7,530	2.3	†	†
2-year	115,670	33.1	115,600	35.5	70	0.3
4-year, non-doctorate-granting	84,620	24.2	77,700	23.9	6,930	28.3
4-year, doctorate-granting	141,840	40.6	124,400	38.3	17,440	71.4
Control and level of institution						
Public less-than-2-year	1,060	0.3	1,060	0.3	†	†
Public 2-year	108,660	31.1	108,590	33.4	70	0.3
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	17,880	5.1	17,740	5.5	140	0.6
Public 4-year, non-doctorate-granting, primarily baccalaureate	26,480	7.6	24,960	7.7	1,520	6.2
Public 4-year, doctorate-granting	99,160	28.4	91,600	28.2	7,560	30.9
Private nonprofit 2-year or less	1,120	0.3	1,120	0.3	†	†
Private nonprofit 4-year, non-doctorate-granting	32,620	9.3	29,130	9.0	3,490	14.3
Private nonprofit 4-year, doctorate-granting	36,280	10.4	29,760	9.1	6,520	26.7
Private for-profit less-than-2-year	6,380	1.8	6,380	2.0	†	†
Private for-profit 2-year	5,970	1.7	5,970	1.8	†	†
Private for-profit 4-year	14,050	4.0	8,910	2.7	5,140	21.0

† Not applicable.
[1] The student sample was drawn from 2,210 eligible institutions that provided enrollment lists.
NOTE: The statistics presented in this table are based on the original student sampling frame classification, not on the student's final classification. The graduate veterans stratum, three master's strata, four doctoral strata, and other graduate stratum have been combined. 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, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

The expected and achieved undergraduate sample sizes for the total undergraduate sample of 325,220 students, by institution stratum and state, are shown in table 8.

Table 8. Expected and achieved undergraduate student sample sizes and percentage achieved, by institution stratum and state: 2017–18

State	All sectors			Institution stratum[1]								
				Public 2-year			Public 4-year			All other sectors		
	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 states	**375,020**	**325,220**	**86.7**	**126,500**	**108,590**	**85.8**	**150,420**	**134,300**	**89.3**	**98,100**	**82,340**	**83.9**
Alabama	7,210	5,790	80.2	2,580	2,280	88.6	3,210	2,420	75.2	1,420	1,090	76.8
Alaska[2]	7,210	6,950	96.3	520	600	114.2	6,240	6,190	99.2	450	170	37.8
Arizona	7,210	5,440	75.4	2,820	2,720	96.5	1,450	1,290	88.4	2,940	1,430	48.6
Arkansas	7,210	6,100	84.5	2,790	2,580	92.4	3,560	2,530	71.0	860	990	115.1
California	7,210	6,940	96.2	4,030	4,110	101.9	2,070	1,990	96.1	1,110	840	75.7
Colorado	7,210	6,610	91.7	1,900	2,110	111.3	3,490	3,280	94.0	1,820	1,220	67.0
Connecticut	7,210	6,180	85.7	2,400	2,240	93.4	1,980	1,470	74.2	2,830	2,470	87.3
Delaware[2]	7,210	7,080	98.1	†	†	†	5,200	5,200	100.0	2,010	1,880	93.5
District of Columbia[2]	7,210	4,810	66.7	†	†	†	1,190	600	50.2	6,020	4,210	69.9
Florida	7,210	6,480	89.9	1,440	1,460	101.1	4,090	3,940	96.3	1,680	1,080	64.3
Georgia	7,210	6,220	86.3	2,020	1,000	49.6	3,770	4,060	107.6	1,420	1,160	81.7
Hawaii	7,210	7,110	98.6	3,360	3,410	101.5	2,510	2,500	99.6	1,340	1,200	89.6
Idaho	7,210	4,430	61.4	1,510	1,210	80.0	2,320	1,620	69.6	3,380	1,600	47.3
Illinois	7,210	6,310	87.5	4,020	3,410	85.0	1,460	1,490	102.2	1,740	1,410	81.0
Indiana	7,210	7,350	102.0	2,380	2,370	99.8	3,300	3,280	99.5	1,530	1,700	111.1
Iowa	7,210	7,050	97.8	3,130	2,990	95.3	1,580	1,560	99.0	2,500	2,500	100.0
Kansas	7,210	5,640	78.3	3,510	2,300	65.6	2,380	2,300	96.7	1,320	1,040	78.8
Kentucky	7,210	5,520	76.6	2,840	1,610	56.7	3,080	2,620	85.1	1,290	1,290	100.0
Louisiana	7,210	6,600	91.5	2,500	2,010	80.5	3,510	3,250	92.7	1,200	1,330	110.8
Maine	7,210	5,140	71.2	2,100	1,380	65.6	2,750	1,990	72.1	2,360	1,770	75.0
Maryland	7,210	5,940	82.3	3,310	2,630	79.4	3,000	2,500	83.3	900	820	91.1
Massachusetts	7,210	7,020	97.4	2,040	1,750	85.7	1,790	1,820	101.6	3,380	3,450	102.1
Michigan	7,210	6,530	90.6	2,560	2,400	93.8	3,400	3,310	97.5	1,260	820	65.1
Minnesota	7,210	6,940	96.3	3,110	3,280	105.5	2,340	2,280	97.4	1,760	1,380	78.4
Mississippi	7,210	6,070	84.1	3,700	2,830	76.6	2,740	2,420	88.3	770	810	105.2
Missouri	7,210	6,160	85.4	2,250	1,820	80.9	2,440	2,030	83.3	2,530	2,310	91.3
Montana	7,210	5,900	81.8	1,670	1,450	86.6	4,760	3,900	81.9	780	550	70.5
Nebraska	7,210	5,200	72.1	3,150	2,470	78.5	2,510	1,200	47.8	1,550	1,530	98.7
Nevada	7,210	7,690	106.6	1,350	2,160	159.6	5,060	4,470	88.4	810	1,070	132.1
New Hampshire	7,210	6,190	85.8	1,380	1,230	88.9	1,350	1,390	103.0	4,480	3,570	79.7

See notes at end of table.

Table 8. Expected and achieved undergraduate student sample sizes and percentage achieved, by institution stratum and state: 2017–18—Continued

State	All sectors			Institution stratum[1]								
				Public 2-year			Public 4-year			All other sectors		
	Number expected	Number achieved	Percent achieved	Number expected	Number achieved	Percent achieved	Number expected	Number achieved	Percent achieved	Number expected	Number achieved	Percent achieved
New Jersey	7,210	6,530	90.5	3,320	2,870	86.3	2,500	2,520	100.8	1,390	1,150	82.7
New Mexico	7,210	6,240	86.5	4,440	3,860	86.9	2,350	1,780	75.5	420	600	142.9
New York	7,210	6,820	94.6	2,320	2,310	99.4	2,080	2,220	106.8	2,810	2,290	81.5
North Carolina	7,210	7,120	98.7	3,670	3,400	92.7	2,280	2,300	100.9	1,270	1,420	111.8
North Dakota	7,210	5,120	71.0	1,290	550	42.4	5,090	3,990	78.3	830	590	71.1
Ohio	7,210	7,000	97.0	2,480	2,580	104.2	3,030	3,020	99.6	1,710	1,400	81.9
Oklahoma	7,210	6,200	86.0	2,620	2,380	91.0	3,250	3,090	95.3	1,350	730	54.1
Oregon	7,210	5,760	79.8	3,870	3,040	78.7	2,520	1,750	69.5	830	960	115.7
Pennsylvania	7,210	6,580	91.2	1,790	1,700	94.8	2,390	2,600	108.7	3,030	2,280	75.2
Puerto Rico	7,210	5,480	75.9	1,060	60	5.7	1,440	1,000	69.5	4,710	4,410	93.6
Rhode Island	7,210	7,560	104.9	1,790	1,790	100.0	2,000	2,000	100.0	3,420	3,770	110.2
South Carolina	7,210	5,560	77.1	3,210	1,740	54.3	2,550	2,370	93.2	1,450	1,440	99.3
South Dakota	7,210	6,160	85.4	1,260	1,280	101.5	4,650	3,960	85.3	1,300	910	70.0
Tennessee	7,210	6,020	83.5	2,680	2,520	94.1	2,530	1,810	71.7	2,010	1,690	84.1
Texas	7,210	7,430	103.0	3,940	4,040	102.4	2,420	2,510	103.8	860	890	103.5
Utah	7,210	6,400	88.7	1,430	1,390	97.0	3,040	3,070	101.2	2,750	1,940	70.5
Vermont	7,210	5,970	82.8	1,450	1,430	98.6	3,030	2,470	81.4	2,730	2,080	76.2
Virginia	7,210	6,570	91.1	2,980	2,510	84.2	2,250	2,260	100.7	1,990	1,800	90.5
Washington	7,210	6,010	83.4	1,530	1,230	80.6	4,860	3,520	72.4	820	1,260	153.7
West Virginia	7,210	5,740	79.5	1,390	720	52.0	2,160	2,160	100.3	3,670	2,850	77.7
Wisconsin	7,210	6,420	89.0	2,440	2,120	87.2	3,600	3,140	87.1	1,170	1,160	99.1
Wyoming	7,210	5,180	71.8	5,180	3,250	62.8	1,890	1,880	99.4	150	50	33.3

† Not applicable.
[1] Institution stratum reflects institution categorization as determined from the 2015–16 Integrated Postsecondary Education Data System files.
[2] Alaska had one institution that was sampled as a public 2-year institution but was recoded to a public 4-year institution after sampling due to updated IPEDS institution characterization for this institution. There were no public 2-year institutions in Delaware or the District of Columbia
NOTE: Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding. For the purposes of this table, Puerto Rico and the District of Columbia are referred to as "states."
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

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

The following chapter describes the design, implementation, and outcomes of institution data collection for NPSAS:18-AC. It includes details 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:18-AC institution data were collected in several stages using systems designed for contacting sample members and managing data collection processes. NPSAS project staff were trained in using an institution contacting system (ICS) to record data on any communications with institution staff. The Postsecondary Data Portal (PDP) is a website that was used for data collection and served as a single location for consolidating information about the study and contact information for the help desk and project staff.

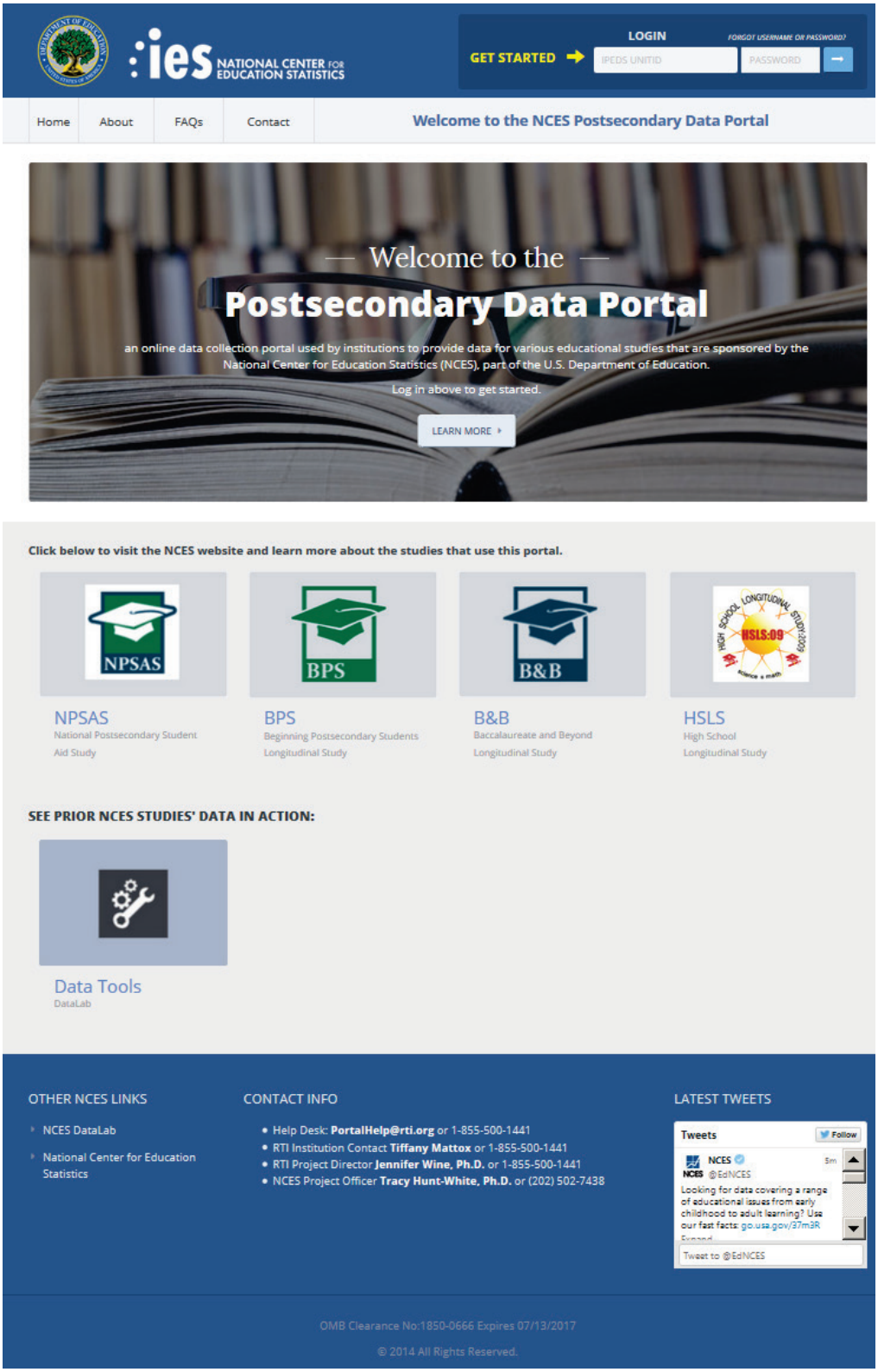
3.1.1 Institution Contacting System (ICS)

The ICS was used 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 or e-mail). The reporting functions of the ICS allowed for the viewing of the overall progress of institution recruitment, enrollment list collection, and student records collection.

3.1.2 Postsecondary Data Portal (PDP)

All institution data collection was conducted through the PDP (figure 1). 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 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. Postsecondary Data Portal home page: 2017–18



SOURCE: U.S. Department of Education, National Center for Education Statistics, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

Before NPSAS:18-AC student records data collection, changes were made to the PDP to enhance usability and the quality of the data collected. For example, instructional videos were added to the PDP. These videos showed institution staff how to provide data in each instrument mode.

3.1.3 Student Records Collection System

The NPSAS:18-AC student records instrument collected student-level data for the 2017–18 academic year. The instrument was organized into five sections:

1. Institution Information (e.g., institution's academic calendar for the period of July 1, 2017, to June 30, 2018);

2. General Student Information (e.g., student demographics, contact information, and other characteristics);

3. Enrollment (e.g., students' degree program, major[s], class level, and enrollment intensity in the 2017–18 academic year);

4. Budget (e.g., estimated costs of attendance); and

5. Financial Aid (e.g., student financial aid received by sampled students for the 2017–18 academic year, including federal, state, institution, graduate, private, other government, and other awards).

To minimize burden on institutions, the NPSAS:18-AC student records instrument maintained consistency with prior student records collections as much as possible. Consequently, the NPSAS:18-AC student records data elements were generally consistent with those collected during the 2012/17 Beginning Postsecondary Students Longitudinal Study (BPS:12/17) student records collection, except that BPS:12/17 covered 6 academic years (2011–12 through 2016–17), while NPSAS:18-AC covered only the 2017–18 academic year.

A few substantive changes were made to individual items to adapt the student records instrument for NPSAS:18-AC. These changes aligned student records data elements across NCES postsecondary studies, clarified item definitions, and made the PDP more user friendly for institutions. For example, based on the results of NPSAS:16, eight items in the Eligibility subsection of the NPSAS:16 student records instrument were collapsed into one item, included in the General Student Information section of the NPSAS:18-AC student records instrument, to collect a single reason that the student was ineligible for NPSAS:18-AC. In addition, Graduate Aid was removed as a separate subsection in the Financial Aid

section. Instead, institutions reported graduate aid based on the source, which was consistent with how other aid types were reported. Appendix D includes a complete list of the items in each section of the student records instrument.

Institutions provided student records data through the PDP described in section 3.1.2. Institutions were offered three modes to submit student records data, according to their own preferences and internal capabilities. Institutions could choose to use any of these methods and could change modes at any time:

1. Excel mode, in which institution staff downloaded a preformatted Excel spreadsheet template from the PDP, completed the template by hand-keying or pasting data, and then uploaded the completed template to the PDP;

2. comma-separated values (CSV) mode, in which institution staff downloaded CSV file specifications from the PDP, prepared .csv-formatted data files offline, and then uploaded the completed files to the PDP; and

3. web mode, in which institution staff entered student data into the instrument one field at a time, using text-entry fields and drop-down boxes.

3.2 Institution Contacting, Recruitment, and Student Enrollment List Acquisition

At the outset of institution data collection, sampled institutions were contacted to secure their participation in the study. Institutions were asked to designate an institution coordinator to act as a primary point of contact for the submission of student enrollment lists. Institution contacting, recruitment, and student enrollment list acquisition activities are described in the following sections.

3.2.1 Institution Contacting and Recruitment

Sampled institutions began to be notified of the data collection in April 2018, and lists were requested a few weeks later from the majority of institutions. The schedule differed from past iterations of NPSAS in which institutions were contacted roughly 4 months before the earliest enrollment list submission deadlines. Sixteen people were trained as institution contactors to contact institutions, including several who were already working on the transcript collection for BPS. Appendix E contains the training agenda for institution

contactors. Their training included an overview of NPSAS:18-AC, guidance on building strong working relationships with institution coordinators, and instruction on assisting with data collection and submission using the PDP. Institution contactors identified some 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 or not that institution was sampled. The identified institutions were reviewed to confirm whether the institution was ineligible for the study.

To encourage participation and confirm the legitimacy of the study before data collection, a list of postsecondary organizations and associations that have endorsed NPSAS:18-AC were printed on all letters to the institutions and posted on the PDP (appendix F). Many of the same organizations endorsed NPSAS in 2016. Institutions were also provided with outreach materials promoting NPSAS and DataLab. Institution outreach and marketing materials can be found in appendix G.

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

- a cover letter printed on NCES letterhead providing background information on NPSAS:18-AC and requesting that the chief administrator designate an institution coordinator via the PDP;
- website access instructions;
- a brochure that summarized NPSAS:18-AC objectives and provided background information and key findings from past NPSAS cycles; and
- a schedule and flow chart of all NPSAS:18-AC data collection activities.

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

Once the institution named an institution coordinator, institution contactors confirmed study participation and set a deadline for submission of the student enrollment list. Deadlines were customized according to the institution's term

structure. For institutions with distinct terms, institution coordinators were asked to provide the start and end dates for the term that included April 30, 2018. Institution contactors set the institution's deadline for 2 weeks after the start of that term, or if that date had passed, for 4 weeks from the time the institution coordinator completed the registration page or received the list request letter. For institutions with continuous enrollment, institution contactors asked institution coordinators to provide lists by May 15. Institution contactors communicated as needed with institution coordinators to offer reminders of the scheduled due date and to find out if they had any questions.

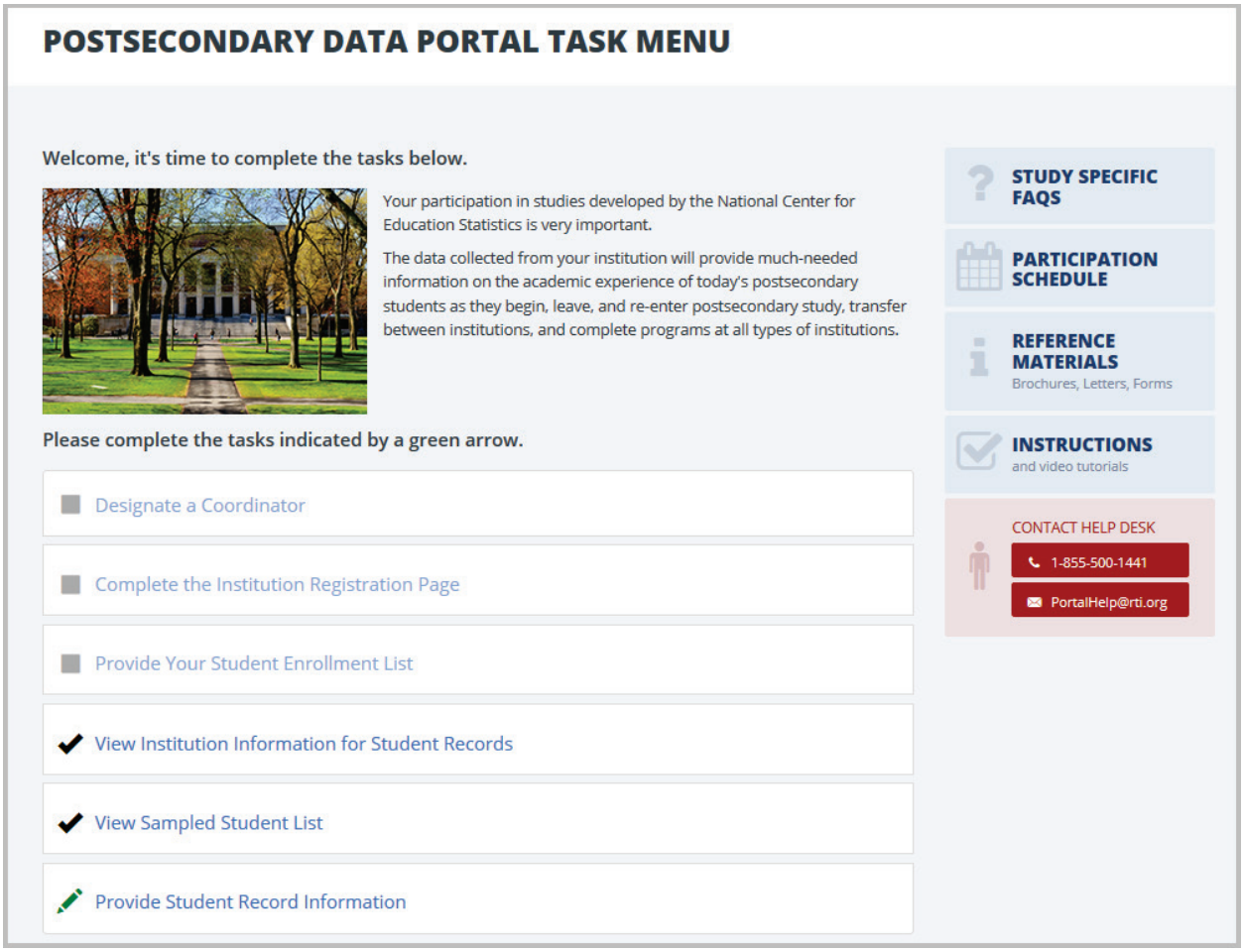
After institution coordinators were designated, institution contactors sent them the following materials (appendix H):

- a cover letter describing the study with information on how to access the PDP and complete the Institution Registration Page (IRP);
- a brochure describing the study; and
- a schedule and flow chart of all NPSAS:18-AC data collection activities.

Institution contactors then followed up by telephone to confirm receipt and prompt for completion of the IRP. After institution coordinators 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 institution coordinators were prompted by telephone before 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. Once logged in, an institution coordinator 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 (figure 2). This design allowed institutions to identify the tasks that were not yet completed and monitor their overall progress.

Figure 2. Postsecondary Data Portal task menu: 2017–18



SOURCE: U.S. Department of Education, National Center for Education Statistics, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

Large campus systems with centralized record keeping were identified at the start of data collection using IPEDS reporting data. Institution coordinators 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 the large campus systems. Institution contactors worked directly with these systems to provide guidance on reporting and to accommodate any ad hoc requests.

3.2.2 Student Enrollment List Acquisition

Institutions were formally asked to provide enrollment list information for all students enrolled at any time between July 1, 2017, and April 30, 2018.[10] The

[10] The NPSAS:18-AC target population consisted of all eligible students enrolled at any time between July 1, 2017, and June 30, 2018. However, most institutions provided enrollment lists that covered the period of July 1, 2017, through April 30, 2018. The date of April 30 was selected to include virtually all students enrolled before the summer term without delaying data collection.

PDP provided institution coordinators 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. 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 institution coordinators submitted enrollment lists, the lists' quality and completeness were checked several times before the student sample was selected. These checks included verifying that institutions used a readable format, that key data needed for sampling (e.g., degree program) were provided, and that the counts of undergraduate and graduate students provided on the list were within a plus-or-minus 50-percent range of the IPEDS counts. 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 3,080 eligible institutions, 87 percent initially agreed to participate by designating an institution coordinator. Of the eligible institutions, 72 percent provided usable student enrollment lists. Seven lists were rejected and omitted from these counts because of issues that the institutions would or could not resolve, including too many missing items for the list to be usable or obvious errors (e.g., mismatched e-mail addresses). In total, 2,210 institutions provided eligible enrollment lists in NPSAS:18-AC. Seventy-nine percent of the 1,580 NPSAS:18-AC sample institutions that had previously participated in NPSAS:16 or NPSAS:12 data collection provided enrollment lists for NPSAS:18-AC, a rate statistically different from the rate among institutions that had not previously participated (65 percent), $\chi^2(1, n = 3{,}080) = 74.6, p < .05$.

The percentage of institutions providing enrollment lists across control and level of institution ranged from 38 percent to 87 percent (table 9). The lowest participation rates were among public less-than-2-year institutions (38 percent) and the private for-profit 2-year sector (44 percent).

Table 9. Number and percentages of institutions providing enrollment lists and number of sampled students, by control and level of institution: 2017–18

Control and level of institution	Eligible institutions	Institutions providing enrollment lists			Sampled students
		Number	Unweighted percent	Weighted percent[1]	
All institutions	**3,080**	**2,210**	**71.7**	**79.3**	**349,650**
Control of institution					
Public	1,760	1,350	76.8	82.5	253,230
Private nonprofit	760	580	75.3	77.0	70,010
Private for-profit	560	290	51.1	58.6	26,410
Level of institution					
Less-than-2-year	220	110	50.2	54.7	7,530
2-year	1,210	830	68.1	75.5	115,670
4-year, non-doctorate-granting	900	660	73.3	75.8	84,620
4-year, doctorate-granting	750	620	82.0	85.0	141,840
Control and level of institution					
Public less-than-2-year	30	10	38.2	44.9	1,060
Public 2-year	980	720	73.5	77.5	108,660
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	140	110	83.0	84.0	17,880
Public 4-year, non-doctorate-granting, primarily baccalaureate	230	180	76.1	79.0	26,480
Public 4-year, doctorate-granting	370	330	87.4	88.9	99,160
Private nonprofit less-than-2-year	40	20	51.4	82.3	1,120
Private nonprofit 4-year, non-doctorate-granting	390	290	76.2	75.3	32,620
Private nonprofit 4-year, doctorate-granting	340	260	76.8	77.3	36,280
Private for-profit less-than-2-year	180	100	52.7	52.5	6,380
Private for-profit 2-year	200	90	44.2	42.4	5,970
Private for-profit 4-year	180	100	56.9	65.2	14,050

[1] The weighted percentages were calculated using the institution sampling weight.
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, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

As mentioned previously, institution-provided enrollment lists were evaluated for accuracy and completeness in several ways, including comparing institution-provided data to IPEDS data for the same institutions. Institutions that had submitted student counts with discrepancies were then contacted to reconcile the data. During quality checks, 84 percent of the enrollment lists had no problems identified.

Institution participation was examined by selected classification categories used for implicit stratification (see section 2.2), including 2010 Carnegie Classification categories (table 10). Of the 2,210 institutions that provided enrollment lists in NPSAS:18-AC, 570 did not have a Carnegie Classification. Of those with a known Carnegie Classification, the number that provided enrollment lists ranged from a high of 210 institutions classified as Associate's-public rural-serving medium to six classification categories with numbers that rounded to 0.

Table 10. Number and percentage of NPSAS:18-AC institutions providing enrollment lists, by 2010 Carnegie Classification: 2017–18

2010 Carnegie Classification	Number	Percent
All institutions	**2,210**	**100.0**
Not classified	570	25.8
Associate's—public rural-serving small	80	3.7
Associate's—public rural-serving medium	210	9.4
Associate's—public rural-serving large	100	4.6
Associate's—public suburban single campus	90	3.8
Associate's—public suburban multicampus	70	3.1
Associate's—public urban single campus	30	1.1
Associate's—public urban multicampus	90	3.9
Associate's—public special use	#	0.1
Associate's—private nonprofit	10	0.4
Associate's—private for-profit	30	1.4
Associate's—public 2-year under 4-year	40	1.9
Associate's—public 4-year primarily associate's	30	1.4
Associate's—private nonprofit 4-year primarily associate's	#	0.2
Associate's—private for-profit 4-year primarily associate's	10	0.3
Research (very high research activity)	70	3.3
Research (high research activity)	70	3.2
Doctor's/research universities	40	1.6
Master's (larger programs)	190	8.7
Master's (medium programs)	70	3.3
Master's (smaller programs)	50	2.2
Bachelor's colleges—arts and sciences	80	3.4
Bachelor's colleges—diverse fields	120	5.6
Bachelor's/associate's colleges	40	1.8
Special focus—theological	20	0.9
Special focus—medical	20	1.1
Special focus—other health professions	20	0.9
Special focus—engineering	#	0.1
Special focus—other technology	#	#
Special focus—business/management	10	0.4
Special focus—art, music, and design	20	0.9
Special focus—law	10	0.2
Special focus—other special-focus	#	0.1
Tribal colleges	#	1.0

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, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

Table 11 shows the number and percentage of HBCUs that provided enrollment lists, also used for implicit stratification, by current and prior NPSAS studies. Sixty HBCUs provided enrollment lists in NPSAS:18-AC.

Table 11. Number and percentage of Historically Black Colleges and Universities (HBCUs) providing enrollment lists, by current and prior National Postsecondary Student Aid Study (NPSAS) studies: 2017–18

Current and prior NPSAS studies	Number of HBCUs providing enrollment lists	HBCUs as a percentage of total number of institutions that provided enrollment lists
1986–87 National Postsecondary Student Aid Study	20	1.9
1989–90 National Postsecondary Student Aid Study	20	1.5
1982–93 National Postsecondary Student Aid Study	30	2.6
1995–96 National Postsecondary Student Aid Study	20	1.9
1999–2000 National Postsecondary Student Aid Study	20	2.3
2003–04 National Postsecondary Student Aid Study	30	2.1
2007–08 National Postsecondary Student Aid Study	40	2.3
2011–12 National Postsecondary Student Aid Study	30	2.0
2015–16 National Postsecondary Student Aid Study	40	2.0
2017–18 National Postsecondary Student Aid Study, Administrative Collection	60	2.4

NOTE: Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

3.3 Student Records Data Collection

Once institutions sent student enrollment lists, the student sample was created as detailed in chapter 2, and institution record data collection for sample members began. These data were used to identify both students and institutions as student records respondents. The final count of student respondents included student records respondents who had at least 1 month of enrollment confirmed in any data source, including administrative data sources detailed in chapter 4. The following sections describe student records collection and outcomes.

3.3.1 Student Records Collection From Institutions

After a student sample was selected for an institution, the designated institution coordinator was sent an information packet on the student records collection process. The packet included instructions for accessing the PDP and a Quick Guide to Providing Student Records Data (appendix H). The PDP 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 PDP—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 available to help if institution staff had questions or encountered problems.

3.3.2 Student Records Collection Outcomes

Of the 2,210 institutions with sampled students, 1,910 (86 percent) provided student records data. Most institutions that provided student records data opted for Excel mode (65 percent), 18 percent uploaded a CSV file, and the remaining 17 percent used web mode. Table 12 shows student records data collection results by collection mode and control and level of institution. From the institutions that provided student records data, these data were obtained for 78 percent of eligible sample members (table 13). This total included 77 percent of the total undergraduate students in the sample and 89 percent of the graduate students.

Table 12. Student records data collection results, by data collection mode and control and level of institution: 2017–18

Control and level of institution	Institutions			Data collection mode					
	Provided enrollment lists	Provided student records		Excel		CSV		Web	
		Number	Percent	Number	Percent	Number	Percent	Number	Percent
Total	**2,210**	**1,910**	**86.4**	**1,230**	**64.6**	**350**	**18.3**	**330**	**17.1**
Control of institution									
Public	1,350	1,170	86.8	720	61.7	270	22.7	180	15.6
Private nonprofit	580	510	88.0	410	80.0	80	15.0	30	4.9
Private for-profit	290	230	81.2	110	45.5	10	3.4	120	51.1
Level of institution									
Less-than-2-year	110	90	81.1	60	63.3	#	#	30	36.7
2-year	830	690	83.0	480	70.4	110	15.6	100	14.0
4-year, non-doctorate-granting	660	580	87.4	360	61.7	90	14.9	140	23.4
4-year, doctorate-granting	620	560	90.7	340	60.6	160	28.1	60	11.3

Rounds to zero.
NOTE: CSV = comma-separated values. 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, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

Table 13. Student records data collection results, by control and level of institution and student type: 2017–18

Control and level of institution and student type	Institutions			Students		
	Provided enrollment lists	Provided student records		Total sample eligible students[1]	Student records collected	
		Number	Percent		Number	Percent
Total	**2,210**	**1,910**	**86.4**	**344,710**	**268,960**	**78.0**
Control of institution						
Public	1,350	1,170	86.8	249,360	195,930	78.6
Private nonprofit	580	510	88.0	69,330	54,660	78.8
Private for-profit	290	230	81.2	26,010	18,370	70.6
Level of institution						
Less-than-2-year	110	90	81.1	7,430	5,040	67.8
2-year	830	690	83.0	113,590	85,900	75.6
4-year, non-doctorate-granting	660	580	87.4	83,300	62,430	74.9
4-year, doctorate-granting	620	560	90.7	140,390	115,600	82.3
Control and level of institution						
Public less-than-2-year	10	10	76.9	1,020	760	73.9
Public 2-year	720	610	84.0	106,600	81,830	76.8
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	110	90	80.4	17,540	11,550	65.8
Public 4-year, non-doctorate-granting, primarily baccalaureate	180	160	90.4	25,900	19,180	74.0
Public 4-year, doctorate-granting	330	310	93.6	98,300	82,620	84.0
Private nonprofit less-than-4-year	20	20	78.9	1,120	910	81.5
Private nonprofit 4-year, non-doctorate-granting	290	260	87.4	32,380	25,550	78.9
Private nonprofit 4-year, doctorate-granting	260	230	89.3	35,840	28,200	78.7
Private for-profit less-than-2-year	100	80	81.3	6,320	4,210	66.5
Private for-profit 2-year	90	70	76.1	5,960	3,230	54.2
Private for-profit 4-year	100	90	85.4	13,730	10,940	79.7
Student type						
Undergraduate	†	†	†	320,160	247,090	77.2
Graduate	†	†	†	24,550	21,870	89.1

† Not applicable.
[1] Total sample eligible students sampled from 2,210 institution enrollment lists.
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, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

Overall institution participation. Table 14 shows the number and percentage of institutions that provided an enrollment list and were student records respondents, by state and institution stratum. Institutions that provided an enrollment list and were student records respondents were considered participating institutions. Appendix I contains maps of enrollment list participation rates and student records participation rates for each state overall and by institution stratum.

Table 14. Number and percentage of institutions that provided an enrollment list and were student records respondents, by institution stratum and state: 2017–18

State	All sectors			Institution stratum[1]								
				Public 2-year[2]			Public 4-year			Other sectors		
	Number eligible	Number participated	Percent participated	Number eligible	Number participated	Percent participated	Number eligible	Number participated	Percent participated	Number eligible	Number participated	Percent participated
All states	**3,080**	**1,910**	**62.0**	**980**	**610**	**61.7**	**740**	**560**	**74.9**	**1,360**	**750**	**55.1**
Alabama	70	40	57.6	20	20	78.3	10	10	71.4	30	10	34.5
Alaska	10	10	77.8	†	†	†	#	#	#	10	#	60.0
Arizona	60	40	60.3	20	20	85.0	10	10	70.0	30	10	39.3
Arkansas	60	40	68.3	20	20	81.8	10	10	63.6	30	20	60.0
California	180	90	48.1	110	50	42.1	50	30	56.5	30	20	56.7
Colorado	60	40	70.0	10	10	92.9	20	10	87.5	30	20	50.0
Connecticut	50	30	55.8	10	10	61.5	10	#	44.4	30	20	56.7
Delaware	20	10	62.5	†	†	†	#	#	100.0	10	10	53.8
District of Columbia	20	10	54.5	†	†	†	#	#	100.0	20	10	50.0
Florida	100	60	60.8	40	20	51.4	40	30	76.9	30	10	50.0
Georgia	80	50	62.5	20	10	30.4	30	30	96.3	30	20	56.7
Hawaii	20	20	66.7	10	10	100.0	#	#	100.0	10	10	42.9
Idaho	40	20	50.0	#	#	50.0	#	#	75.0	30	10	46.4
Illinois	90	60	71.9	50	40	72.9	10	10	91.7	30	20	62.1
Indiana	40	30	72.7	#	#	100.0	10	10	100.0	30	20	58.6
Iowa	50	40	73.5	20	10	81.3	#	#	100.0	30	20	66.7
Kansas	60	40	60.7	30	10	52.0	10	10	87.5	30	20	60.7
Kentucky	50	40	69.8	20	10	68.8	10	10	87.5	30	20	65.5
Louisiana	60	40	65.6	20	10	66.7	20	10	76.5	30	20	58.6
Maine	40	20	54.1	10	#	57.1	10	#	37.5	20	10	59.1
Maryland	60	30	52.5	20	10	37.5	10	10	84.6	30	10	46.7
Massachusetts	60	50	80.3	20	10	68.8	20	10	93.3	30	20	80.0
Michigan	80	50	68.4	30	20	80.0	20	20	85.7	30	10	46.7
Minnesota	70	60	85.9	30	30	93.5	10	10	100.0	30	20	71.4
Mississippi	60	30	44.6	20	10	40.0	10	10	87.5	30	10	36.4
Missouri	60	40	71.2	20	10	58.8	10	10	84.6	30	20	72.4
Montana	30	20	48.4	10	10	45.5	10	10	83.3	10	10	35.7
Nebraska	50	30	65.3	10	10	75.0	10	#	42.9	30	20	67.6
Nevada	40	20	46.2	#	#	100.0	10	#	50.0	30	10	43.8
New Hampshire	40	20	46.2	10	#	14.3	10	10	83.3	30	10	46.2

See notes at end of table.

Table 14. Number and percentage of institutions that provided an enrollment list and were student records respondents, by institution stratum and state: 2017–18—Continued

State	All sectors			Institution stratum[1]								
				Public 2-year[2]			Public 4-year			Other sectors		
	Number eligible	Number participated	Percent participated	Number eligible	Number participated	Percent participated	Number eligible	Number participated	Percent participated	Number eligible	Number participated	Percent participated
New Jersey	60	40	59.7	20	10	63.2	10	10	69.2	30	20	53.3
New Mexico	50	30	56.5	20	10	73.7	10	#	44.4	20	10	44.4
New York	110	80	75.7	40	30	73.7	40	40	86.0	30	20	63.3
North Carolina	110	70	75.7	60	40	66.1	20	10	87.5	30	20	66.7
North Dakota	30	10	37.9	10	#	40.0	10	#	44.4	20	10	33.3
Ohio	100	60	62.0	40	20	68.6	40	20	57.1	30	20	60.0
Oklahoma	70	30	48.6	20	10	58.3	20	10	70.6	30	10	27.6
Oregon	60	30	52.7	20	10	52.9	10	10	55.6	30	20	51.7
Pennsylvania	90	60	69.6	20	10	64.7	50	30	75.6	30	20	63.3
Puerto Rico	50	30	58.3	10	#	20.0	10	10	35.7	30	20	75.9
Rhode Island	20	10	50.0	#	#	#	#	#	100.0	20	10	47.4
South Carolina	60	40	60.3	20	10	35.0	10	10	84.6	30	20	66.7
South Dakota	30	20	57.1	10	#	60.0	10	10	71.4	20	10	50.0
Tennessee	80	40	55.8	40	20	55.3	10	10	50.0	30	20	58.6
Texas	140	100	70.1	60	40	63.3	50	40	89.4	30	20	53.3
Utah	40	30	61.0	#	#	100.0	10	10	100.0	30	10	46.7
Vermont	30	10	48.0	#	#	100.0	10	#	60.0	20	10	42.1
Virginia	70	50	64.3	20	10	58.3	20	10	75.0	30	20	63.3
Washington	70	50	65.3	10	10	66.7	30	20	51.6	30	20	79.3
West Virginia	50	30	52.8	10	10	45.5	10	10	84.6	30	10	41.4
Wisconsin	60	40	60.7	20	10	75.0	20	10	73.3	30	10	46.7
Wyoming	10	10	60.7	10	#	42.9	#	#	100.0	#	#	50.0

† Not applicable.
# Rounds to zero.
[1] Institution stratum reflects institution categorization as determined from the 2016–17 Integrated Postsecondary Education Data System files.
[2] There were no public 2-year schools in Alaska, Delaware, or the District of Columbia.
NOTE: Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding. For the purposes of this table, Puerto Rico and the District of Columbia are referred to as "states."
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

3.4 Institution Data Evaluation

As each institution submitted student records, the data were reviewed for quality and completeness. First, automated programs assessed the missingness of critical data elements, invalid or out-of-range data, and the inclusion of sample members' personally identifiable information (PII). After administrative records matching (described in chapter 4), student records data from sample members were checked to confirm that the students were enrolled for at least 1 month between July 1, 2017, and June 30, 2018. If no enrollment could be confirmed in institution data or any administrative data source, the student was not considered a student respondent. Out of the 268,960 student record respondents, 2,050 students did not have at least 1 month of enrollment confirmed and were not considered student respondents. The final count of student respondents was 266,910. Table 15 shows the completion rates of key student records data elements by data collection mode—Excel, CSV, or web. The data elements requested are available at a large majority of institutions, which is reflected in the high completion rates of key student records data elements. Variation in the item-level responses 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.

Items with the highest completion rates were class level (100 percent) and degree program (100 percent), which are critical data elements for institutions to determine students' eligibility for financial aid. Of the key data elements listed in table 15, the items with the lowest completion rates were citizenship status (79 percent) and race (82 percent). Some institutions reported not being able to provide race because their data management systems record the data in a different format than was requested for NPSAS:18-AC.

Table 15. Student records item-level completion rates, by data collection mode and data element: 2017–18

	Total student respondents[1]		Data collection mode					
			Excel		CSV		Web	
Data element	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Total	**266,910**	**100.0**	**132,860**	**49.8**	**99,500**	**37.3**	**34,550**	**12.9**
Student characteristics								
Sex	266,320	99.8	132,450	99.7	99,440	99.9	34,430	99.6
Marital status	239,890	89.9	115,230	86.7	91,740	92.2	32,910	95.3
Citizenship	209,630	78.5	97,090	73.1	84,480	84.9	28,060	81.2
High school completion type	234,850	88.0	109,070	82.1	92,990	93.5	32,790	94.9
Race	218,370	81.8	105,850	79.7	84,260	84.7	28,270	81.8
Ethnicity	265,830	99.6	132,060	99.4	99,400	99.9	34,370	99.5
Enrollment								
Degree program	266,700	99.9	132,680	99.9	99,470	100.0	34,550	100.0
Student class level	266,840	100.0	132,840	100.0	99,450	99.9	34,550	100.0
Residency for tuition purposes	244,440	91.6	118,770	89.4	93,680	94.1	32,000	92.6
Total tuition and fees charged	240,910	90.3	117,520	88.5	91,980	92.4	31,410	90.9
Budget								
Tuition and fees	244,720	91.7	123,170	92.7	92,950	93.4	28,600	82.8
Financial aid								
Any aid received[2]	264,680	99.2	131,250	98.8	98,890	99.4	34,540	100.0

[1] Total student respondent count comprises the 268,960 student records respondents, minus 2,050 cases that did not have at least 1 month of enrollment verified in any administrative data source.
[2] The item is counted as completed if the institution indicated whether or not the student had received any aid. This row is not a measure of how many respondents received aid.
NOTE: CSV = comma-separated values. All percentages are unweighted and based on the number of eligible students within the row. All nonmissing responses, including responses of "Unknown," are counted as complete. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

The automated data review programs produced reports summarizing the data quality assessments and a pass or fail result. When data problems were detected, the automated reports were reviewed, the source of data problems was identified, and the problems were resolved whenever possible. If the data problems could not be resolved, the issue was documented for institution contactors to follow up with the institution. Institutions would indicate that the data should be used as submitted, or they would be asked to provide updated data. The most common data problems that institutions were contacted for were missing critical data elements or missing entire instrument sections.

Chapter 4. Administrative Records Matching Overview and Outcomes

In addition to the student records collection, student data for NPSAS:18-AC came from administrative databases, including two from the U.S. Department of Education Federal Student Aid (FSA) Office: CPS and NSLDS. Additional data sources included NSC, ACT, College Board, and VBA. These additional data sources were useful in providing information that could not be collected from institutions and when assessing the accuracy of similar information from other sources. This chapter provides detail on administrative data matching processes and outcomes.

4.1 Administrative Records Matching

4.1.1 Central Processing System (CPS)

To reduce institution burden, federal financial aid application data were obtained from 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.

Record matching was conducted for NPSAS:18-AC against CPS data for the 2017–18 financial aid year using the sample member's CPS ID—the sample member's SSN concatenated with the first two letters of the sample member's last name. Records were not submitted to CPS for matching for sample members without available SSNs. Data were transmitted to FSA using its Secure Sockets Layer (SSL) encrypted website, and from FSA to project staff using EdConnect, a software program provided by the U.S. Department of Education to securely transmit data.

Two academic years of FAFSA data were accessible at any time: the current academic year and the following academic year. In September of each year, access to the academic year most recently completed is lost and access for the following academic year is added. Thus, after September 2018, the 2017–18 FAFSA data were no longer accessible through FSA's CPS. Because NPSAS:18-AC sample selection was not completed by that time, only about 38 percent of the sample members were matched to CPS for 2017–18. For the sample members selected after September 2018, FAFSA data were obtained from tables in NSLDS instead. Data from both sources were then merged.

4.1.2 National Student Loan Data System (NSLDS)

Student-level data on Pell Grants and federal student loans were obtained by matching sample members to the NSLDS database. In cooperation with FSA, a records match was initiated between NPSAS:18-AC records and the NSLDS database after the entire sample was selected in order to retrieve the most current NSLDS data. As with CPS, records for sample members missing SSNs were not part of the match. NPSAS student respondents 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.

4.1.3 National Student Clearinghouse (NSC)

Enrollment data were obtained for the student sample from the NSC StudentTracker service. This administrative record matching provided information on institutions attended, enrollment dates, and degree completions. An individual student record would match with NSC only if the student's institution was a participant in NSC.[11] StudentTracker data were requested after 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. PII used for the match included sample member name, SSN, and DOB.

[11] For more information on NSC participation, visit https://www.studentclearinghouse.org.

4.1.4 ACT

NPSAS sample member data files were merged with ACT data files to obtain admissions test data. The ACT files used contained survey data and a record of the highest test score registered by each student between the 2011–12 and 2016–17 academic years. This record matching was performed after data collection to use the most updated personally identifying data (first and last name, middle initial, DOB, and last four digits of SSN) as matching criteria. If the file merge yielded multiple test records per student, it returned only the most recent record. All data transfers used a password-protected NCES system transmitting over an encrypted SSL connection.

4.1.5 SAT

To obtain SAT test scores and questionnaire data, sample member data files were merged with College Board records spanning high school graduation years from 2012 to 2017. If the file merge yielded multiple test records per student, it returned only the most recent record. As with ACT, files were merged after the end of data collection. Matching was done using name, DOB, SSN, sex, graduation year, and permanent zip code. The file transfers were secured through an NCES system that required a log-in, a password, and an encrypted SSL connection.

4.1.6 Veterans Benefits Administration (VBA)

A file match was performed with VBA to identify veterans, amounts of federal veterans education benefits, and any associated enrollment information. After the entire sample was selected, a file containing SSN, name, and DOB was provided to 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 ACT and SAT file matching, all data transmission used an NCES secure file transfer system.

4.1.7 Administrative Records Quality Checks

Several checks were performed to verify the quality of the administrative data received. File layouts and code to read in the files were compared to ensure that the code was current and accurate. File completeness was also checked; for example, with NSLDS, the file received had to have a header and a footer, otherwise the file and data were incomplete. If a source provided PII from its database, names and dates of births were compared to what was in the study database to make sure the data were for the correct person. If the names and dates

of birth did not match, the case was excluded from the administrative file. Project staff examined individual data files by running basic summary statistics such as number of records and value ranges (e.g., dates and amounts) to check for potential outliers or abnormalities. Then the files were checked to confirm they were related to one another as expected. For example, student veterans who received veterans education benefit payments were expected to have at least one military service record.

4.2 Administrative Records Matching Outcomes

4.2.1 Central Processing System (CPS)

Table 16 summarizes the results of matching sample members to CPS overall and by institution and student characteristics. Data were sent for matching to CPS 2017–18 for the sample members for whom there was a valid SSN, which was about 91 percent of all sample members. The overall matching rate for the 2017–18 academic year was 62 percent. Match rates varied by control and level of institution, ranging from a low of 42 percent for public less-than-2-year institutions to a high of 78 percent for private for-profit less-than-2-year institutions. These match rates met expectations, which were based on trend estimates of Pell Grant and federal student loan funding recipients over time.[12]

[12] Trend estimates were generated using NCES TrendGenerator (https://nces.ed.gov/ipeds/TrendGenerator).

Table 16. Results of Central Processing System (CPS) matching for 2017–18, by control and level of institution and student type: 2017–18

Control and level of institution and student type	Sample	Sent to CPS 2017–18[1]		Matched to CPS 2017–18	
		Number	Percent	Number	Percent[2]
Total	**349,650**	**318,770**	**91.2**	**216,620**	**62.0**
Control of institution					
Public	253,230	232,180	91.7	153,100	60.5
Private nonprofit	70,010	61,800	88.3	44,070	62.9
Private for-profit	26,410	24,780	93.8	19,450	73.7
Level of Institution					
Less-than-2-year	7,530	6,900	91.6	5,500	73.1
2-year	115,670	107,530	93.0	70,420	60.9
4-year, non-doctorate-granting	84,620	77,950	92.1	55,780	65.9
4-year, doctorate-granting	141,840	126,390	89.1	84,910	59.9
Control and level of institution					
Public less-than-2-year	1,060	900	84.7	450	42.4
Public 2-year	108,660	101,050	93.0	65,180	60.0
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	17,880	16,760	93.8	10,830	60.6
Public 4-year, non-doctorate-granting, primarily baccalaureate	26,480	24,460	92.4	17,090	64.5
Public 4-year, doctorate-granting	99,160	89,010	89.8	59,550	60.1
Private nonprofit less-than-4-year	1,120	1,060	95.1	760	68.1
Private nonprofit 4-year, non-doctorate-granting	32,620	29,630	90.8	22,410	68.7
Private nonprofit 4-year, doctorate-granting	36,280	31,110	85.8	20,900	57.6
Private for-profit less-than-2-year	6,380	5,920	92.7	5,000	78.3
Private for-profit 2-year	5,970	5,490	92.0	4,530	75.8
Private for-profit 4-year	14,050	13,370	95.1	9,930	70.6
Student type					
Undergraduate	324,790	296,520	91.3	205,360	63.2
Graduate	24,870	22,250	89.5	11,260	45.3

[1] "Sent to CPS" includes cases sent for matching to CPS 2017–18 as well as cases selected after September 2018 whose Free Application for Federal Student Aid data were taken from tables in the National Student Loan Data System after CPS 2017–18 was no longer available for matching.
[2] Percentage of cases sent to CPS.
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, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

Of all undergraduate students, 63 percent matched to CPS for the 2017–18 academic year, while only 45 percent of graduate students matched. 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 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 completion of the federal financial aid forms, are not available as part of CPS.

4.2.2 National Student Loan Data System (NSLDS)

Sample members with valid SSNs—about 91 percent of all sample members—were matched with NSLDS. As stated previously, NSLDS matching only returned records of sample members who, at some point, had received Pell Grant or federal student loan funding. The NSLDS database is historical and includes information not only for the 2017–18 academic year but also for prior years. Table 17 shows the overall NSLDS match rates for sample members. A match indicates that a student had at least one loan or Pell Grant, although not necessarily during 2017–18. Based on existing estimates from FSA[13] and NCES,[14] approximately 44 percent of college students received federal grants or loans during 2017–18.

[13] According to FSA's Aid Recipient Summary estimates (https://studentaid.gov/data-center/student/title-iv), there were approximately 11,464,130 aid recipients in the 2017–18 academic year.
[14] According to the NCES TrendGenerator (https://nces.ed.gov/ipeds/TrendGenerator), there were 26,321,518 students enrolled in postsecondary institutions in the 2017–18 academic year.

Table 17. Results of National Student Loan Data System (NSLDS) loan and Pell Grant matching, by control and level of institution and student type: 2017–18

Control and level of institution and student type	Sample	Sent to NSLDS		Matched to NSLDS loan		Matched to NSLDS Pell Grant	
		Number	Percent	Number	Percent[1]	Number	Percent[1]
Total	**349,650**	**318,760**	**91.2**	**193,340**	**55.3**	**174,880**	**50.0**
Control of institution							
Public	253,230	232,180	91.7	130,770	51.6	126,290	49.9
Private nonprofit	70,010	61,800	88.3	42,630	60.9	29,720	42.5
Private for-profit	26,410	24,780	93.8	19,940	75.5	18,870	71.5
Level of institution							
Less-than-2-year	7,530	6,900	91.6	5,100	67.8	5,460	72.6
2-year	115,670	107,530	93.0	54,260	46.9	66,260	57.3
4-year, non-doctorate-granting	84,620	77,950	92.1	51,040	60.3	45,230	53.4
4-year, doctorate-granting	141,840	126,390	89.1	82,950	58.5	57,930	40.8
Control and level of institution							
Public less-than-2-year	1,060	900	84.7	350	32.8	540	50.3
Public 2-year	108,660	101,040	93.0	49,190	45.3	61,190	56.3
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	17,880	16,760	93.8	7,760	43.4	10,160	56.8
Public 4-year, non-doctorate-granting, primarily baccalaureate	26,480	24,460	92.4	16,610	62.7	13,650	51.6
Public 4-year, doctorate-granting	99,160	89,010	89.8	56,870	57.3	40,750	41.1
Private nonprofit less-than-4-year	1,120	1,060	95.1	660	58.6	680	60.7
Private nonprofit 4-year, non-doctorate-granting	32,620	29,630	90.8	21,120	64.8	15,900	48.8
Private nonprofit 4-year, doctorate-granting	36,280	31,110	85.8	20,850	57.5	13,140	36.2
Private for-profit less-than-2-year	6,380	5,920	92.7	4,730	74.1	4,870	76.2
Private for-profit 2-year	5,970	5,490	92.0	4,440	74.2	4,450	74.5
Private for-profit 4-year	14,050	13,370	95.1	10,780	76.7	9,560	68.0
Student type							
Undergraduate	324,790	296,510	91.3	177,080	54.5	165,520	51.0
Graduate	24,870	22,250	89.5	16,270	65.4	9,360	37.6

[1] Percentage of cases sent to NSLDS.
NOTE: Both institution and student classifications were verified to correct classification errors on the sampling frame. Matching was completed on historical files that included awards made in 2017–18 and prior years. 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, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

In total, approximately 193,340 sample members were matched to loan data, accounting for 55 percent of total sample members with a recorded SSN. NSLDS loan match rates by control of institution ranged from 52 percent of public institutions to 76 percent of private for-profit institutions. The match rates by level of institution ranged from 47 percent for 2-year institutions to 68 percent for less-than-2-year institutions. Match rates for control and level of institution ranged from a low of 33 percent for public less-than-2-year institutions to a high of 77 percent for private for-profit 4-year institutions. Of undergraduate students, 55 percent matched to the loan database, while 65 percent of graduate students matched.

NSLDS match yielded Pell Grant matches for 174,880 sample members (50 percent). Match rates by control and level of institution ranged from 36 percent for private nonprofit 4-year, doctorate-granting institutions to 76 percent for private for-profit less-than-2-year institutions. Of undergraduate students, 51 percent matched to the Pell Grant database, while 38 percent of graduate students had a match.

4.2.3 National Student Clearinghouse (NSC)

Approximately 97 percent of students nationwide attend institutions that report enrollment and degree data to NSC. NSC matches used enrollment and degree records for the 2017–18 academic year. Individual student record matching was possible only if an institution the student attended was a participant in NSC. NSC matches for sample members included their NPSAS-sampled institution and any other participating institutions they attended during the 2017–18 academic year.

Of the total sample members, 289,220 (83 percent) matched to NSC for their NPSAS-sampled institution. By control and level of institution, the match rate ranged from 1 percent for private for-profit less-than-2-year institutions to 91 percent for public 4-year, doctorate-granting institutions. Match rates by level of institution ranged from 6 percent for less-than-2-year institutions to 89 percent for 4-year, doctorate-granting institutions. Match rates by control of institution ranged from 30 percent for private for-profit institutions to 88 percent for public institutions. Matches to institutions other than sample members' NPSAS institutions yielded results for 57,250 sample members (16 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 18 shows NSC match rates by control and level of institution and student type.

Table 18. Results of National Student Clearinghouse matching, by control and level of institution and student type: 2017–18

Control and level of institution and student type	Sample	Matched for NPSAS school[1]		Matched for another school[1]	
		Number	Percent	Number	Percent
Total	**349,650**	**289,220**	**82.7**	**57,250**	**16.4**
Control of institution					
Public	253,230	221,990	87.7	44,540	17.6
Private nonprofit	70,010	59,360	84.8	9,030	12.9
Private for-profit	26,410	7,870	29.8	3,680	13.9
Level of institution					
Less-than-2-year	7,530	460	6.1	590	7.8
2-year	115,670	95,400	82.5	20,380	17.6
4-year, non-doctorate-granting	84,620	67,280	79.5	15,050	17.8
4-year, doctorate-granting	141,840	126,090	88.9	21,240	15.0
Control and level of institution					
Public less-than-2-year	1,060	400	37.8	150	14.0
Public 2-year	108,660	94,010	86.5	19,470	17.9
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	17,880	14,960	83.7	3,040	17.0
Public 4-year, non-doctorate-granting, primarily baccalaureate	26,480	22,230	84.0	5,620	21.2
Public 4-year, doctorate-granting	99,160	90,390	91.2	16,260	16.4
Private nonprofit less-than-4-year	1,120	440	39.7	150	13.0
Private nonprofit 4-year, non-doctorate-granting	32,620	27,210	83.4	4,440	13.6
Private nonprofit 4-year, doctorate-granting	36,280	31,710	87.4	4,450	12.3
Private for-profit less-than-2-year	6,380	60	0.9	430	6.8
Private for-profit 2-year	5,970	950	15.9	770	12.8
Private for-profit 4-year	14,050	6,860	48.8	2,480	17.7
Student type					
Undergraduate	324,790	270,320	83.2	54,880	16.9
Graduate	24,870	18,900	76.0	2,370	9.5

[1] Sample members matched to only the NPSAS-year enrollment period (July1, 2017–June 30, 2018).
NOTE: 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, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

4.2.4 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 2011–12 and 2016–17 academic years. In total, 127,940 sample members (37 percent) matched to ACT (table 19). The match rate by control and level of institution ranged from 8 percent for students sampled from private for-profit 4-year institutions to 49 percent for students sampled from public 4-year, doctorate-granting institutions. Match rates also varied by student type, with 39 percent of undergraduate students having an ACT record on file for the matched years, and only 6 percent of the graduate students having records in the database.

NPSAS staff also obtained the most recent student records of SAT and questionnaire data for high school graduation years 2012–17. As shown in table 19, staff matched SAT data records for 66,060 sample members (19 percent). Match rates by control and level of institution ranged from 3 percent of students from private for-profit 4-year institutions and from public less-than-2-year institutions to 31 percent of students from private nonprofit 4-year, doctorate-granting institutions.

Table 19. Results of ACT and SAT matching, by control and level of institution and student type: 2017–18

Control and level of institution and student type	Sample	Matched to ACT[1]		Matched to SAT[2]	
		Number	Percent	Number	Percent
Total	**349,650**	**127,940**	**36.6**	**66,060**	**18.9**
Control of institution					
Public	253,230	97,770	38.6	45,170	17.8
Private nonprofit	70,010	26,800	38.3	19,600	28.0
Private for-profit	26,410	3,370	12.8	1,290	4.9
Level of institution					
Less-than-2-year	7,530	1,320	17.6	540	7.2
2-year	115,670	35,010	30.3	12,670	11.0
4-year, non-doctorate-granting	84,620	28,810	34.0	16,630	19.6
4-year, doctorate-granting	141,840	62,810	44.3	36,230	25.5
Control and level of institution					
Public less-than-2-year	1,060	290	27.5	40	3.3
Public 2-year	108,660	33,560	30.9	12,200	11.2
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	17,880	5,100	28.5	2,580	14.4
Public 4-year, non-doctorate-granting, primarily baccalaureate	26,480	10,230	38.6	5,430	20.5
Public 4-year, doctorate-granting	99,160	48,590	49.0	24,930	25.1
Private nonprofit less-than-4-year	1,120	280	25.1	160	13.9
Private nonprofit 4-year, non-doctorate-granting	32,620	12,770	39.1	8,300	25.5
Private nonprofit 4-year, doctorate-granting	36,280	13,750	37.9	11,140	30.7
Private for-profit less-than-2-year	6,380	1,020	15.9	510	7.9
Private for-profit 2-year	5,970	1,170	19.7	310	5.3
Private for-profit 4-year	14,050	1,180	8.4	470	3.3
Student type					
Undergraduate	324,790	126,430	38.9	65,250	20.1
Graduate	24,870	1,510	6.1	810	3.3

[1] Sample members were matched to the 2011–12 through 2016–17 academic years for ACT scores.
[2] Sample members were matched to high school graduation years 2012–17 for SAT scores.
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, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

4.2.5 Veterans Benefits Administration (VBA)

Veterans education benefits information was obtained for 23,130 of the sample members with a recorded SSN (7 percent), as shown in table 20. Match rates across control and level of institution ranged from 5 percent of cases sampled

from private for-profit less-than-2-year institutions to 21 percent at private for-profit 4-year institutions. Undergraduate students matched to VBA data at a rate of 7 percent, and graduate students had a match rate of 16 percent.

Table 20. Results of Veterans Benefits Administration (VBA) matching, by control and level of institution and student type: 2017–18

Control and level of institution and student type	Sample	Sent to VBA	Matched to VBA	
			Number	Percent
Total	**349,650**	**318,740**	**23,130**	**7.3**
Control of institution				
Public	253,230	232,160	15,120	6.5
Private nonprofit	70,010	61,790	4,560	7.4
Private for-profit	26,410	24,790	3,440	13.9
Level of Institution				
Less-than-2-year	7,530	6,900	400	5.7
2-year	115,670	107,530	7,350	6.8
4-year, non-doctorate-granting	84,620	77,940	6,950	8.9
4-year, doctorate-granting	141,840	126,370	8,430	6.7
Control and level of institution				
Public less-than-2-year	1,060	900	90	9.4
Public 2-year	108,660	101,040	6,890	6.8
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	17,880	16,770	1,110	6.6
Public 4-year, non-doctorate-granting, primarily baccalaureate	26,480	24,450	1,730	7.1
Public 4-year, doctorate-granting	99,160	89,000	5,310	6.0
Private nonprofit less-than-4-year	1,120	1,060	80	7.4
Private nonprofit 4-year, non-doctorate-granting	32,620	29,620	2,520	8.5
Private nonprofit 4-year, doctorate-granting	36,280	31,100	1,960	6.3
Private for-profit less-than-2-year	6,380	5,920	310	5.2
Private for-profit 2-year	5,970	5,490	380	6.9
Private for-profit 4-year	14,050	13,370	2,750	20.6
Student type				
Undergraduate	324,790	296,490	19,580	6.6
Graduate	24,870	22,250	3,550	15.9

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, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

Chapter 5. Data File Processing and Preparation

NPSAS:18-AC student- and institution-level data were compiled from institution student records 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 details on data file processing, editing, and documentation.

5.1 Overview of the NPSAS:18-AC Study Files

The primary analysis files (or derived files) for NPSAS:18-AC contain data for 266,910 student respondents and include more than 450 variables. Data were subject to quality checks during editing of institution record data. A primary analysis file was created for both undergraduate and graduate student respondents.

Complete data for NPSAS:18-AC are in restricted-use files (RUFs) and are documented in detail in the associated codebooks. RUFs are available to researchers who have applied for and received authorization from NCES to access RUFs. Researchers may obtain authorization by contacting the Institute of Education Sciences (IES) Data Security Office.[15] The NPSAS:18-AC RUFs are listed in table 21. SAT, ACT, NSC, and VBA data were also used to create derived variables, in combination with data from other sources (institution record data, CPS, NSLDS). The SAT, ACT, NSC, and VBA data files are not available as source files due to agreements decided upon with the organizations that own the data.

[15] More information on obtaining RUFs can be found at https://nces.ed.gov/statprog/rudman.

Table 21. NPSAS:18-AC restricted-use files, by file name, description, and file path: 2017–20

File name	Description	File path
NPSAS undergraduate analysis (derived) file	Contains analytic variables derived from all NPSAS:18-AC data sources for the 245,200 undergraduate student respondents.	/DATA/DERIVED/UNDERGRADUATE/np18derivedug_datafile.csv
NPSAS graduate analysis (derived) file	Contains analytic variables derived from all NPSAS:18-AC data sources for the 21,720 graduate student respondents.	/DATA/DERIVED/GRADUATE/np18derivedgr_datafile.csv
Student records data file	Contains data collected from institution records of the 266,910 student respondents.	/DATA/SOURCE/NP18STUDRECS/np18studrecs_datafile.csv
Imputation flag	Contains imputation flags for any NPSAS:18-AC derived variable that was imputed for 266,910 NPSAS:18-AC student respondents.	/DATA/SOURCE/NP18FLAG/np18flag_datafile.csv
Institution data	Contains institution-level data for NPSAS:18-AC sample members' NPSAS institution. There are 1,910 institutions represented.	/DATA/SOURCE/NP18INSTITUTION/np18institution_datafile.csv
CPS 2017–18 data	Contains data received from CPS for the 162,280 student respondents who matched to the 2017–18 financial aid application files.	/DATA/SOURCE/NP18CPS18/np18cps18_datafile.csv
CPS 2018–19 data	Contains data received from CPS for the 128,730 student respondents who matched to the 2018–19 financial aid application files.	/DATA/SOURCE/NP18CPS19/np18cps19_datafile.csv
CPS 2019–20 data	Contains data received from CPS for the 68,250 student respondents who matched to the 2019–20 financial aid application files.	/DATA/SOURCE/NP18CPS20/np18cps20_datafile.csv
GIS census tract data	Contains one observation per student respondent with census tract-level characteristics based on geocoded location information of each student respondent's residence in 2017–18 matched to the 2013–17 (5-year) estimates from the America Community Survey from the U.S. Census Bureau.	/DATA/SOURCE/NP18GIS/np18gis_datafile.csv
NSLDS loan	Contains loan-level data received from NSLDS for 146,760 matched student respondents who had received federal loans as of early 2019. This file includes one record for each federal loan received by these student respondents and provides the most recent information for that loan.	/DATA/SOURCE/NP18NSLDSLOAN/np18nsldsloan_datafile.csv
NSLDS loan disbursement	Contains loan-disbursement level data from NSLDS for 146,320 matched student respondents who borrowed federal loans as of early 2019. This file includes one record for each disbursement made on a federal loan to student respondents.	/DATA/SOURCE/NP18NSLDSLOANDIS/np18nsldsloandis_datafile.csv
NSLDS award origin	Contains student-award-year-level data on federal Direct Loans awarded to 143,770 student respondents as of early 2019. This file includes one record for each student and year during which the student was awarded a federal Direct Loan between 2012 and 2019. 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/NP18NSLDSAWARDORIGIN/np18nsldsawardorigin_datafile.csv
NSLDS program enrollment status	Contains student-school-program-level enrollment information from NSLDS for 179,340 student respondents. This file includes one record for each program and enrollment status change for a student respondent as reported to NSLDS.	/DATA/SOURCE/NP18NSLDSENROLLPROG/np18nsldsenrollprog_datafile.csv
NSLDS Pell Grant data	Contains Pell Grant data received from NSLDS for 128,930 student respondents who received a federal grant as of early 2019. This file includes one record for each federal grant received by these student respondents and provides distribution information for that grant.	/DATA/SOURCE/NP18NSLDSPELL/np18nsldspell_datafile.csv

See notes at end of table.

Table 21. NPSAS:18-AC restricted-use files, by file name, description, and file path: 2017–20—Continued

File name	Description	File path
NSLDS certification data	Contains student-school-level information from NSLDS on the certification of the accuracy of enrollment data for 183,700 student respondents, as reported through the Student Status Confirmation Report process. This file contains one record for each institution a student attended.	/DATA/SOURCE/NP18NSLDSCERT/np18nsldscert_datafile.csv
NSLDS enrollment data	Contains student-school-level enrollment information from NSLDS for 185,590 student respondents. This file includes one record for each institution a student attended.	/DATA/SOURCE/NP18NSLDENROLL/np18nsldsenroll_datafile.csv
NSLDS enrollment program status data	Contains student-school-program-level enrollment information from NSLDS for 179,340 student respondents. This file includes one record for each program a student enrolled in at each institution attended.	/DATA/SOURCE/NP18NSLDSENROLLPROG/np18nsldsenrollprog_datafile.csv
NSLDS financial profile data	Contains records of the financial profile of 185,680 student respondents and their relatives. Relatives include parents, father, mother, and spouse. This file includes one record for each student.	/DATA/SOURCE/NP18NSLDSFINANCPROF/np18nsldsfinancprof_datafile.csv
NSLDS gainful employment data	Contains student-school-program-level information for 31,940 student respondents on the Gainful Employment data elements of records submitted to NSLDS.	/DATA/SOURCE/NP18NSLDSGE/np18nsldsge_datafile.csv
NSLDS student demographic data	Contains award-year-level information on student demographics for 186,000 student respondents. This file contains results from CPS's daily compute process for a student in NSLDS. It retains all transactions processed by CPS for award years 2007–08 and forward and only the most recent transaction for award years 2006–07 and earlier.	/DATA/SOURCE/NP18NSLDSSTUDDEM/np18nsldsstudem_datafile.csv
NSLDS student demographic supplement data	Contains award-year-level information on supplemental student demographics for 162,180 student respondents.	/DATA/SOURCE/NP18NSLDSSTUDDEMSUPP/np18nsldsstudemsupp_datafile.csv
NPSAS:18-AC weights	Contains final NPSAS:18-AC weight and variance estimation variables as a separate record for each student respondent.	/DATA/SOURCE/NP18WEIGHTS/np18weights_datafile.csv
NPSAS:18-AC weights history	Contains intermediate weight adjustment factors as well as final student weights and variance estimation variables as a separate record for each student respondent.	/DATA/SOURCE/NP18WEIGHTH/np18weighth_datafile.csv

NOTE: COD = Common Origination and Disbursement; CPS = Central Processing System; GIS = geographic information system; NPSAS = National Postsecondary Student Aid Study; NSLDS = National Student Loan Data System.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2018 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

5.2 Post–Data Collection Editing

During data collection, quality control checks were performed on all information collected from the institution student records to ensure the quality and accuracy of data. This review included identifying consistently missing items, which often led to correcting the submitted data through a discussion with the institution. Once an institution's data passed the initial quality control review stage, they were extracted and placed into a SAS dataset for further processing, which included various activities to ensure consistency across student records data.

The data cleaning and editing process for the NPSAS:18-AC data files was a multistage process that consisted of the following:

- **Sanitization.** Verbatim character strings, such as financial aid program information and major studies, were sanitized to ensure the integrity of the data and confidentiality of the respondents. Character strings provided by institutions were censored by redacting PII that could be used to identify respondents.

- **Value recoding.** Data were reviewed to ensure that each variable contained valid and consistent values. Invalid, yet understandable, entries were recoded as needed to ensure that data were not lost. For example, an institution might have reported a student's enrollment status to be "Full" instead of the requested value, "Full-Time." In such cases, the enrollment status was recoded to "Full-Time." This process was executed programmatically so that the value "Full" was only converted once. Consistency with prior student records collections, such as NPSAS:16, was ensured by using a collection of common recodes created during earlier student records collections.

- **Financial aid program review.** Financial aid programs were thoroughly reviewed to ensure consistent and accurate categorization. For example, if the name of an aid program indicated that it was a state merit grant but the award was inadvertently categorized by the institution as an institution merit grant, the source for the award was changed from "institution" to "state." This process was also executed programmatically so that aid programs were reviewed once per institution.

- **CIP code review.** All major fields of study that contained invalid or blank CIP codes were systematically reviewed using a coding application. This application was used to review the major text field and provide a valid CIP code where possible. This process supported the encoding of valid, consistent, and accurate majors.

In addition to the preceding activities, the following steps were conducted in editing the student records data files for release:

- Logical recodes of the data were performed when the value of missing items could be determined from answers to previous questions. For example, if the institution reported a student as earning an undergraduate degree but a response to the student's grade level was missing, the value for the student's grade level was set to "undergraduate, unclassified" and the edit was documented in the codebook.

- All missing data were examined to assign specific values indicating the cause of the missing data (table 22). For example, gate–nest question

relationships were defined, and data were examined for adherence to logic established in the student records instrument design. ("Gate" questions are those that must logically be answered first before subsequent "nest" questions are asked.)

- Minimum, maximum, mean, and median values of continuous variables were examined to assess reasonableness of responses. Anomalous distributions and values were investigated and corrected or documented. If the value in question was unacceptable, the value was replaced with a missing data code (table 22) and the edit was documented in the codebook.

- Similar and related items were cross-tabulated to verify the proper relationships between variables and reserve codes held. Note that, in some cases, institutions provided values that appeared inconsistent, but values were not always edited if it was unclear which value should be retained and which was in error.

Table 22. Description of missing data codes: 2017–18

Missing data code	Data label	Description
−1	Don't know	Institution did not have this information for the student.
−3	Not applicable	Item does not apply to the student.
−4	Missing—unable to determine applicability	Institution either did not provide an answer to a gate item or did not have information on the gate item; therefore, it cannot be determined whether any nested items apply to the student.
–5	Missing—implied no/zero	Item was left blank by institution, but a response was provided for either other students at the institution or other items in the group; when all provided responses are nonmissing, no/zero could be implied as responses for missing or blank items.
–6	Out of range	Institution provided a response that was outside of the determined range for this item.
−9	Missing—response not provided	Institution did not provide an answer.

NOTE: In the institution file, Integrated Postsecondary Education Data System data use a value of -2 to indicate "not applicable."
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

While data were processed and edited, question wording, response options, logical recoding, and the "applies to" text were documented for each delivered variable from the student records data collection (appendix D).

5.3 Composite and Derived Variable Construction

The derived variables were derived 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, derived or composite variables were created by assigning the value from the available source with the highest priority. In other cases, items were recoded or

combined to create a derived variable. For a listing of the variables derived for NPSAS:18-AC, see appendix J. Further detail on variable derivation is available in PowerStats on the "Get more info" tab for each variable and in the RUF codebooks.

Due to the exclusion of the student survey, NPSAS:18-AC represents a departure in multiple areas from prior years of NPSAS. This section outlines the changes and their reasons.

5.3.1 Dropped Variables

Some derived variables that traditionally appear in NPSAS iterations and that include the student survey as a significant response source could not be created. The causes of missingness for these derived variables included the following:

- The variable had a single source from the student survey and had no alternative administrative sources.
- A significant number of responses for the variable came from the student survey and could not be filled in with administrative data.
- Responses to the student survey provided additional information needed to create the variable from administrative sources.

Table 23 shows the name and label of every NPSAS variable excluded from NPSAS:18-AC for these reasons. The variable name and label are drawn from NPSAS:16, the most recent NPSAS study prior to NPSAS:18-AC.

Table 23. NPSAS:16 derived variables excluded from NPSAS:18-AC: 2017–18

NPSAS:16 variable name	NPSAS:16 variable label
ASIANTYPE	Asian type
AWAREIDR	Aware of income-driven student loan repayment plans
AWARELFP	Aware of student loan forgiveness programs
BANK1	Bank accounts: had checking or savings account
BANK2	Bank accounts: individual or shared
BORAMT1	Cumulative amount borrowed for undergrad
BORAMT2	Cumulative amount borrowed for grad
BORAMT3	Cumulative amount borrowed for undergrad and grad
CAGI	Adjusted Gross Income (AGI)
CRBALCR	Credit cards: balance carried over each month
CRBALDUE	Credit cards: balance due on all credit cards
CRNUMCRD	Credit cards: number of credit cards in own name
CRTUIT	Credit cards: used credit cards to pay tuition and fees in 2015–16
CRTUIT2	Credit cards: only source available to pay tuition and fees in 2015–16
CSTBKSDG	Cost of digital textbooks (student reported)
CSTBKSPR	Cost of print-only textbooks (student reported)
CSTOTHER	Cost of other required course materials (student reported)
CSTSUPP	Cost of required supplies (student reported)
CSTTECH	Cost of required technology (student reported)
CSTTOTAL	Total cost of required course materials (student reported)
CUMLNTP1	Cumulative loan type for undergrad
CUMLNTP2	Cumulative loan type for grad
CUMLNTP3	Cumulative loan type for undergrad and grad
DECMAJ	Formally declared major field of study
DEGPR	Prior degree earned since high school
DEGPRAA	Prior degree: associate’s degree
DEGPRBA	Prior degree: 4-year bachelor’s degree
DEGPRCRT	Prior degree: undergraduate certificate/diploma
DEGPRDOC	Prior degree: doctorate or professional degree
DEGPRHIGH	Highest prior degree earned since high school
DEGPRMS	Prior degree: master's degree
DEGPRPTB	Prior degree: postbachelor’s certificate
DEGPRPTM	Prior degree: postmaster's certificate
DEPCARE	Dependents: child in paid childcare
DEPCOLCS	Amount contributed to college costs for dependents in 2015–16
DEPCOST	Dependents: children in paid childcare—monthly costs
DEPNUMCH	Number of dependent children
DEPNUMOT	Number of dependent(s) other than children
DEPOTCST	Dependents: monthly cost of supporting dependents other than children
DEPYNG	Dependents: age of youngest child
DIS16A	Disability: deaf or serious difficulty hearing
DIS16B	Disability: blind or serious difficulty seeing
DIS17B	Disability: serious difficulty walking or climbing stairs
DISABIL2	Disability: has some type of disability
DISMNTAL	Disability: serious difficulty concentrating, remembering, deciding
DISTALL	Distance education: entire program at NPSAS was online
DISTEDUC	Distance education: any classes taken at NPSAS taught entirely online
DISTYPE2	Disability: main type of condition or impairment
DSTUINC	Dependent students' income
ELAPSE	Number of months between high school completion and postsecondary entry
EVER2PUB	Ever attended community college
EVER4YR	Ever attended 4-year institution
FAMHELP	Other financial support received

See notes at end of table.

Table 23. NPSAS:16 derived variables excluded from NPSAS:18-AC: 2017–18—Continued

NPSAS:16 variable name	NPSAS:16 variable label
FAMHPAM	Help from family and friends: total amount in 2015–16
FEDBEN	Received any federal benefits
FEDBENA	Received food stamp benefits
FEDBENB	Received free or reduced-price lunch benefits
FEDBENC	Received supplemental security income benefits
FEDBEND	Received Temporary Assistance to Needy Families benefits
FEDBENE	Received WIC benefits
FIN2000	Financial security: $2,000 within the next month
FINLIT1	Financial literacy: Effect of inflation on purchasing
FINLIT2	Financial literacy: Effect of interest on savings
FINLIT3	Financial literacy: Effect of diversification on risk
FINLITALL	Financial literacy: Correctly answered all three items
FINLITTOT	Financial literacy: Number of correct responses
GAINSUR	Graduate assistantship: included health insurance
GRADGAP	Years between bachelor's and graduate school
GRADGAPM	Months between bachelor's and graduate school
GRADPYM	Year and month began graduate degree
GRADPYR	Year began graduate degree
GRJOBHR	Graduate school job: hours worked per week
GRJOBWK	Graduate school job: proportion of weeks worked
HIGHLVEX	Highest level of education ever expected
HSTYPE	Type of high school last attended
HOMESTUD	Student owns home or pays mortgage
HSCRDAP	Took Advanced Placement courses while in high school
HSCRDCOL	Took college-level courses while in high school
HSCRDIB	Took International Baccalaureate courses while in high school
HSGRADYM	High school completion date (year and month)
HSTKANY	Took any college credits in high school
IMMIAGE	Age when arrived in the United States
IMMIGEN	Immigrant generational status
IMMIGRA	Immigrant status
ISTUINC	Independent students' income
JOBANY	Job: Had job while enrolled in 2015–16 (excluding work-study)
JOBANY_GR	Job: Had job while enrolled in 2015–16 (excluding assistantship/traineeship)
JOBANY2	Job: Had job while enrolled in 2015–16 (including work-study)
JOBANY2_GR	Job: Had job while enrolled in 2015–16 (incl. assistantship/traineeship)
JOBANY3	Job: Had job at any time in 2015–16 (including work-study)
JOBANY3_GR	Job: Had job at any time in 2015–16 (including assistantship/traineeship)
JOBENR	Jobs while enrolled: work intensity (excluding work-study)
JOBENR_GR	Jobs while enrolled: work intensity (excluding assistantship/traineeship)
JOBENR2	Jobs while enrolled: work intensity (including work-study)
JOBENR2_GR	Jobs while enrolled: work intensity (including assistantship/traineeship)
JOBENR3	Job: work intensity (including work-study)
JOBHOUR	Jobs while enrolled: Hours worked per week (excluding work-study)
JOBHOUR_GR	Jobs while enrolled: Hours per week (excluding assistantship/traineeship)
JOBHOUR2	Jobs while enrolled: Hours worked per week (including work-study)
JOBHOUR2_GR	Jobs while enrolled: Hours per week (including assistantship/traineeship)
JOBHOUR3	Job: Hours worked per week (including work-study)
JOBMAJOR	Jobs while enrolled: Related to major or field of study (excluding work-study)
JOBMAJOR2	Jobs while enrolled: Related to major or field of study (including work-study)
JOBMAJOR3	Job: Related to major or field of study (including work-study)
JOBONOFF	Jobs while enrolled: Located primarily on or off campus (excluding work-study)
JOBRATE	Jobs while enrolled: Hourly earnings rate (excluding work-study) in 2015–16
JOBRATE_GR	Jobs while enrolled: Hourly earnings (excluding assistantship/traineeship)

See notes at end of table.

Table 23. NPSAS:16 derived variables excluded from NPSAS:18-AC: 2017–18—Continued

NPSAS:16 variable name	NPSAS:16 variable label
JOBRATE2	Jobs while enrolled: Hourly earnings rate (including work-study) in 2015–16
JOBRATE3	Job: Hourly earnings rate (including work-study) in 2015–16
JOBROLE	Job: Primary role as student or employee (excluding work-study)
JOBROLE_GR	Job: Primary role as student or employee
JOBROLE2	Job: Primary role as student or employee (including work-study)
JOBTYPE2	Type of job student had while enrolled (including work-study)
JOBTYPE2_GR	Type of job student had while enrolled (including graduate school jobs)
LNREPAY	Expect help with repaying student loans
LOANLIT1	Loan literacy: Govt. can report unpaid debt to credit bureaus
LOANLIT2	Loan literacy: Govt. can garnish wages for unpaid fed. loan debt
LOANLIT3	Loan literacy: Govt. can retain tax refunds, Soc Sec for loan debt
LOANLITALL	Loan literacy: Correctly answered all three items
LOANLITTOT	Loan literacy: Number of correct responses
LOCALEST	Degree of urbanization of student's permanent address
MAJCHGFQ	Majors: frequency of formally changed
MNTRENT	Monthly mortgage or rent amount
NFEDCUM1	Cumulative nonfederal loan amount for undergrad
NFEDCUM2	Cumulative nonfederal loan amount for grad
NFEDCUM3	Cumulative nonfederal loan amount for undergrad and grad
NUMJOB	Jobs while enrolled: Number in 2015–16 (excluding work-study)
NUMJOB_GR	Jobs while enrolled: Number in 2015–16 (excluding assistantship/traineeship)
NUMJOB2	Jobs while enrolled: Number in 2015–16 (including work-study)
NUMJOB2_GR	Jobs while enrolled: Number in 2015–16 (including assistantship/traineeship)
NUMJOB3	Job: Number in 2015–16 (including work-study)
NUMJOB3_GR	Job: Number in 2015–16 (including assistantship/traineeship)
OWEAMT1	Amount still owed on undergrad
OWEAMT2	Amount still owed on grad
OWEAMT3	Amount still owed on undergrad and grad
PAGI	Parents' adjusted gross income
PAR1	Parent type 1 (for parents' highest education)
PAR2	Parent type 2 (for parents' highest education)
PARBORN	Parent born in U.S., Puerto Rico, or U.S. territory
PARFEMED	Female parent's highest education level
PARHELP	Help from parents: housing, tuition, and other expenses
PARHELP2	Help from parents: housing, tuition, other expenses
PARHPAMT	Help from parents: amount parents helped pay for expenses in 2015–16
PARHPAMT2	Help from parents: amount parents helped pay for expenses
PARMALED	Male parent's highest education level
PDADED	Father's highest education level
PFEDTAX	Parents' federal tax paid
PMOMED	Mother's highest education level
PRIMLANG	Primary language spoken
PRIMLGFQ	Frequency of speaking non-English language with primary caregiver
PRIVAMT	Total private sources grants and loans
PRIVLOAN	Private (alternative) loans
PRIVLRAT	Ratio of private loans to total loans
PRIVPACK	Package of private and nonprivate loans
PRIVPCT	Ratio of private loans to total aid
PTAXFILE	Parents' federal tax filed
RAINDTRB	Race: American Indian or Alaska Native recognized tribe
REANOAPA	Reason for not applying: did not want to take on the debt
REANOAPB	Reason for not applying: forms were too much work
REANOAPC	Reason for not applying: no information about how to apply

See notes at end of table.

Table 23. NPSAS:16 derived variables excluded from NPSAS:18-AC: 2017–18—Continued

NPSAS:16 variable name	NPSAS:16 variable label
REANOAPD	Reason for not applying: no need
REANOAPE	Reason for not applying: thought ineligible
REANOAPF	Reason for not applying: other
REFUND1	Received a refund of scholarships, grants, or loans from NPSAS
REFUND2	Method of receiving refund from NPSAS
REMETOOK	Remedial courses: took in 2015–16
REMEVER	Remedial courses: ever taken
REMMATH	Remedial courses: number taken in math in 2015–16
REMRW	Remedial courses: number taken in reading/writing in 2015–16
REMSTAT	Remedial courses: Year taken and subjects taken in 2015–16
RISKIND2	Index of risk and nontraditional students, modified
RISKINDX	Index of risk and nontraditional students
SAGI	Independent students: Adjusted gross income
SFEDTAX	Independent students: Federal tax paid
SFEDTAXD	Dependent student: Federal tax paid
SIBCOLFT	First sibling to go to college
SJHOURS	Work-study job: hours worked per week
SJMAJOR	Work-study job: related to major or coursework
SJONOFF	Work-study job: located primarily on or off campus
SJRATE	Work-study job: hourly earnings rate
SJWKST	Work-study job: had work-study job in 2015–16
SPSINC	Independent students' spouse's income
STABEVR	Ever studied abroad during undergraduate education
STABREG	Study abroad region
STABTIME	Length of time studied abroad
STAXFILE	Independent students' federal tax filed
STAXFILED	Dependent student: Federal tax filed
TETOOK	Took SAT or ACT exams
TOTDUE1	Cumulative amount owed, principal and interest, all loans for undergrad
TOTDUE2	Cumulative amount owed, principal and interest, all loans for grad
TOTDUE3	Cumulative amount owed, principal and interest, all loans for undergrad and grad
USBORN	Born in the U.S. (student)
USEIDR	Likelihood of using income-driven student loan repayment plans
USELFP	Likelihood of using loan forgiveness program

NOTE: NPSAS = National Postsecondary Student Aid Study; WIC = Special Supplemental Nutrition Program for Women, Infants, and Children.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16), and 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

5.3.2 Renamed Variables

For some derived variables, administrative sources did not provide the same information provided by the student survey (i.e., the same response options or degrees of missingness). Rather than exclude these variables entirely from the dataset, alternative approaches were developed during the derivation process, including

- incorporating new sources of administrative data into the derived variable specifications;
- changing the conditions for skipped cases of the derived variable, thus limiting the amount of missingness; and

- excluding missing components of derived variables, thus changing the definition of the renamed derived variable.

Because these variables have different specifications, different logic for skipped cases, and different data sources than in the past, they are not comparable with variables of the same name in prior NPSAS studies. Table 24 reports the old variable name from NPSAS:16; the new variable name from NPSAS:18-AC; and the skipped cases and variable values for both studies, noting where they differ.

Table 24. NPSAS:16 derived variables renamed in NPSAS:18-AC: 2017–18

NPSAS:16 variable name	NPSAS:18-AC variable name	Variable label	Skipped cases in NPSAS:16	Skipped cases in NPSAS:18-AC	NPSAS:16 variable values	NPSAS:18-AC variable values	Notes
GRADLVL	GRADLVL_AC	Graduate class level	Applies to all graduate respondents	Applies to all graduate respondents	0: Not in a degree program 1: First year 2: Second year 3: Third year 4: Fourth year or higher	0: Not in a degree program 1: First year 2: Second year or higher	
HOMELESS2	HOMELESS	At risk of homelessness	Applies to all respondents	AGE >= 24	0: No 1: Yes	0: No 1: Yes	The NPSAS:18-AC variable is comparable to the NPSAS:12 HOMELESS variable.
ORPHAN2	ORPHAN	Orphan, ward of court, emancipated minor, in legal guardianship	Applies to all respondents	AGE >= 24	0: No 1: Yes	0: No 1: Yes	The NPSAS:18-AC variable is comparable to the NPSAS:12 ORPHAN variable.
PARED1	PARED1_AC	Parent 1 highest education level	Applies to all respondents	Applies to all respondents	0: Don't know 1: Did not complete high school 2: High school diploma or equivalent 3: Vocational/technical training 4: Associate's degree 5: Some college but no degree 6: Bachelor's degree 7: Master's degree or equivalent 8: Doctor's degree–professional 9: Doctor's degree–research	1: Middle school/junior high 2: High school 3: College or beyond	
PARED2	PARED2_AC	Parent 2 highest education level	Applies to all respondents	Applies to all respondents	See PARED1.	See PARED1.	
PAREDUC	PAREDUC_AC	Parents' highest education level	Applies to all respondents	Applies to all respondents	See PARED1.	See PARED1.	

See notes at end of table.

Table 24. NPSAS:16 derived variables renamed in NPSAS:18-AC: 2017–18—Continued

NPSAS:16 variable name	NPSAS:18-AC variable name	Variable label	Skipped cases in NPSAS:16	Skipped cases in NPSAS:18-AC	NPSAS:16 variable values	NPSAS:18-AC variable values	Notes
PROGSTAT	PROGSTAT_AC	Completed degree program in NPSAS year	Applies to all respondents	Applies to all respondents	0: No 1: Yes	0: No 1: Yes	The NPSAS:18-AC variable no longer only measures whether the student intended or expected to graduate but also if the student did or did not graduate.
TEACTDER	TEACTDER_AC	ACT derived composite score	AGE >= 30 or TETOOK <= 0	AGE >= 30	Continuous	Continuous	†
TESATDER	TESATDER_AC	SAT derived composite score	AGE >= 30 or TETOOK <= 0	AGE >= 30	Continuous	Continuous	†
ZPARED1	ZPARED1_AC	Source variable for PARED1_AC	†	†	†	†	†
ZPARED2	ZPARED2_AC	Source variable for PARED2_AC	†	†	†	†	†
ZPROGSTAT	ZPROGSTAT_AC	Source variable for PROGSTAT_AC	†	†	†	†	†
ZTEACTDER	ZTEACTDER_AC	Source variable for TEACTDER_AC	†	†	†	†	†
ZTESATDER	ZTESATDER_AC	Source variable for TESATDER_AC	†	†	†	†	†

† Not applicable.
NOTE: NPSAS = National Postsecondary Student Aid Study; NPSAS:12 = 2011–12 National Postsecondary Student Aid Study.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16), and 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

5.3.3 Changes in Total Aid and Aid Ratios

With private loans (PRIVLOAN) excluded in NPSAS:18-AC, several other total aid and aid ratio variables were affected. Importantly, TOTAID8 (Total aid excluding private loans) is equivalent to the new variable TOTAID_AC, and TOTLOAN3 (Total loans excluding private loans) is equivalent to the new variable TOTLOAN_AC. The aid ratio variables on the NPSAS:18-AC derived data files were calculated with total aid and total loan variables unaffected by the exclusion of private loans: Ratios that previously used TOTAID as the denominator use TOTAID8, and ratios that previously used TOTLOAN in the numerator or denominator use TOTLOAN3.

Table 25 lists all total aid and aid ratio variables directly or indirectly affected by missing private loans and indicates how their derivation in NPSAS:18-AC differs from the most recent study, NPSAS:16. Variables that do not include private loans, in either the total or the ratio calculations, are not changed in NPSAS:18-AC. For instance, because the variables TOTAID2, TOTAID8, and TOTLOAN3 do not include PRIVLOAN, they are not affected by the change in NPSAS:18-AC to exclude private loan data.

Table 25. Total aid and aid ratio derived variables affected by missing private loans in NPSAS:18-AC: 2017–18

NPSAS:16 variable name	NPSAS:18-AC variable name	Variable label	NPSAS:18-AC adjustment
AIDCST	AIDCST_AC	Ratio of aid to student budget	Uses TOTAID8 instead of TOTAID
AIDSNEED	AIDSNEED_AC	Aid amount exceeding federal need	Uses TOTAID8 instead of TOTAID
AIDSRC	AIDSRC_AC	Aid package by source of aid	Uses TOTAID8 instead of TOTAID
AIDSRCG	AIDSRCG_AC	Graduate aid package by source of aid	Uses TOTAID8 instead of TOTAID
AIDTYPE	AIDTYPE_AC	Aid package by type of aid	Uses TOTAID8 instead of TOTAID
AIDTYPEG	AIDTYPEG_AC	Graduate aid package by type of aid	Uses TOTAID8 instead of TOTAID
EFFORT20	EFFORT20_AC	Net price after grants and loans as percentage of income	Uses TOTLOAN3 instead of TOTLOAN
FEDGRPCT	FEDGRPCT_AC	Ratio of federal grants to total aid	Uses TOTAID8 instead of TOTAID
FEDPCT	FEDPCT_AC	Ratio of federal aid to total aid	Uses TOTAID8 instead of TOTAID
GLOANSRC	GLOANSRC_AC	Loan package by type of loan	Uses TOTLOAN3 instead of TOTLOAN
GPLUSRAT	GPLUSRAT_AC	Ratio of GPLUS Loans to total loans	Uses TOTLOAN3 instead of TOTLOAN
GRTPCT	GRTPCT_AC	Ratio of grants to total aid	Uses TOTAID8 instead of TOTAID
INSTGPCT	INSTGPCT_AC	Ratio of institution grants to total aid	Uses TOTAID8 instead of TOTAID
INSTPACK	INSTPACK_AC	Aid package with institution aid	Uses TOTAID8 instead of TOTAID
INSTPCT	INSTPCT_AC	Ratio of institution aid to total aid	Uses TOTAID8 instead of TOTAID
LOANCST2	LOANCST2_AC	Ratio of student loans to budget (includes Parent PLUS Loans)	Does not include PRIVLOAN
LOANPCT	LOANPCT_AC	Ratio of loans to total aid (excluding Parent PLUS)	Uses TOTAID8 instead of TOTAID
LOANPCT2	LOANPCT2_AC	Ratio of loans to total aid (including Parent PLUS)	Uses TOTAID8 instead of TOTAID
LOANSRC	LOANSRC_AC	Loan package by source of loan	Uses TOTLOAN3 instead of TOTLOAN
NETCST1	NETCST1_AC	Student budget minus all aid	Uses TOTAID8 instead of TOTAID
NETCST17	NETCST17_AC	Student budget minus all grants and loans	Uses TOTLOAN3 instead of TOTLOAN
NETCST20	NETCST20_AC	Student budget minus all grants and loans (includes Parent PLUS Loans)	Does not include PRIVLOAN
NETCST43	NETCST43_AC	Student budget minus all aid except loans	Uses TOTAID8 instead of TOTAID and TOTLOAN3 instead of TOTLOAN
OTHRSCR	OTHRSCR_AC	Total financial aid from outside sources	Does not include PRIVLOAN
PELLRAT1	PELLRAT1_AC	Ratio of Pell Grant to total aid	Uses TOTAID8 instead of TOTAID
PLUSPCT	PLUSPCT_AC	Ratio of PLUS Loans to total aid	Uses TOTAID8 instead of TOTAID
SNEED2	SNEED2_AC	Student budget minus expected family contribution minus total aid	Uses TOTAID8 instead of TOTAID
STAPCT	STAPCT_AC	Ratio of state aid to total aid	Uses TOTAID8 instead of TOTAID
STGRPCT	STGRPCT_AC	Ratio of state grants to total aid	Uses TOTAID8 instead of TOTAID
TGRTNLN	TGRTNLN_AC	Total loans and grants	Does not include PRIVLOAN
TNFEDAID	TNFEDAID_AC	Total nonfederal loans	Does not include PRIVLOAN
TOTAID	TOTAID_AC	Total aid amount	Does not include PRIVLOAN
TOTAID4	TOTAID4_AC	Total aid excluding Parent PLUS	Does not include PRIVLOAN
TOTAID6	TOTAID6_AC	Total aid excluding Parent PLUS and veterans/DOD	Does not include PRIVLOAN
TOTAID7	TOTAID7_AC	Total aid excluding veterans/DOD	Does not include PRIVLOAN
TOTLOAN	TOTLOAN_AC	Total loans excluding Parent PLUS	Does not include PRIVLOAN
TOTLOAN2	TOTLOAN2_AC	Total loans including Parent PLUS	Does not include PRIVLOAN
UNSBLOAN	UNSBLOAN_AC	Total unsubsidized loans (all sources)	Does not include PRIVLOAN
WORKPCT	WORKPCT_AC	Ratio of work-study to total aid	Uses TOTAID8 instead of TOTAID
ZOTHRSCR	ZOTHRSCR_AC	Source variable for OTHRSCR_AC	†
ZTOTAID	ZTOTAID_AC	Source variable for TOTAID_AC	†
ZTOTLOAN	ZTOTLOAN_AC	Source variable for TOTLOAN_AC	†

† Not applicable.
NOTE: DOD = Department of Defense; GPLUS = Graduate PLUS.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16), and 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

5.3.4 Income Imputation

NPSAS:18-AC does not have information on income that student respondents provide in the student survey. Thus, income is a single-source variable from CPS and is only available for students who filed a FAFSA, so the pattern of missingness in the income variables is not random. Higher income students are less likely to file a FAFSA, due to the lower need for financial assistance, which results in fewer observed income values for higher income students. The result is that the observed data have an insufficient number of students to provide data (donors) for imputation (section 6.4) for higher income NPSAS respondents with missing income information. The normal approach to income imputation would produce an artificially low estimate of the income values for non-FAFSA filers, given that the donors for these nonfilers would be students with lower income values than most non-FAFSA filers.

To account for this issue of nonrandom missingness, a regression-based imputation approach was used to model student income. The models for imputation were trained using income data from NPSAS:12, NPSAS:16, and the U.S. Census Bureau. Models were run separately for dependent undergraduates, independent undergraduates, and graduate students. The model was used to predict the NPSAS:18-AC median income values within groups.[16] These predicted values were then used to impute the actual missing values based on the lognormal probability distribution function.

Quality control of the resulting imputed data entailed comparisons between observed and imputed income measures to ensure that imputed values were reasonably higher than the observed values, particularly for dependent and independent undergraduates. For graduate students, income generally measures graduate assistantships, which do not always require a FAFSA. Thus, the differences between observed and imputed income for graduate students are less substantial.

Due to the differences in how income variables were imputed in NPSAS:16 and NPSAS:18-AC, the continuous income variables for NPSAS:18-AC are available only in the RUF. New categorical income variables were created and included in both the RUF and PowerStats. Table 26 lists the historical continuous income variable names, the new categorical income variable names, the variable label, and the values of the categorical income variable.

[16] These groups were constructed differently for dependent undergraduates, independent undergraduates, and graduate students and include combinations of race, FAFSA filing status, urbanicity, sex, demographics, and degree.

Table 26. NPSAS:16 continuous derived variables changed to categorical variables for NPSAS:18-AC: 2017–18

NPSAS:16 continuous variable name	NPSAS:18-AC new categorical variable name	NPSAS:18-AC variable label	New categorical variable values
CINCOME	BINCOME	Total income	1: Less than $5,000 2: $5,000–$9,999 3: $10,000–$14,999 4: $15,000–$19,999 5: $20,000–$24,999 6: $25,000–$29,999 7: $30,000–$34,999 8: $35,000–$39,999 9: $40,000–$44,999 10: $45,000–$49,999 11: $50,000–$54,999 12: $55,000–$59,999 13: $60,000–$64,999 14: $65,000–$69,999 15: $70,000–$74,999 16: $75,000–$79,999 17: $80,000–$89,999 18: $90,000–$99,999 19: $100,000–$109,999 20: $110,000–$119,999 21: $120,000–$129,999 22: $130,000–$139,999 23: $140,000–$149,999 24: $150,000–$159,999 25: $160,000–$169,999 26: $170,000–$179,999 27: $180,000–$189,999 28: $190,000–$199,999 29: $200,000 or more
DEPINC	DEPINC2	Dependent students: parents' income	See CINCOME.
INDEPINC	INDEPINC2	Independent students: student and spouse's income	See CINCOME.

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2015–16 National Postsecondary Student Aid Study (NPSAS:16), and 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

5.4 State Representativeness

As mentioned in chapter 1, NPSAS:18-AC was designed to be state representative for undergraduate students at the state level. After data collection was completed, state representativeness was determined for all states overall and within state for public 2-year and public 4-year institution sectors. The determination of state representativeness of undergraduates overall and for both institution sectors was informed by institution response rates, student respondent counts, nonresponse bias analysis results (see section 6.2), and substantive knowledge about states. These criteria were a combination of objective and subjective measures that varied by state. Generally, the following informed the determination:

- Institution response rate needed to be at least 50 percent.
- The student response counts were based on power calculations.

- The percentage of characteristics with significant bias and effect size needed to be reasonable relative to other states.
- Key estimates needed to be reasonable compared to external estimates.
- Most of the important and unique institutions (based on multiple characteristics) needed to have participated in the NPSAS:18-AC institution data collection.

Thirty states are representative for undergraduate students overall, 36 states are representative for undergraduate students in public 2-year institutions, and 45 states are representative for undergraduate students in public 4-year institutions. Table 27 lists the final determinations of state representativeness and within-state representativeness for each state in the study. Additionally, appendix I contains maps identifying which states and sectors within states are representative.

Table 27. NPSAS:18-AC state and within-state representativeness, by institution sector: 2017–18

State	Overall	Institution sector	
		Public 2-year	Public 4-year
Alabama	Yes	Yes	Yes
Alaska	No	†	Yes
Arizona	Yes	Yes	Yes
Arkansas	Yes	Yes	Yes
California	Yes	Yes	Yes
Colorado	Yes	Yes	Yes
Connecticut	No	Yes	No
Delaware	Yes	†	Yes
District of Columbia	No	†	No
Florida	Yes	Yes	Yes
Georgia	No	No	Yes
Hawaii	Yes	Yes	Yes
Idaho	No	No	No
Illinois	Yes	Yes	Yes
Indiana	Yes	Yes	Yes
Iowa	Yes	Yes	Yes
Kansas	Yes	Yes	Yes
Kentucky	Yes	Yes	Yes
Louisiana	Yes	Yes	Yes
Maine	No	No	No
Maryland	No	No	Yes
Massachusetts	Yes	Yes	Yes
Michigan	Yes	Yes	Yes
Minnesota	Yes	Yes	Yes
Mississippi	Yes	Yes	Yes
Missouri	Yes	Yes	Yes
Montana	No	No	Yes
Nebraska	No	Yes	No
Nevada	No	Yes	Yes
New Hampshire	No	No	Yes
New Jersey	Yes	Yes	Yes
New Mexico	No	Yes	Yes
New York	Yes	Yes	Yes
North Carolina	Yes	Yes	Yes
North Dakota	No	No	Yes
Ohio	Yes	Yes	Yes
Oklahoma	No	Yes	Yes
Oregon	No	Yes	No
Pennsylvania	Yes	Yes	Yes
Puerto Rico	No	No	No
Rhode Island	No	No	Yes
South Carolina	No	No	Yes
South Dakota	No	No	Yes
Tennessee	Yes	Yes	Yes
Texas	Yes	Yes	Yes
Utah	Yes	Yes	Yes
Vermont	Yes	Yes	Yes
Virginia	Yes	Yes	Yes
Washington	No	No	Yes

See notes at end of table.

Table 27. NPSAS:18-AC state and within-state representativeness, by institution sector: 2017–18 —Continued

State	Overall	Institution sector	
		Public 2-year	Public 4-year
West Virginia	No	No	Yes
Wisconsin	Yes	Yes	Yes
Wyoming	No	Yes	Yes

† Not applicable.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

Chapter 6. Weighting and Variance Estimation

The following chapter provides a general overview of the statistical analysis weights for the NPSAS:18-AC sample, including institution- and student-level weights. Unit-level and item-level nonresponse bias analysis, variances, imputation for missing data, and disclosure procedures to protect student confidentiality are also covered.

6.1 Weighting

A single statistical analysis weight (one weight per student) was computed for students so that student respondents represent the target population. The statistical analysis weight compensates for the unequal probability of selection of institutions and students into the NPSAS:18-AC sample.

The statistical analysis weight was calculated so that it can be used for undergraduate- and graduate-level analysis at the national level and within each category of control and level of institution. In addition, the statistical weight for undergraduates was calculated so that it can be used for analysis at the state level for states in which the sample was sufficient for state-level representation and for analysis within the public 2-year and public 4-year institution sectors within states for sectors in which the sample was sufficient for sector-level representation.

Some states and public 2-year and public 4-year institution sectors within states did not have enough student respondents to be representative. The determination of state representativeness and the representativeness of the public 2-year and public 4-year institution sectors within states was informed by institution response rates, student respondent counts, nonresponse bias analysis results, and substantive knowledge about states.

The weights were adjusted for multiplicity at the student level, unknown student eligibility, and nonresponse and poststratification at both the institution and the student levels. Although the weights for undergraduates and graduates were

computed separately, the weighting process steps described in this section were followed for both undergraduates and graduates.

The steps in developing the weight components are described in sections 6.1.1 and 6.1.2 and include

- creating the institution sampling weight;
- performing institution-level weight adjustments including nonresponse and poststratification adjustments;
- creating the student sampling weight; and
- performing student-level weight adjustments including multiplicity adjustment, unknown eligibility adjustment, nonresponse adjustment, and poststratification adjustment.

The institution weight was first computed and then used as a component to construct the student weight.

Table 28 lists all the components used to construct the student-level analysis weight. Each weight component represents either a probability of selection, a weight adjustment, or an interim weight. These components are described in detail in sections 6.1.1 and 6.1.2.

Table 28. Components of final analysis weight: 2017–18

Weight	Description
WT1	Institution sampling weight
adj1	Institution nonresponse adjustment
WT2	Institution base weight, adjusted for nonresponse; WT2 = WT1 * adj1
adj2	Institution poststratification adjustment
WT3	Institution weight; WT3 = WT2 * adj2
WT4	Student within-institution sampling weight
WT5	Student base weight; WT5 = WT3 * WT4
adj3	Student multiplicity adjustment
WT6	Student base weight, adjusted for multiplicity; WT6 = WT5 * adj3
adj4	Student unknown eligibility status adjustment
WT7	Student base weight, adjusted for multiplicity and unknown eligibility; WT7 = WT6 * adj4
adj5	Student nonresponse adjustment
WT8	Student base weight, adjusted for multiplicity, unknown eligibility, and nonresponse; WT8 = WT7 * adj5
adj6	Student poststratification adjustment
WT9	Final analysis weight, WT9 = WT8 * adj6

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

All nonresponse and poststratification adjustments were computed using the WTADJUST procedure in SUDAAN (RTI International 2012). The WTADJUST

procedure used a constrained logistic model to predict response. The β-parameters of the logistic model, the lower and upper bounds set on the factors, and the centering constant were used to determine the institution nonresponse adjustment (adj1) and all other weight adjustment factors computed by the SUDAAN WTADJUST procedure (the institution poststratification adjustment [adj2], the student nonresponse adjustment [adj5], and the student poststratification adjustment [adj6]).

A key feature of the WTADJUST procedure is that the weight adjustments and weight trimming and smoothing are all accomplished in one step. Two sets of bounds can be defined when using WTADJUST. One set of bounds, which trims the input weight, can be specified on the weight going into the WTADJUST procedure. These bounds trim extremely small and extremely large weights before the WTADJUST procedure. The second set of bounds restrains the values of the weight adjustment factors that are determined by the WTADJUST procedure. Bounds used for nonresponse and poststratification adjustments for the undergraduate and graduate weights are described in the following sections for each weight adjustment.

During model refinement, which involved collapsing categories of candidate predictor variables and/or excluding candidate predictor variables, the WTADJUST procedure ran with no upper or lower limits. Once convergence of the model was achieved, the weight adjustment bounds were tightened to reduce the magnitude of the weight adjustment factors and the variance inflation caused by unequal weighting (referred to as the unequal weighting effect or UWE). The UWE can be inflated by weight adjustments, in addition to unequal probabilities of selection during sampling. The UWE is defined as follows:

$$UWE = n\frac{\sum w^2}{(\sum w)^2},$$

where

n = the sample size and

w = the final student weight.

To be cautious that the bounds for the adjustments were not overtightened, the upper and lower bounds for the adjustment factors were loosened and variables were dropped from the models, as necessary, to ensure that the models still converged, while maintaining reasonable values for the variances, UWEs, and minimum and maximum adjustment factors.

In this way, extreme weights were controlled and the design effects due to unequal weighting were reduced. The WTADJUST procedure was 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 Language Manual* (RTI International 2012).

6.1.1 Institution Weight

The institution weight was calculated in four steps. In the first step, the weight associated with the probability of selection into the sample was calculated. In the second and third steps, nonresponse and poststratification weight adjustments were performed and were incorporated into the final institution weight. In the fourth step, the final institution weight was computed and used as a component of the final student weight.

Institution sampling weight (WT1). The first weight factor, the institution sampling weight, is associated with the selection process of the NPSAS:18-AC institution sample. The institution frame was constructed from the IPEDS 2016–17 IC-H, 2016–17 IC, 2015–16 E12, and 2016–17 C files. A sample of 3,130 institutions was selected that included a census of all public 2-year and all public 4-year institutions and a sample of 1,390 institutions from the "other institutions" stratum. Within the "all other sectors" stratum, institutions were selected using stratified random sampling with probabilities proportional to a composite measure of size, which was based on full-year enrollment by student type (Folsom, Potter, and Williams 1987).

The sampling weight for each sample institution is the reciprocal of its probability of selection in the sample of 3,130 institutions. The probability of selection for institution i was

$$\pi(i) = \begin{cases} \frac{nS(i)}{S(+)} & \text{for noncertainty selections in the "all other sectors" stratum} \\ 1 & \text{for certainty selections,}[17] \end{cases}$$

[17] Noncertainty institutions are institutions selected with a probability less than one, and certainty institutions are institutions selected with a probability of one. The latter include public 2-year and public 4-year institutions, which are a census.

where

n = the sample size in the "other institutions" stratum;

$S(i)$ = the measure of size for the ith institution in the "other institutions" stratum; and

$S(+)$ = the total measure of size of all institutions in the "other institutions" stratum.

Therefore, the institution sampling weight was assigned as follows: WT1 = 1 / π(i).

Institution nonresponse adjustment (adj1). An institution respondent was defined as any sampled institution that provided a student enrollment list from which a student sample was selected.

Of the 3,130 sampled institutions, 3,080 were determined to be eligible. Of the 3,080 eligible institutions, 3,000 enrolled undergraduates and 1,300 enrolled graduates. There were 1,220 eligible institutions that enrolled both undergraduates and graduates. Of the 3,000 eligible sample institutions that enrolled undergraduate students, 2,170 provided enrollment lists (72 unweighted percent). Of the 1,300 eligible sample institutions that enrolled graduate students, 1,020 provided enrollment lists (79 unweighted percent). Of the 1,220 eligible institutions that enrolled both undergraduates and graduates, 980 provided enrollment lists

Institution-level response rates, by control and level of institution, for undergraduate- and graduate-enrolling institutions are shown in table 29.

Institution-level response rates for undergraduate-enrolling institutions, by state and institution stratum, are shown in table 30.

Table 29. Institution-level response rates for undergraduate- and graduate-enrolling institutions, by control and level of institution: 2017–18

Control and level of institution	Undergraduate-enrolling institutions				Graduate-enrolling institutions			
	Number of eligible sampled institutions	Number of respondents[1]	Response rates		Number of eligible sampled institutions	Number of respondents[1]	Response rates	
			Unweighted	Weighted			Unweighted	Weighted
Total	**3,000**	**2,170**	**72.4**	**64.8**	**1,300**	**1,020**	**78.5**	**74.9**
Public less-than-2-year	40	20	47.4	12.6	†	†	†	†
Public 2-year	960	710	73.6	73.1	†	†	†	†
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	150	120	82.0	82.0	30	30	93.8	93.8
Public 4-year, non-doctorate-granting, primarily baccalaureate	230	170	75.2	75.2	190	150	78.9	78.9
Public 4-year, doctorate-granting	360	320	87.5	87.5	380	330	87.5	87.5
Private nonprofit less-than-4-year	30	20	60.6	89.8	†	†	†	†
Private nonprofit 4-year, non-doctorate-granting	370	280	75.7	69.8	260	200	74.7	80.4
Private nonprofit 4-year, doctorate-granting	320	250	78.1	63.0	350	270	77.0	65.1
Private for-profit less-than-2-year	180	100	52.5	62.5	†	†	†	†
Private for-profit 2-year	190	90	46.1	50.7	†	†	†	†
Private for-profit 4-year	170	100	61.4	59.6	100	60	55.3	58.5

† Not applicable.
[1] An institution respondent is defined as any sampled institution that provided a student enrollment list from which a student sample was selected.
NOTE: The weighted response rates were calculated using the institution sampling weight. 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, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

Table 30. Institution-level response rates for undergraduate-enrolling institutions, by state and institution stratum: 2017–18

State and institution stratum	Undergraduate-enrolling institutions			
	Number of eligible sampled institutions	Number of respondents[1]	Response rates	
			Unweighted	Weighted
Alabama				
Overall	60	40	65.6	65.8
Public 2-year	20	20	87.0	87.0
Public 4-year	10	10	71.4	71.4
Alaska				
Overall	10	10	77.8	77.8
Public 2-year	†	†	†	†
Public 4-year	#	#	100.0	100.0
Arizona				
Overall	50	30	61.1	42.9
Public 2-year	20	20	85.0	85.0
Public 4-year	10	10	66.7	66.7
Arkansas				
Overall	60	50	72.6	75.8
Public 2-year	20	20	90.9	90.9
Public 4-year	10	10	63.6	63.6
California				
Overall	180	120	64.2	57.8
Public 2-year	100	60	61.5	61.5
Public 4-year	50	30	72.3	72.3
Colorado				
Overall	60	40	74.6	63.4
Public 2-year	10	10	92.3	92.3
Public 4-year	20	20	88.2	88.2
Connecticut				
Overall	50	40	68.6	71.8
Public 2-year	10	10	76.9	76.9
Public 4-year	10	10	55.6	55.6
Delaware				
Overall	20	10	75.0	75.0
Public 2-year	†	†	†	†
Public 4-year	#	#	100.0	100.0
District of Columbia				
Overall	20	10	64.7	64.7
Public 2-year	†	†	†	†
Public 4-year	#	#	#	100.0
Florida				
Overall	100	70	71.6	50.0
Public 2-year	30	20	60.0	60.0
Public 4-year	40	40	87.8	87.8
Georgia				
Overall	80	60	71.4	81.6
Public 2-year	20	10	39.1	39.1
Public 4-year	30	30	96.3	96.3
Hawaii				
Overall	20	20	77.3	77.3
Public 2-year	10	10	100.0	100.0
Public 4-year	#	#	100.0	100.0

See notes at end of table.

Table 30. Institution-level response rates for undergraduate-enrolling institutions, by state and institution stratum: 2017–18—Continued

State and institution stratum	Undergraduate-enrolling institutions			
	Number of eligible sampled institutions	Number of respondents[1]	Response rates	
			Weighted	Unweighted
Idaho				
Overall	40	20	65.7	65.7
Public 2-year	#	#	75.0	75.0
Public 4-year	#	#	75.0	75.0
Illinois				
Overall	90	70	77.0	77.6
Public 2-year	50	40	77.1	77.1
Public 4-year	10	10	91.7	91.7
Indiana				
Overall	40	40	90.7	74.4
Public 2-year	#	#	100.0	100.0
Public 4-year	10	10	100.0	100.0
Iowa				
Overall	50	40	91.7	93.3
Public 2-year	20	20	93.8	93.8
Public 4-year	#	#	100.0	100.0
Kansas				
Overall	60	40	66.7	57.0
Public 2-year	30	20	64.0	64.0
Public 4-year	10	10	87.5	87.5
Kentucky				
Overall	50	40	76.5	70.3
Public 2-year	20	10	68.8	68.8
Public 4-year	10	10	87.5	87.5
Louisiana				
Overall	60	40	71.7	78.1
Public 2-year	20	10	73.3	73.3
Public 4-year	20	10	81.3	81.3
Maine				
Overall	40	30	72.2	72.2
Public 2-year	10	10	71.4	71.4
Public 4-year	10	10	62.5	62.5
Maryland				
Overall	60	40	64.4	54.1
Public 2-year	20	10	68.8	68.8
Public 4-year	10	10	84.6	84.6
Massachusetts				
Overall	60	50	83.3	72.8
Public 2-year	20	10	75.0	75.0
Public 4-year	10	10	92.9	92.9
Michigan				
Overall	70	60	75.7	66.8
Public 2-year	30	20	92.0	92.0
Public 4-year	20	20	90.5	90.5
Minnesota				
Overall	70	60	88.6	89.5
Public 2-year	30	30	93.5	93.5
Public 4-year	10	10	100.0	100.0

See notes at end of table.

Table 30. Institution-level response rates for undergraduate-enrolling institutions, by state and institution stratum: 2017–18—Continued

State and institution stratum	Undergraduate-enrolling institutions			
	Number of eligible sampled institutions	Number of respondents[1]	Response rates	
			Weighted	Unweighted
Mississippi				
Overall	50	30	57.4	57.4
Public 2-year	20	10	73.3	73.3
Public 4-year	10	10	87.5	87.5
Missouri				
Overall	60	50	77.6	71.5
Public 2-year	20	10	64.7	64.7
Public 4-year	10	10	84.6	84.6
Montana				
Overall	30	20	64.5	64.5
Public 2-year	10	10	80.0	80.0
Public 4-year	10	10	71.4	71.4
Nebraska				
Overall	50	40	76.6	76.6
Public 2-year	10	10	77.8	77.8
Public 4-year	#	#	42.9	42.9
Nevada				
Overall	40	20	63.2	63.2
Public 2-year	#	#	100.0	100.0
Public 4-year	10	10	83.3	83.3
New Hampshire				
Overall	40	30	72.2	72.2
Public 2-year	10	10	85.7	85.7
Public 4-year	10	10	100.0	100.0
New Jersey				
Overall	60	40	71.0	48.2
Public 2-year	20	20	78.9	78.9
Public 4-year	10	10	92.3	92.3
New Mexico				
Overall	40	30	79.1	79.1
Public 2-year	20	20	84.2	84.2
Public 4-year	10	10	77.8	77.8
New York				
Overall	110	90	82.2	60.9
Public 2-year	40	30	81.6	81.6
Public 4-year	40	40	90.2	90.2
North Carolina				
Overall	110	80	76.2	79.7
Public 2-year	60	40	72.9	72.9
Public 4-year	20	10	87.5	87.5
North Dakota				
Overall	30	20	55.2	55.2
Public 2-year	#	#	40.0	40.0
Public 4-year	10	10	88.9	88.9
Ohio				
Overall	100	80	77.8	58.8
Public 2-year	30	30	84.8	84.8
Public 4-year	30	30	82.4	82.4

See notes at end of table.

Table 30. Institution-level response rates for undergraduate-enrolling institutions, by state and institution stratum: 2017–18—Continued

State and institution stratum	Undergraduate-enrolling institutions			
	Number of eligible sampled institutions	Number of respondents[1]	Response rates	
			Weighted	Unweighted
Oklahoma				
Overall	60	40	57.8	54.1
Public 2-year	20	20	68.2	68.2
Public 4-year	20	10	75.0	75.0
Oregon				
Overall	50	40	70.6	60.5
Public 2-year	20	10	70.6	70.6
Public 4-year	10	10	77.8	77.8
Pennsylvania				
Overall	90	70	74.7	63.1
Public 2-year	20	10	70.6	70.6
Public 4-year	40	40	81.8	81.8
Puerto Rico				
Overall	50	30	70.8	89.3
Public 2-year	#	#	20.0	20.0
Public 4-year	10	10	50.0	50.0
Rhode Island				
Overall	20	20	81.0	81.0
Public 2-year	#	#	100.0	100.0
Public 4-year	#	#	100.0	100.0
South Carolina				
Overall	60	40	66.7	74.7
Public 2-year	20	10	45.0	45.0
Public 4-year	10	10	84.6	84.6
South Dakota				
Overall	30	20	70.4	70.4
Public 2-year	#	#	80.0	80.0
Public 4-year	10	10	85.7	85.7
Tennessee				
Overall	80	50	66.2	58.1
Public 2-year	40	30	71.1	71.1
Public 4-year	10	10	50.0	50.0
Texas				
Overall	130	110	78.9	48.5
Public 2-year	60	50	78.3	78.3
Public 4-year	40	40	90.9	90.9
Utah				
Overall	40	30	63.4	51.3
Public 2-year	#	#	100.0	100.0
Public 4-year	10	10	100.0	100.0
Vermont				
Overall	20	10	59.1	59.1
Public 2-year	#	#	100.0	100.0
Public 4-year	#	#	80.0	80.0
Virginia				
Overall	70	50	73.1	74.6
Public 2-year	20	20	70.8	70.8
Public 4-year	20	10	80.0	80.0

See notes at end of table.

Table 30. Institution-level response rates for undergraduate-enrolling institutions, by state and institution stratum: 2017–18—Continued

State and institution stratum	Undergraduate-enrolling institutions			
	Number of eligible sampled institutions	Number of respondents[1]	Response rates	
			Weighted	Unweighted
Washington				
Overall	70	60	78.6	84.6
Public 2-year	10	10	75.0	75.0
Public 4-year	40	20	65.7	65.7
West Virginia				
Overall	50	30	59.6	55.5
Public 2-year	10	10	45.5	29.1
Public 4-year	10	10	100.0	100.0
Wisconsin				
Overall	60	50	76.7	76.5
Public 2-year	20	10	87.5	87.5
Public 4-year	20	10	86.7	86.7
Wyoming				
Overall	10	10	60.0	60.0
Public 2-year	#	#	57.1	57.1
Public 4-year	#	#	100.0	100.0

† Not applicable.
# Rounds to zero.
[1] An institution respondent is defined as any sampled institution that provided a student enrollment list from which a student sample was selected.
NOTE: The weighted response rates were calculated using the institution sampling weight. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. For the purposes of this table, Puerto Rico and the District of Columbia are referred to as "states."
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

Institution nonresponse adjustment model. Candidate predictor variables for the institution nonresponse adjustment model were selected because they were thought to predict response status based on knowledge of NPSAS data and were nonmissing data for most respondents and nonrespondents. If needed, missing data were recoded into an "unknown" category when there was a small percentage of missing data. The weight used in the model was WT1 multiplied by the institution's undergraduate or graduate enrollment totals determined by the IPEDS 2017–18 IC-H, IC, and E12 files.

The candidate predictor variables for the nonresponse model included many of the variables used in the institution-level nonresponse bias analysis (section 6.2.1). The following descriptive variables were used for the institution nonresponse adjustment as well as the institution nonresponse bias analysis for both undergraduate- and graduate-enrolling institutions:

- 2015 Carnegie Basic Classification (categorical);
- region of institution (categorical);
- HBCU status (yes/no); and

- HSI status (yes/no).[18]

Variables related to the institutions' student population were created separately for undergraduate- and graduate-enrolling institutions and reflected characteristics of the undergraduate and graduate populations, respectively, at each institution.[19] These additional variables were also used for the institution nonresponse adjustment and bias analysis for both undergraduate- and graduate-enrolling institutions and included the following:

- total undergraduate or graduate student enrollment (categorical);
- male undergraduate or graduate student enrollment (categorical);
- female undergraduate or graduate student enrollment (categorical);
- percentage of undergraduate or graduate students enrolled who were Black, non-Hispanic (categorical);
- percentage of undergraduate or graduate students enrolled who were Asian or Pacific Islander, non-Hispanic (categorical); and
- percentage of undergraduate or graduate students enrolled who were Hispanic (categorical).

The following variables[20] were used for the institution nonresponse adjustment and nonresponse bias analysis for only undergraduate-enrolling institutions:

- public 2-year, public 4-year, and other institution sectors within the state where the institution was located;
- number of full-time, first-time undergraduate students living on campus who were receiving Title IV aid (categorical);
- number of full-time, first-time undergraduate students with incomes up to $30,000 who were receiving Title IV aid (categorical);
- percentage of full-time, first-time undergraduate students receiving federal grant aid (categorical);

[18] An HSI indicator is no longer available from IPEDS, so an HSI proxy was created following the definition of HSI provided by the U.S. Department of Education (https://www2.ed.gov/programs/idueshsi/definition.html) and using IPEDS Hispanic enrollment data.
[19] For the continuous variables, categories were formed based on quartiles.
[20] For the continuous variables, categories were formed based on quartiles.

- percentage of full-time, first-time undergraduate students receiving state or local grant aid (categorical);
- percentage of full-time, first-time undergraduate students receiving institution grant aid (categorical);
- percentage of full-time, first-time undergraduate students receiving student loan aid (categorical); and
- average net price among full-time, first-time undergraduates receiving grant or scholarship aid (categorical).

The following variables[21] were used for the institution nonresponse adjustment and nonresponse bias analysis for only graduate-enrolling institutions:

- control and level of institution (categorical);
- degree of urbanization (categorical);[22] and
- total institution employee count of office and administrative support staff (categorical).

Predictors used in the nonresponse modeling included the predictor variables listed above as well as potentially important two-way and three-way interactions. To identify these interactions, the chi-square automatic interaction detection (CHAID) algorithm (Kass 1980) was used. 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. The most significant optimal partition was retained, and the CHAID algorithm was applied to the members of that partition to find further partitions using the remaining predictors. The algorithm was stopped 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.

All predictor variables were retained in the institution nonresponse weight adjustment models unless the models failed to converge. When the models failed to converge, the nonresponse adjustment models were reduced sequentially using a backward stepwise method until converging models were attained.

[21] For the continuous variables, categories were formed based on quartiles.
[22] Degree of urbanization is an IPEDS variable representing the urbanicity (city/suburb/rural) by population size of the institution's location.

The following variables[23] were considered for the nonresponse weight adjustment models but were excluded from the final models:

- average amount of grant and scholarship aid received (categorical) by undergraduate students; and
- percentage of undergraduate students receiving any grant aid (categorical).

The final model on which the nonresponse adjustment, adj1, was based is a generalized exponential model. The exact formula for the weight adjustment factors calculated by the SUDAAN WTADJUST procedure is in the *SUDAAN Language Manual* (RTI International 2012).

For the institution nonresponse weight adjustment, a lower bound of 1 was set on the weight adjustment factors coming out of the nonresponse weight adjustment procedure for both undergraduate- and graduate-enrolling institutions, and an upper bound of 25 was set for graduate-enrolling institutions. The undergraduate-enrolling institution nonresponse weight adjustment factors did not have an upper bound.

Table 31 and table 32 show the average weight adjustment and the final model predictor variables used to determine the institution nonresponse weight adjustments in the undergraduate- and graduate-enrolling institution models, respectively. Summary statistics of the weight adjustment factors follow:

- minimum: 1.00 for both undergraduate- and graduate-enrolling institutions;
- median[24]: 1.17 for both undergraduate- and graduate-enrolling institutions; and
- maximum: 117.56 for undergraduate-enrolling institutions and 15.39 for graduate-enrolling institutions.

[23] For the continuous variables, categories were formed based on quartiles.
[24] The median includes institutions that have no student respondents.

Table 31. Weight adjustment factors for institution nonresponse adjustment for undergraduate-enrolling institutions, by state and institution stratum: 2017–18

State and institution stratum	Number of respondents[1]	Weighted response rate[2]	Average weight adjustment factor (adj1)
Total	**2,170**	**79.0**	**1.69**
Alabama			
Public 2-year	20	85.3	1.15
Public 4-year	10	73.0	1.37
Other	10	68.6	2.15
Alaska			
Public 4-year	#	100.0	1.00
Other	#	76.9	1.30
Arizona			
Public 2-year	20	92.3	1.08
Public 4-year	10	76.8	1.74
Other	10	88.4	1.50
Arkansas			
Public 2-year	20	88.9	1.12
Public 4-year	10	71.2	1.76
Other	20	77.2	1.49
California			
Public 2-year	60	60.9	1.60
Public 4-year	30	76.0	1.31
Other	20	53.3	2.23
Colorado			
Public 2-year	10	97.5	1.02
Public 4-year	20	91.8	1.09
Other	20	33.4	3.43
Connecticut			
Public 2-year	10	87.4	1.12
Public 4-year	10	57.9	1.73
Other	20	65.3	1.80
Delaware			
Public 4-year	#	100.0	1.00
Other	10	94.9	1.10
District of Columbia			
Public 4-year	#	100.0	1.00
Other	10	66.6	2.06
Florida			
Public 2-year	20	70.7	1.99
Public 4-year	40	91.4	1.12
Other	20	56.0	1.85
Georgia			
Public 2-year	10	45.8	2.21
Public 4-year	30	94.9	1.06
Other	20	65.8	1.53
Hawaii			
Public 2-year	10	100.0	1.00
Public 4-year	#	100.0	1.00
Other	10	88.2	1.13
Idaho			
Public 2-year	#	96.9	1.04
Public 4-year	#	55.0	2.57
Other	20	9.7	31.42

See notes at end of table.

Table 31. Weight adjustment factors for institution nonresponse adjustment for undergraduate-enrolling institutions, by state and institution stratum: 2017–18—Continued

State and institution stratum	Number of respondents[1]	Weighted response rate[2]	Average weight adjustment factor (adj1)
Illinois			
Public 2-year	40	78.8	1.33
Public 4-year	10	97.8	1.03
Other	20	67.6	1.41
Indiana			
Public 2-year	#	100.0	1.00
Public 4-year	10	100.0	1.00
Other	20	91.4	1.09
Iowa			
Public 2-year	20	98.0	1.02
Public 4-year	#	100.0	1.00
Other	30	94.5	1.04
Kansas			
Public 2-year	20	53.2	1.95
Public 4-year	10	92.6	1.10
Other	20	53.6	2.92
Kentucky			
Public 2-year	10	58.8	1.70
Public 4-year	10	85.9	1.20
Other	20	84.6	1.21
Louisiana			
Public 2-year	10	77.1	1.36
Public 4-year	10	89.3	1.12
Other	20	70.6	1.55
Maine			
Public 2-year	10	70.8	1.26
Public 4-year	10	83.7	1.23
Other	20	79.8	1.53
Maryland			
Public 2-year	10	65.9	1.46
Public 4-year	10	83.5	1.45
Other	20	51.3	2.17
Massachusetts			
Public 2-year	10	74.3	1.36
Public 4-year	10	99.2	1.01
Other	30	85.4	1.17
Michigan			
Public 2-year	20	76.6	1.21
Public 4-year	20	93.1	1.07
Other	10	45.7	2.30
Minnesota			
Public 2-year	30	98.4	1.01
Public 4-year	10	100.0	1.00
Other	20	61.4	1.83
Mississippi			
Public 2-year	10	73.9	1.37
Public 4-year	10	81.6	1.21
Other	10	67.4	2.60

See notes at end of table.

Table 31. Weight adjustment factors for institution nonresponse adjustment for undergraduate-enrolling institutions, by state and institution stratum: 2017–18—Continued

State and institution stratum	Number of respondents[1]	Weighted response rate[2]	Average weight adjustment factor (adj1)
Missouri			
Public 2-year	10	74.5	1.45
Public 4-year	10	78.8	1.37
Other	20	89.0	1.14
Montana			
Public 2-year	10	88.9	1.19
Public 4-year	10	83.5	1.19
Other	10	41.0	5.71
Nebraska			
Public 2-year	10	75.6	1.66
Public 4-year	#	73.1	1.68
Other	30	87.7	1.32
Nevada			
Public 2-year	#	100.0	1.00
Public 4-year	10	94.9	1.06
Other	20	69.4	1.46
New Hampshire			
Public 2-year	10	85.9	1.15
Public 4-year	10	100.0	1.00
Other	20	97.2	1.22
New Jersey			
Public 2-year	20	78.8	1.26
Public 4-year	10	95.8	1.06
Other	20	57.8	1.82
New Mexico			
Public 2-year	20	89.2	1.12
Public 4-year	10	83.5	1.29
Other	10	81.4	1.22
New York			
Public 2-year	30	86.6	1.20
Public 4-year	40	92.4	1.09
Other	20	75.6	1.30
North Carolina			
Public 2-year	40	81.0	1.27
Public 4-year	10	95.3	1.06
Other	20	73.7	1.37
North Dakota			
Public 2-year	#	41.1	3.46
Public 4-year	10	91.1	1.18
Other	10	65.9	5.79
Ohio			
Public 2-year	30	96.4	1.03
Public 4-year	30	77.6	1.58
Other	20	66.1	1.51
Oklahoma			
Public 2-year	20	84.3	1.26
Public 4-year	10	94.4	1.11
Other	10	43.5	2.80

See notes at end of table.

Table 31. Weight adjustment factors for institution nonresponse adjustment for undergraduate-enrolling institutions, by state and institution stratum: 2017–18—Continued

State and institution stratum	Number of respondents[1]	Weighted response rate[2]	Average weight adjustment factor (adj1)
Oregon			
Public 2-year	10	87.3	1.17
Public 4-year	10	72.2	1.95
Other	20	66.2	1.75
Pennsylvania			
Public 2-year	10	88.1	1.15
Public 4-year	40	78.0	1.40
Other	20	65.3	1.50
Puerto Rico			
Public 2-year and public 4-year	10	60.9	1.74
Other	30	84.9	1.17
Rhode Island			
Public 2-year	#	100.0	1.00
Public 4-year	#	100.0	1.00
Other	10	97.2	1.13
South Carolina			
Public 2-year	10	42.6	2.98
Public 4-year	10	86.5	1.18
Other	20	74.3	1.35
South Dakota			
Public 2-year	#	96.7	1.03
Public 4-year	10	86.9	1.14
Other	10	37.8	5.83
Tennessee			
Public 2-year	30	75.0	1.51
Public 4-year	10	66.3	1.55
Other	20	70.9	1.41
Texas			
Public 2-year	50	85.1	1.16
Public 4-year	40	92.7	1.09
Other	20	59.0	1.68
Utah			
Public 2-year	#	100.0	1.00
Public 4-year	10	100.0	1.00
Other	20	74.4	1.64
Vermont			
Public 2-year	#	100.0	1.00
Public 4-year	#	90.1	1.21
Other	10	70.8	2.05
Virginia			
Public 2-year	20	82.6	1.32
Public 4-year	10	92.5	1.16
Other	20	81.7	1.49
Washington			
Public 2-year	10	82.2	1.19
Public 4-year	20	74.4	1.35
Other	30	99.1	1.01
West Virginia			
Public 2-year	10	63.7	1.73
Public 4-year	10	100.0	1.00
Other	10	93.2	1.70

See notes at end of table.

Table 31. Weight adjustment factors for institution nonresponse adjustment for undergraduate-enrolling institutions, by state and institution stratum: 2017–18—Continued

State and institution stratum	Number of respondents[1]	Weighted response rate[2]	Average weight adjustment factor (adj1)
Wisconsin			
Public 2-year	10	84.4	1.17
Public 4-year	10	86.0	1.19
Other	20	55.7	1.80
Wyoming			
Public 2-year	#	55.0	1.77
Public 4-year	#	100.0	1.00
Other	#	9.5	10.50

Rounds to zero.
[1] An institution respondent is defined as any sampled institution that provided a student enrollment list from which a student sample was selected.
[2] The weighted response rates were calculated using the institution base weight multiplied by full-year student enrollment from the 2017–18 Integrated Postsecondary Education Data System files. Response rates are expressed as percentages.
NOTE: Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding. For the purposes of this table, Puerto Rico and the District of Columbia are referred to as "states."
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

Table 32. Weight adjustment factors for institution nonresponse adjustment for graduate-enrolling institutions: 2017–18

Model predictor variables	Number of respondents[1]	Weighted response rate[2]	Average weight adjustment factor (adj1)
Total	**1,020**	**81.1**	**1.43**
Institution stratum			
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	30	70.6	1.18
Public 4-year, non-doctorate-granting, primarily baccalaureate	150	77.8	1.34
Public 4-year, doctorate-granting	330	88.4	1.15
Private nonprofit 4-year, non-doctorate-granting	200	70.7	1.73
Private nonprofit 4-year, doctorate-granting	270	77.4	1.36
Private for-profit 4-year	60	72.4	2.67
Carnegie Classification			
Associate's	20	97.2	1.02
Research and doctoral	260	85.2	1.17
Master's	460	77.0	1.37
Baccalaureate	160	67.7	1.67
Special focus and other	120	72.6	1.86
Unavailable or unknown	10	69.7	2.21
Degree of urbanization			
Large city	260	78.7	1.49
Mid-size city	140	86.1	1.37
Small city	150	84.0	1.33
Large suburb	170	78.4	1.44
Mid-size suburb	40	91.4	1.32
Small suburb	30	94.5	1.11
Urban area on fringe of town	40	81.5	1.12
Urban area distant from town	90	80.0	1.54
Urban area remote from town	70	82.4	1.38
Rural area on fringe of town	30	67.0	2.00
Rural area distant from town	10	68.7	1.93
Rural area remote from town	10	89.2	1.25

See notes at end of table.

Table 32. Weight adjustment factors for institution nonresponse adjustment for graduate-enrolling institutions: 2017–18—Continued

Model predictor variables	Number of respondents[1]	Weighted response rate[2]	Average weight adjustment factor (adj1)
Bureau of Economic Analysis region[3]			
New England	90	89.5	1.19
Mideast	150	75.9	1.51
Great Lakes	150	86.4	1.23
Plains	120	79.4	1.52
Southeast	250	82.5	1.34
Southwest	90	88.1	1.17
Rocky Mountains	40	80.9	1.42
Far West	110	70.5	2.06
Outlying Areas	20	70.5	1.65
Historically Black College or University			
Yes	30	70.8	1.51
No or Unavailable or unknown	990	81.2	1.42
Hispanic-Serving Institution			
Yes	120	80.3	1.39
No	910	81.2	1.43
Total graduate enrollment			
1–308	240	64.9	1.90
309–1,084	260	76.2	1.34
1,085–3,046	250	75.7	1.33
3,047 or more	280	83.8	1.19
Total male graduate enrollment			
0-96	240	70.9	1.79
97–356	250	74.1	1.44
357–1,180	250	75.9	1.34
1,181 or more	280	83.9	1.18
Total female graduate enrollment			
0–184	240	67.1	1.89
185–683	260	73.2	1.38
684–1,869	250	77.8	1.28
1,870 or more	270	83.6	1.20
Percentage of graduate students enrolled who were Black, non-Hispanic			
0–3	300	79.0	1.55
4–7	270	87.2	1.26
8–14	200	77.6	1.48
15-100	230	79.2	1.39
Unknown	30	42.7	1.73
Percentage of graduate students enrolled who were Asian or Pacific Islander, non-Hispanic			
0–1	280	72.1	1.59
2–3	270	79.3	1.29
4–6	230	89.2	1.18
7-90	210	78.5	1.62
Unknown	30	42.7	1.73

See notes at end of table.

Table 32. Weight adjustment factors for institution nonresponse adjustment for graduate-enrolling institutions: 2017–18—Continued

Model predictor variables	Number of respondents[1]	Weighted response rate[2]	Average weight adjustment factor (adj1)
Percentage of graduate students enrolled who were Hispanic			
0–2	240	74.6	1.55
3–4	260	87.9	1.20
5–9	270	78.8	1.58
10–100	220	82.2	1.32
Unknown	30	42.7	1.73
Total institution employee count of office and administrative support staff			
1–32	230	74.0	1.59
33–82	250	69.1	1.53
83–206	260	80.5	1.24
207–3,274	270	86.3	1.16
Unknown	10	53.4	5.69
CHAID segments			
Public 4-year, non-doctorate-granting, primarily subbaccalaureate or Public 4-year, doctorate-granting institution	360	88.4	1.15
Private nonprofit 4-year, doctorate-granting or Private for-profit 4-year institution where the Percentage of graduate students enrolled who were Asian or Pacific Islander, non-Hispanic is between 0 and 3, and the total institution employee count of office and administrative support staff is between 1 and 32	50	77.4	1.28
Private nonprofit 4-year, doctorate-granting or Private for-profit 4-year institution where the Percentage of graduate students enrolled who were Asian or Pacific Islander, non-Hispanic is between 0 and 3, and the total institution employee count of office and administrative support staff is between 33 and 3,274 or is unknown	100	69.5	1.72
Private nonprofit 4-year, doctorate-granting or Private for-profit 4-year institution where the Percentage of graduate students enrolled who were Asian or Pacific Islander, non-Hispanic is between 4 and 6	90	88.2	1.10
Private nonprofit 4-year, doctorate-granting or Private for-profit 4-year institution where the Percentage of graduate students enrolled who were Asian or Pacific Islander, non-Hispanic is between 7 and 90 or is unknown, and the total male graduate enrollment is between 0 and 1,180	40	51.0	3.27
Private nonprofit 4-year, doctorate-granting or Private for-profit 4-year institution where the Percentage of graduate students enrolled who were Asian or Pacific Islander, non-Hispanic is between 7 and 90 or is unknown, and the total male graduate enrollment is greater than or equal to 1,181 or is unknown	50	79.0	1.25
Public 4-year, non-doctorate-granting, primarily baccalaureate or Private nonprofit 4-year, non-doctorate-granting institution where the Degree of urbanization is not equal to large city, large suburb, or urban area on fringe of town, and the Percentage of graduate students enrolled who were Asian or Pacific Islander, non-Hispanic is between 0 and 3	170	64.5	1.69
Public 4-year, non-doctorate-granting, primarily baccalaureate or Private nonprofit 4-year, non-doctorate-granting institution where the Degree of urbanization is not equal to large city, large suburb, or urban area on fringe of town, and the Percentage of graduate students enrolled who were Asian or Pacific Islander, non-Hispanic is between 4 and 90 or is unknown	40	61.3	2.16

See notes at end of table.

Table 32. Weight adjustment factors for institution nonresponse adjustment for graduate-enrolling institutions: 2017–18—Continued

Model predictor variables	Number of respondents[1]	Weighted response rate[2]	Average weight adjustment factor (adj1)
Public 4-year, non-doctorate-granting, primarily baccalaureate institution where the Degree of urbanization is equal to large city, large suburb, or urban area on fringe of town	50	79.9	1.30
Private nonprofit 4-year, non-doctorate-granting institution where the Degree of urbanization is equal to large city, large suburb, or urban area on fringe of town	80	86.7	1.12

[1] An institution respondent is defined as any sampled institution that provided a student enrollment list from which a student sample was selected.
[2] The weighted response rates were calculated using the institution base weight multiplied by full-year student enrollment from the 2017–18 Integrated Postsecondary Education Data System files. Response rates are expressed as percentages.
[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, Louisiana, 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.
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, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

The institution nonresponse adjustment, adj1, was multiplied by the institution sampling weight, WT1, to obtain the institution base weight, adjusted for nonresponse, WT2 = WT1 * adj1.

Institution poststratification adjustment (adj2). After adjusting for the inverse of the probability of selection into the sample and nonresponse, the nonresponse-adjusted institution weight (WT2) was further adjusted using the WTADJUST procedure to meet enrollment totals (control totals). The enrollment totals were obtained from the IPEDS 2017–18 IC-H, IC, and E12 files. Enrollment totals were determined for public 2-year, public 4-year, and other institutions within state for undergraduate-enrolling institutions and by institution size[25] (*small* vs. *large*) within control and level of institution for graduate-enrolling institutions except for public 4-year, non-doctorate-granting, primarily subbaccalaureate institutions, which were not split by institution size due to small sample size.

This adjustment ensured that the resultant weight adequately represents the student target population. The weight used in the poststratification model was the product of WT1 and adj1 multiplied by the institution's full-year undergraduate or graduate enrollment totals using IPEDS 2017–18 IC-H, IC, and E12 files.

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 "other institutions" were selected with probability proportional to size, with the size being counts of students. This

[25] Institution size was determined based on the median total enrollment as a cut point within each control and level of institution.

method of sampling and weighting does not yield an accurate estimate of institutions, overall or within control and level of institution.

Before the graduate-enrolling institution poststratification adjustment, upper and lower bounds were set on extreme weights to trim the weights. A lower bound of 1.25 was set for 10 graduate-enrolling institutions with extreme low weights, and an upper bound of 6,000 was set for 1 graduate-enrolling institution with an extreme high weight. No trimming was done for the undergraduate-enrolling institution poststratification weights.

The WTADJUST procedure produced the poststratification adjustment factor (adj2). For the graduate-enrolling institution poststratification weight factor, a lower bound of 0.1 and an upper bound of 10 were set on the weight adjustment factors coming out of the poststratification weight adjustment. No bounds were applied to the undergraduate-enrolling institution poststratification weight adjustment factors.

Table 33 and table 34 show the characteristics associated with the control totals and the average weight adjustments factors by these characteristics for the undergraduate- and graduate-enrolling institutions, respectively. Summary statistics of the weight adjustment factors follow:

- minimum: 0.85 for undergraduate-enrolling institutions and 0.19 for graduate-enrolling institutions;
- median: 1.00 for undergraduate-enrolling institutions and 0.98 for graduate-enrolling institutions; and
- maximum: 13.58 for undergraduate-enrolling institutions and 1.22 for graduate-enrolling institutions.

Table 33. Weight adjustment factors for institution poststratification adjustment for undergraduate-enrolling institutions, by state and institution stratum: 2017–18

State and institution stratum	Control total[1]	Average weight adjustment factor (adj2)
Total	**22,891,860**	**1.00**
Alabama		
Public 2-year	116,430	1.01
Public 4-year	153,080	1.00
Other	61,680	1.00
Alaska		
Public 4-year	38,720	1.00
Other	2,960	1.00
Arizona		
Public 2-year	291,690	1.00
Public 4-year	171,960	1.00
Other	253,710	1.02
Arkansas		
Public 2-year	65,090	1.00
Public 4-year	91,280	1.00
Other	21,530	1.01
California		
Public 2-year	1,837,570	1.00
Public 4-year	1,064,130	1.00
Other	474,630	0.85
Colorado		
Public 2-year	94,530	1.00
Public 4-year	217,930	1.00
Other	102,020	0.98
Connecticut		
Public 2-year	69,360	1.00
Public 4-year	59,580	1.03
Other	85,630	0.99
Delaware		
Public 4-year	43,990	1.00
Other	15,980	1.00
District of Columbia		
Public 4-year	5,040	1.00
Other	55,850	1.00
Florida		
Public 2-year	78,790	1.01
Public 4-year	929,000	1.00
Other	351,540	1.05
Georgia		
Public 2-year	163,640	1.19
Public 4-year	317,720	1.00
Other	112,910	1.08
Hawaii		
Public 2-year	34,750	1.00
Public 4-year	27,150	1.00
Other	13,470	1.00
Idaho		
Public 2-year	41,660	1.00
Public 4-year	53,880	1.00
Other	88,910	1.00

See notes at end of table.

Table 33. Weight adjustment factors for institution poststratification adjustment for undergraduate-enrolling institutions, by state and institution stratum: 2017–18—Continued

State and institution stratum	Control total[1]	Average weight adjustment factor (adj2)
Illinois		
Public 2-year	519,320	1.00
Public 4-year	150,590	1.00
Other	238,800	0.99
Indiana		
Public 2-year	162,760	1.00
Public 4-year	209,060	1.01
Other	88,970	0.96
Iowa		
Public 2-year	132,040	1.00
Public 4-year	70,650	1.00
Other	95,340	0.99
Kansas		
Public 2-year	128,240	1.00
Public 4-year	86,660	1.00
Other	40,950	1.01
Kentucky		
Public 2-year	106,830	1.00
Public 4-year	114,790	1.00
Other	49,490	1.01
Louisiana		
Public 2-year	90,800	1.00
Public 4-year	137,650	1.00
Other	47,070	0.97
Maine		
Public 2-year	22,850	1.00
Public 4-year	31,650	1.00
Other	27,490	1.00
Maryland		
Public 2-year	165,380	1.00
Public 4-year	175,320	1.00
Other	45,530	0.97
Massachusetts		
Public 2-year	122,970	1.00
Public 4-year	115,630	1.00
Other	215,190	0.98
Michigan		
Public 2-year	215,270	1.00
Public 4-year	303,850	1.00
Other	94,060	1.01
Minnesota		
Public 2-year	167,410	1.00
Public 4-year	129,700	1.00
Other	98,720	1.03
Mississippi		
Public 2-year	98,120	1.00
Public 4-year	73,770	1.00
Other	21,860	1.00

See notes at end of table.

Table 33. Weight adjustment factors for institution poststratification adjustment for undergraduate-enrolling institutions, by state and institution stratum: 2017–18—Continued

State and institution stratum	Control total[1]	Average weight adjustment factor (adj2)
Missouri		
Public 2-year	124,920	1.00
Public 4-year	140,670	1.00
Other	126,510	1.01
Montana		
Public 2-year	12,150	1.00
Public 4-year	38,190	1.00
Other	5,910	0.98
Nebraska		
Public 2-year	61,710	1.00
Public 4-year	51,650	1.00
Other	30,270	1.00
Nevada		
Public 2-year	15,830	1.00
Public 4-year	113,250	1.00
Other	15,970	1.01
New Hampshire		
Public 2-year	22,320	1.00
Public 4-year	26,050	1.00
Other	122,650	1.00
New Jersey		
Public 2-year	213,670	1.00
Public 4-year	172,870	1.00
Other	92,730	1.02
New Mexico		
Public 2-year	101,180	1.00
Public 4-year	52,210	1.00
Other	6,020	1.00
New York		
Public 2-year	420,270	1.00
Public 4-year	412,160	1.00
Other	502,030	0.98
North Carolina		
Public 2-year	310,940	1.00
Public 4-year	206,770	1.00
Other	107,110	1.00
North Dakota		
Public 2-year	9,770	1.00
Public 4-year	40,240	1.00
Other	6,880	0.99
Ohio		
Public 2-year	262,740	1.00
Public 4-year	308,740	1.00
Other	169,170	1.00
Oklahoma		
Public 2-year	91,430	1.00
Public 4-year	113,700	1.00
Other	46,380	1.00

See notes at end of table.

Table 33. Weight adjustment factors for institution poststratification adjustment for undergraduate-enrolling institutions, by state and institution stratum: 2017–18—Continued

State and institution stratum	Control total[1]	Average weight adjustment factor (adj2)
Oregon		
Public 2-year	160,720	1.00
Public 4-year	105,220	1.00
Other	32,000	1.02
Pennsylvania		
Public 2-year	183,050	1.00
Public 4-year	251,010	1.00
Other	297,410	0.98
Puerto Rico		
Public 2-year	2,650	13.58
Public 4-year	57,190	0.96
Other	170,940	0.93
Rhode Island		
Public 2-year	19,390	1.00
Public 4-year	24,970	1.00
Other	40,870	1.00
South Carolina		
Public 2-year	115,970	1.00
Public 4-year	101,920	1.00
Other	55,510	0.99
South Dakota		
Public 2-year	8,640	1.00
Public 4-year	41,230	1.00
Other	17,560	1.00
Tennessee		
Public 2-year	134,550	1.01
Public 4-year	127,740	1.00
Other	97,470	1.02
Texas		
Public 2-year	1,085,050	1.01
Public 4-year	676,700	1.00
Other	216,660	0.86
Utah		
Public 2-year	54,340	1.00
Public 4-year	173,900	1.00
Other	181,170	1.00
Vermont		
Public 2-year	9,870	1.00
Public 4-year	20,480	0.94
Other	18,460	1.00
Virginia		
Public 2-year	237,250	1.00
Public 4-year	195,010	1.00
Other	155,790	1.02
Washington		
Public 2-year	58,200	1.00
Public 4-year	358,000	1.00
Other	49,100	1.01

See notes at end of table.

Table 33. Weight adjustment factors for institution poststratification adjustment for undergraduate-enrolling institutions, by state and institution stratum: 2017–18—Continued

State and institution stratum	Control total[1]	Average weight adjustment factor (adj2)
West Virginia		
Public 2-year	21,920	0.99
Public 4-year	63,210	1.00
Other	88,730	1.00
Wisconsin		
Public 2-year	135,210	1.00
Public 4-year	191,640	1.00
Other	56,710	1.02
Wyoming		
Public 2-year	29,300	1.00
Public 4-year	10,950	1.00
Other	630	1.00

[1] The weighted response rates were calculated using the institution base weight multiplied by full-year student enrollment from the 2017–18 Integrated Postsecondary Education Data System files.
NOTE: Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding. For the purposes of this table, Puerto Rico and the District of Columbia are referred to as "states."
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

Table 34. Weight adjustment factors for institution poststratification adjustment for graduate-enrolling institutions: 2017–18

Control and level of institution[1]	Control total[2]	Average weight adjustment factor (adj2)
Total	**3,910,440**	**1.00**
Public 4-year, non-doctorate-granting, primarily subbaccalaureate, all	1,650	1.05
Public 4-year, non-doctorate-granting, primarily baccalaureate, small	20,590	0.92
Public 4-year, non-doctorate-granting, primarily baccalaureate, large	143,430	1.01
Public 4-year, doctorate-granting small	297,380	0.98
Public 4-year, doctorate-granting, large	1,349,220	1.01
Private nonprofit 4-year, non-doctorate-granting, small	16,800	0.87
Private nonprofit 4-year, non-doctorate-granting, large	238,040	1.02
Private nonprofit 4-year, doctorate-granting, small	132,300	1.15
Private nonprofit 4-year, doctorate-granting, large	1,296,300	0.98
Private for-profit 4-year, small	8,400	0.95
Private for-profit 4-year, large	406,340	1.06

[1] Size for poststratification weighting classes was based on the median enrollment within sector or state for the institutions on the sampling frame.
[2] Control totals are the sum of enrollment across institutions based on Integrated Postsecondary Education Data System 2017–18 enrollment 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, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

The poststratification adjustment, adj2, was used to define the final institution weight as WT3 = WT2 * adj2 or, equivalently, WT3 = WT1 * adj1 * adj2. The final institution weight was used to develop all student-level analysis weights.

6.1.2 Student Weight

The final analysis weight was calculated through seven steps. First, the student within-institution sampling weight was created. This sampling weight was then multiplied by the institution weight to create the student base weight. This was followed by four weight adjustments, for multiplicity, unknown eligibility, nonresponse, and poststratification. The final student weight was then calculated. Each step is described in this section.

Student within-institution sampling weight (WT4). The overall student sampling strata were defined for undergraduates as the institution sampling strata and for graduates by crossing the institution sampling strata with the student strata.[26] The sampled students were systematically selected from the enrollment lists at institution-specific rates that were inversely proportional to the institution's probability of selection. Specifically, the initial 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 institution-specific student sampling rates were set using IPEDS data before enrollment lists were received. A minimum sample size constraint of 60 students was imposed to ensure sufficient yield for variance estimation. The sample for some institutions was less than 60 because the number of students on the enrollment list was insufficient to achieve a larger sample size. 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 institution-specific 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, the sampling rates were adjusted, when necessary, so that the number of students selected within an institution usually did not exceed 600.

The student within-institution sampling weight was the reciprocal of the adjusted institution-specific student stratum sampling rates, or

[26] See chapter 2 for definitions of the sampling strata.

$$WT4 = 1/f_{s|i}.$$

Student base weight (WT5). The student base weight, denoted WT5, was defined as the product of the institution weight (WT3) and student within-institution sampling weight (WT4): WT5 = WT3 * WT4.

Student multiplicity adjustment (adj3). Students who attended more than one eligible institution during the 2017–18 academic year had multiple chances of being selected. That is, they could have been selected from any of the institutions they attended. Therefore, these students had a higher probability of being selected than was represented in their sampling weight.

This multiplicity was adjusted for by dividing the student base weight (WT5) by the number of institutions attended that were eligible for sample selection. Specifically, the student multiplicity weight adjustment factor was defined as

$$adj3 = 1 / M,$$

where M is the multiplicity, or number of eligible institutions attended. The multiplicity was determined based on the derived variable STUDMULT, which contains a count of institutions attended based on data from NSLDS and NSC. STUDMULT and student multiplicity was set to 1 if there was no evidence to the contrary.

The student multiplicity adjustment factor, denoted adj3, was multiplied by the student base weight (WT5) to form the multiplicity-adjusted student weight (WT6), WT6 = WT5 * adj3.

Summary statistics of the student multiplicity weight adjustment factors follow:

- minimum: 0.11 for undergraduate students and 0.17 for graduate students;
- median: 1.00 for both undergraduate and graduate students; and
- maximum: 1.00 for both undergraduate and graduate students.

Student unknown eligibility status adjustment (adj4). In past NPSAS studies, student eligibility was determined during the student survey. However, because NPSAS:18-AC does not include a student survey, student eligibility for NPSAS:18-AC was based on student records information. Students who had incomplete or conflicting student records data that made it impossible to determine their final eligibility status were treated as eligible. However, their weights (before nonresponse adjustment) were reduced to compensate for the

proportion of known students who were actually ineligible (as described as follows).

Weighting classes were defined by control and level of institution. These weight adjustment factors were simply the eligibility rate estimated among students with known eligibility status. For the known eligible students, the weight adjustment factor was set to 1.

Summary statistics of the student unknown eligibility status weight adjustment factors follow:

- minimum: 0.91 for undergraduate students and 0.70 for graduate students;
- median: 1.00 for both undergraduate and graduate students; and
- maximum: 1.00 for both undergraduate and graduate students.

Table 35 shows the weight adjustment factors applied to the students with unknown eligibility for undergraduate and graduate students.

Table 35. Weight adjustment factors for unknown student eligibility status for undergraduate and graduate students with unknown eligibility, by control and level of institution: 2017–18

	Undergraduate students		Graduate students	
Control and level of institution	Number adjusted for unknown eligibility	Weight adjustment factor (adj4)	Number adjusted for unknown eligibility	Weight adjustment factor (adj4)
Public less-than-2-year	410	0.91	†	†
Public 2-year	25,100	0.97	†	†
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	6,650	0.95	40	0.70
Public 4-year, non-doctorate-granting, primarily baccalaureate	6,590	0.97	130	0.98
Public 4-year, doctorate-granting	15,420	0.99	490	0.99
Private nonprofit less-than-4-year	270	1.00	†	†
Private nonprofit 4-year, non-doctorate-granting	6,220	0.99	660	0.98
Private nonprofit 4-year, doctorate-granting	7,220	0.99	610	0.98
Private for-profit less-than-2-year	2,170	1.00	†	†
Private for-profit 2-year	2,840	0.96	†	†
Private for-profit 4-year	2,090	0.96	920	0.99

† Not applicable.
NOTE: Sample sizes rounded to the nearest 10.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

The student unknown eligibility status adjustment factor, denoted adj4, was multiplied by the multiplicity-adjusted student weight (WT6) to form the student weight, adjusted for multiplicity and unknown eligibility (WT7), WT7 = WT6 * adj4.

Student nonresponse adjustment (adj5), response definition. The rules for defining student respondents were dependent on the student records data elements that were collected and were established separately for undergraduate and graduate students.

An undergraduate student respondent was defined as any sample member who was determined to be eligible for the study, for whom there was evidence of at least 1 month of enrollment,[27] and who, at a minimum, had valid data from student records for the following items:

- federal work-study amount awarded;
- state aid recipient indicator;
- state program and name or type;
- state aid amount by program or type;
- institution aid recipient indicator;
- institution program name or type; and
- institution aid by program name or type.

A graduate student respondent was defined as any sample member who was determined to be eligible for the study, for whom there was evidence of at least 1 month of enrollment from any source, and who, at a minimum, had valid data from student records for the following items:

- institution aid recipient indicator;
- institution program name or type; and
- institution aid by program name or type.

These were the minimum data requirements, but the vast majority of students had considerably more complete data.

Of the 320,160 eligible undergraduate students, 245,200 were respondents (77 unweighted percent, 76 weighted percent). Of the 24,550 eligible graduate students, 21,720 were respondents (89 unweighted percent, 90 weighted percent). Student-level response rates, by control and level of institution, for undergraduate and graduate students are shown in table 36.

[27] Enrollment is determined from student records, NSLDS, and NSC.

Table 36. Student-level response rates for undergraduate and graduate students, by control and level of institution: 2017–18

Control and level of institution	Undergraduate students				Graduate students			
	Number of eligible sampled students	Number of student respondents	Response rates		Number of eligible sampled students	Number of student respondents	Response rates	
			Unweighted	Weighted			Unweighted	Weighted
Total	**320,160**	**245,200**	**76.6**	**75.9**	**24,550**	**21,720**	**88.5**	**90.2**
Public less-than-2-year	1,300	890	68.5	57.3	†	†	†	†
Public 2-year	104,990	79,900	76.1	73.4	†	†	†	†
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	18,980	12,320	64.9	69.7	160	120	77.2	66.0
Public 4-year, non-doctorate-granting, primarily baccalaureate	24,010	17,430	72.6	76.0	1,470	1,340	91.3	87.9
Public 4-year, doctorate-granting	90,690	75,270	83.0	80.5	7,840	7,350	93.8	91.9
Private nonprofit less-than-4-year	1,180	910	77.4	73.3	†	†	†	†
Private nonprofit 4-year, non-doctorate-granting	27,960	21,750	77.8	79.0	3,030	2,370	78.3	84.8
Private nonprofit 4-year, doctorate-granting	30,040	22,810	76.0	83.1	7,070	6,460	91.3	88.6
Private for-profit less-than-2-year	6,240	4,070	65.2	55.9	†	†	†	†
Private for-profit 2-year	6,040	3,210	53.1	34.0	†	†	†	†
Private for-profit 4-year	8,740	6,650	76.1	79.0	5,000	4,080	81.7	93.3

† Not applicable.
NOTE: The weighted response rates were calculated using the student base weight adjusted for multiplicity and unknown eligibility. 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, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

Student-level response rates for undergraduate students overall and by state and institution stratum are shown in table 37.

Table 37. Student-level response rates for undergraduate students, by state and institution stratum: 2017–18

State and institution stratum	Number of eligible sampled students	Number of student respondents	Response rates	
			Unweighted	Weighted
Alabama				
Overall	5,740	4,250	74.1	70.1
Public 2-year	2,260	1,710	75.8	74.0
Public 4-year	2,410	1,740	72.3	65.6
Alaska[1]				
Overall	6,710	5,500	82.0	88.3
Public 2-year	†	†	†	†
Public 4-year	6,540	5,390	82.4	88.9
Arizona				
Overall	5,390	5,070	94.1	87.2
Public 2-year	2,700	2,630	97.5	96.4
Public 4-year	1,280	1,280	99.9	99.9
Arkansas				
Overall	5,840	4,800	82.2	78.5
Public 2-year	2,460	2,060	83.9	82.9
Public 4-year	2,400	1,840	76.8	74.4
California				
Overall	6,860	4,450	64.9	65.1
Public 2-year	3,980	2,470	62.1	58.7
Public 4-year	2,070	1,310	63.4	66.7
Colorado				
Overall	6,430	5,370	83.5	80.5
Public 2-year	1,850	1,470	79.4	71.1
Public 4-year	3,360	2,860	85.3	84.5
Connecticut				
Overall	6,160	4,550	73.8	75.7
Public 2-year	2,240	1,770	79.0	76.4
Public 4-year	1,460	1,220	83.9	83.7
Delaware[1]				
Overall	7,010	6,180	88.2	78.7
Public 2-year	†	†	†	†
Public 4-year	5,130	4,970	96.9	96.7
District of Columbia[1]				
Overall	4,760	2,240	47.0	55.5
Public 2-year	†	†	†	†
Public 4-year	600	270	45.3	45.1
Florida				
Overall	6,270	4,900	78.1	78.9
Public 2-year	1,090	830	76.4	86.7
Public 4-year	3,970	3,300	83.1	83.1
Georgia				
Overall	6,210	3,250	52.4	57.0
Public 2-year	1,000	790	78.6	78.9
Public 4-year	4,060	1,550	38.2	39.1

See notes at end of table.

Table 37. Student-level response rates for undergraduate students, by state and institution stratum: 2017–18—Continued

State and institution stratum	Number of eligible sampled students	Number of student respondents	Response rates	
			Unweighted	Weighted
Hawaii				
Overall	7,070	6,880	97.2	92.1
Public 2-year	3,380	3,350	99.3	99.3
Public 4-year	2,490	2,490	99.8	99.9
Idaho				
Overall	4,270	2,890	67.8	51.2
Public 2-year	1,210	680	56.5	72.5
Public 4-year	1,470	940	63.9	75.2
Illinois				
Overall	6,040	5,420	89.6	87.7
Public 2-year	3,170	2,960	93.4	91.6
Public 4-year	1,480	1,420	95.7	96.3
Indiana				
Overall	7,290	6,280	86.2	85.8
Public 2-year	2,370	2,370	99.8	99.9
Public 4-year	3,240	2,880	88.9	90.8
Iowa				
Overall	6,900	4,950	71.7	65.2
Public 2-year	2,880	2,120	73.6	76.0
Public 4-year	1,530	1,520	99.5	99.5
Kansas				
Overall	5,610	4,990	88.9	89.1
Public 2-year	2,300	1,770	76.8	78.8
Public 4-year	2,290	2,270	99.1	99.0
Kentucky				
Overall	5,470	4,480	81.9	80.7
Public 2-year	1,610	1,030	63.6	60.6
Public 4-year	2,600	2,410	93.0	94.5
Louisiana				
Overall	6,480	5,350	82.5	82.6
Public 2-year	2,000	1,610	80.3	78.9
Public 4-year	3,170	2,650	83.6	85.2
Maine				
Overall	5,100	3,730	73.3	74.7
Public 2-year	1,380	1,300	94.3	97.1
Public 4-year	1,950	1,230	63.1	74.5
Maryland				
Overall	5,930	3,820	64.5	68.3
Public 2-year	2,630	1,240	47.0	47.9
Public 4-year	2,490	1,940	77.9	82.9
Massachusetts				
Overall	6,920	5,740	83.0	83.4
Public 2-year	1,730	1,340	77.5	76.9
Public 4-year	1,780	1,690	94.7	95.5
Michigan				
Overall	6,440	5,640	87.6	88.0
Public 2-year	2,330	1,920	82.1	83.9
Public 4-year	3,300	3,020	91.7	90.8
Minnesota				
Overall	6,940	6,390	92.2	91.1
Public 2-year	3,280	3,140	95.8	95.6
Public 4-year	2,280	2,120	92.9	92.0

See notes at end of table.

Table 37. Student-level response rates for undergraduate students, by state and institution stratum: 2017–18—Continued

State and institution stratum	Number of eligible sampled students	Number of student respondents	Response rates	
			Weighted	Unweighted
Mississippi				
Overall	6,050	4,050	66.9	73.8
Public 2-year	2,820	1,500	53.2	56.5
Public 4-year	2,420	1,830	75.6	88.4
Missouri				
Overall	6,080	5,130	84.3	85.4
Public 2-year	1,800	1,630	90.8	90.2
Public 4-year	2,000	1,700	85.2	87.5
Montana				
Overall	5,810	4,180	72.0	72.5
Public 2-year	1,430	680	47.9	38.3
Public 4-year	3,840	3,240	84.6	81.7
Nebraska				
Overall	4,880	4,270	87.5	91.2
Public 2-year	2,240	2,010	89.9	97.1
Public 4-year	1,200	1,200	99.6	99.7
Nevada				
Overall	7,590	4,450	58.7	35.9
Public 2-year	2,160	2,140	99.3	99.2
Public 4-year	4,380	1,780	40.6	25.2
New Hampshire				
Overall	6,130	3,260	53.2	79.2
Public 2-year	1,220	#	#	#
Public 4-year	1,370	1,140	83.7	80.4
New Jersey				
Overall	6,450	4,680	72.6	68.3
Public 2-year	2,850	1,980	69.5	60.5
Public 4-year	2,480	1,760	70.7	69.9
New Mexico				
Overall	6,000	4,220	70.2	71.6
Public 2-year	3,720	2,730	73.4	73.1
Public 4-year	1,740	1,100	63.2	70.0
New York				
Overall	6,780	4,630	68.4	65.3
Public 2-year	2,290	1,620	70.9	68.5
Public 4-year	2,210	1,640	74.1	67.6
North Carolina				
Overall	7,110	5,890	82.8	82.7
Public 2-year	3,410	2,780	81.4	80.9
Public 4-year	2,290	2,030	88.9	89.0
North Dakota				
Overall	5,040	2,160	42.9	58.1
Public 2-year	510	500	98.2	96.8
Public 4-year	3,990	1,300	32.5	49.3
Ohio				
Overall	6,930	4,470	64.5	66.5
Public 2-year	2,500	2,120	84.5	78.2
Public 4-year	3,010	1,220	40.5	44.1
Oklahoma				
Overall	6,140	5,670	92.3	92.7
Public 2-year	2,360	2,080	87.9	91.5
Public 4-year	3,090	3,000	97.1	96.8

See notes at end of table.

Table 37. Student-level response rates for undergraduate students, by state and institution stratum: 2017–18—Continued

State and institution stratum	Number of eligible sampled students	Number of student respondents	Response rates	
			Weighted	Unweighted
Oregon				
Overall	5,640	4,100	72.8	76.0
Public 2-year	2,980	2,420	81.1	87.7
Public 4-year	1,700	960	56.4	60.8
Pennsylvania				
Overall	6,490	6,060	93.4	87.0
Public 2-year	1,660	1,530	92.5	90.4
Public 4-year	2,600	2,390	91.9	77.2
Puerto Rico				
Overall	5,350	3,540	66.2	59.9
Public 2-year	60	60	100.0	100.0
Public 4-year	970	540	55.5	53.1
Rhode Island				
Overall	7,490	4,140	55.3	57.7
Public 2-year	1,790	#	#	#
Public 4-year	1,990	1,990	99.9	99.9
South Carolina				
Overall	5,470	4,210	77.1	75.2
Public 2-year	1,700	810	47.6	54.6
Public 4-year	2,340	2,190	93.8	93.3
South Dakota				
Overall	6,040	4,180	69.3	68.3
Public 2-year	1,280	960	75.0	61.9
Public 4-year	3,880	2,450	63.1	61.7
Tennessee				
Overall	5,980	4,990	83.5	80.5
Public 2-year	2,500	1,720	68.8	62.9
Public 4-year	1,810	1,800	99.9	99.8
Texas				
Overall	7,460	5,840	78.3	78.7
Public 2-year	4,000	2,950	73.7	72.7
Public 4-year	2,580	2,230	86.7	88.5
Utah				
Overall	6,060	5,450	89.9	93.1
Public 2-year	1,380	1,360	98.9	98.9
Public 4-year	2,770	2,260	81.6	84.8
Vermont				
Overall	5,920	5,320	89.8	95.5
Public 2-year	1,430	1,430	99.9	99.9
Public 4-year	2,440	1,920	78.6	93.3
Virginia				
Overall	6,510	5,640	86.6	89.1
Public 2-year	2,490	2,080	83.7	84.9
Public 4-year	2,250	1,920	85.3	88.0
Washington				
Overall	5,890	4,640	78.7	77.3
Public 2-year	720	560	78.3	77.8
Public 4-year	3,920	3,050	77.9	75.6
West Virginia				
Overall	5,660	4,350	76.9	88.7
Public 2-year	630	450	70.7	71.8
Public 4-year	2,140	1,610	75.5	79.4

See notes at end of table.

Table 37. Student-level response rates for undergraduate students, by state and institution stratum: 2017–18—Continued

State and institution stratum	Number of eligible sampled students	Number of student respondents	Response rates	
			Weighted	Unweighted
Wisconsin				
Overall	6,220	5,160	82.9	81.3
Public 2-year	1,960	1,720	87.9	89.9
Public 4-year	3,110	2,650	85.1	84.3
Wyoming				
Overall	5,180	3,380	65.4	70.9
Public 2-year	3,250	1,500	46.3	50.8
Public 4-year	1,880	1,830	97.4	97.2

† Not applicable.
# Rounds to zero.
[1] There were no public 2-year institutions in Alaska, Delaware, or the District of Columbia.
NOTE: The weighted response rates were calculated using the student base weight adjusted for multiplicity and unknown eligibility. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. For the purposes of this table, Puerto Rico and the District of Columbia are referred to as "states."
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

Student nonresponse adjustment (adj5) model. The model for the adjustment for nonresponse was developed in the following three stages similar to the nonresponse weight adjustments for institutions described previously:

1. identification of model predictors;
2. CHAID analysis to determine significant interactions between model predictors; and
3. WTADJUST procedure to calculate nonresponse adjustment.

Candidate predictor variables for the student nonresponse adjustment model were selected because they were thought to predict response status, based on knowledge of NPSAS data, and were nonmissing for respondents and nonrespondents.

The candidate predictor variables for the student nonresponse model included many of the variables used in the student-level nonresponse bias analysis (section 6.2.2). The following variables[28] were used in the nonresponse adjustment models for both undergraduate and graduate students:

- undergraduate or graduate full-year enrollment from IPEDS 2017–18 file (categorical);
- student age as of December 31, 2017 (categorical);
- sex (categorical);

[28] For the continuous variables, categories were formed based on quartiles. Institution-level variables came from IPEDS 2017–18 files, and student-level variables came from NPSAS:18-AC institution enrollment lists and CPS for aid applicants.

- veteran status (yes/no);
- federal aid recipient (yes/no);
- Direct Loan amount disbursed (categorical);
- state aid recipient (yes/no); and
- institution aid recipient (yes/no).

The following variables were used in the nonresponse adjustment model for only undergraduates:

- state of institution;
- public 2-year and public 4-year institution sectors within state; and
- federal Pell Grant amount awarded (categorical).

The following variables were used in the nonresponse adjustment model for only graduates:

- control and level of institution (categorical);
- region of institution (categorical);
- student type (graduate [excluding doctoral–professional practice] or doctoral–professional practice); and
- race/ethnicity (categorical).

Similar to what was done for the institution nonresponse adjustment, the CHAID algorithm was applied and provided interaction terms for the student nonresponse adjustment models. CHAID was run for up to three levels, identifying two-way and three-way interactions that were the interaction terms in the models used for the WTADJUST procedure.

As with the institution nonresponse models, all variables were retained in the student nonresponse models unless the models failed to converge. The models were reduced sequentially using a backward stepwise method until converging models were attained.

Different bounds on the nonresponse weight adjustments, depending on whether the weight was classified as high extreme,[29] low extreme,[30] or nonextreme, were used to accomplish nonresponse adjustment, truncation, and smoothing in one

[29] High extreme weights were identified as those greater than the median weight plus three times the interquartile range for weights.
[30] Low extreme weights were identified as those less than the median weight minus three times the interquartile range for weights.

step. In addition to using the option in the WTADJUST procedure for weight trimming and smoothing, the data were inspected for extreme weights by checking for outliers by state and institution stratum within state.

Undergraduate nonresponse model. The initial undergraduate nonresponse model included variables that were defined by public 2-year, public 4-year, and other institutions within state. For states with convergence issues, the variables were first collapsed for undergraduate students enrolled in public 2-year and public 4-year institutions before further collapsing into overall state-level variables. Potential representativeness status at the public 2-year, public 4-year, and state levels was used to inform the decision about the collapsing of strata within state.[31] The institution strata used in the final model included

- twelve states with variables defined by public 2-year, public 4-year, and other institutions within state;
- fifteen states with variables defined by public 2-year and public 4-year combined and other institutions within state; and
- twenty-five states with variables defined by state overall.

Table 38 and table 39 show the final predictor variables used to determine the student-level nonresponse weight adjustments and the average weight adjustment factors resulting from the undergraduate student model. Not all final predictor variables were retained for every institution stratum due to removal of variables to resolve convergence issues. CHAID segments for the undergraduate nonresponse adjustment were created separately within each state and institution stratum listed in table 38 and were too numerous and varied to include in table 39. As described above, the models were reduced sequentially using a backward stepwise method until a converging model was attained.

No trimming was done on weights going into the undergraduate student nonresponse adjustment. For the undergraduate student nonresponse weight adjustment, a lower bound of 1 and an upper bound of 5 were set on the weight adjustment factors coming out of the nonresponse weight adjustment. Summary statistics of the nonresponse weight adjustment factors for undergraduate students follow:

- minimum: 1.00;
- median: 1.11; and
- maximum: 132.02.

[31] Final representativeness at the public 2-year, public 4-year, and state levels is discussed in section 5.4.

Table 38. Weight adjustment factors for undergraduate student nonresponse adjustment, by state and institution stratum: 2017–18

State and institution stratum	Number of respondents	Weighted response rate[1]	Average weight adjustment factor (adj5)
Total	**245,200**	**75.9**	**1.35**
Alabama			
Public 2-year and public 4-year	3,450	68.8	1.46
Other	800	75.1	1.36
Alaska			
All	5,500	88.3	1.14
Arizona			
All	5,070	87.2	1.15
Arkansas			
All	4,800	78.5	1.25
California			
Public 2-year	2,470	58.7	1.74
Public 4-year	1,310	66.7	1.61
Other	670	85.1	1.26
Colorado			
Public 2-year and public 4-year	4,330	81.0	1.27
Other	1,030	78.3	1.32
Connecticut			
Public 2-year and public 4-year	3,000	79.9	1.26
Other	1,550	69.7	1.50
Delaware			
All	6,180	78.7	1.30
District of Columbia			
All	2,240	55.5	2.05
Florida			
Public 2-year	830	86.7	1.26
Public 4-year	3,300	83.1	1.20
Other	770	65.2	2.09
Georgia			
Public 2-year	790	78.9	1.27
Public 4-year	1,550	39.1	2.72
Other	910	81.9	1.22
Hawaii			
Public 2-year	3,350	99.3	1.01
Public 4-year	2,490	99.9	1.00
Other	1,040	59.0	1.70
Idaho			
Public 2-year	680	72.5	1.53
Public 4-year	940	75.2	1.50
Other	1,270	32.5	3.76
Illinois			
All	5,420	87.7	1.12
Indiana			
All	6,280	85.8	1.17
Iowa			
All	4,950	65.2	1.64
Kansas			
Public 2-year	1,770	78.8	1.31
Public 4-year	2,270	99.0	1.01
Other	950	93.1	1.08

See notes at end of table.

Table 38. Weight adjustment factors for undergraduate student nonresponse adjustment, by state and institution stratum: 2017–18—Continued

State and institution stratum	Number of respondents	Weighted response rate[1]	Average weight adjustment factor (adj5)
Kentucky			
Public 2-year	1,030	60.6	1.66
Public 4-year	2,410	94.5	1.06
Other	1,040	81.1	1.24
Louisiana			
Public 2-year and public 4-year	4,250	82.8	1.22
Other	1,100	81.4	1.27
Maine			
All	3,730	74.7	1.37
Maryland			
Public 2-year	1,240	47.9	2.38
Public 4-year	1,940	82.9	1.25
Other	650	81.8	1.35
Massachusetts			
All	5,740	83.4	1.21
Michigan			
Public 2-year and public 4-year	4,940	88.2	1.14
Other	700	86.9	1.16
Minnesota			
Public 2-year	3,140	95.6	1.04
Public 4-year	2,120	92.0	1.09
Other	1,130	84.3	1.28
Mississippi			
All	4,050	73.8	1.36
Missouri			
All	5,130	85.4	1.18
Montana			
All	4,180	72.5	1.73
Nebraska			
All	4,270	91.2	1.11
Nevada			
All	4,450	35.9	2.36
New Hampshire			
All	3,260	79.2	1.57
New Jersey			
All	4,680	68.3	1.50
New Mexico			
Public 2-year and public 4-year	3,830	71.8	1.48
Other	390	66.3	1.57
New York			
Public 2-year and public 4-year	3,260	68.0	1.49
Other	1,370	60.6	1.64
North Carolina			
All	5,890	82.7	1.23
North Dakota			
Public 2-year and public 4-year	1,800	57.5	1.73
Other	360	62.2	1.80
Ohio			
Public 2-year and public 4-year	3,340	58.7	1.64
Other	1,130	89.1	1.12

See notes at end of table.

Table 38. Weight adjustment factors for undergraduate student nonresponse adjustment, by state and institution stratum: 2017–18—Continued

State and institution stratum	Number of respondents	Weighted response rate[1]	Average weight adjustment factor (adj5)
Oklahoma			
All	5,670	92.7	1.08
Oregon			
Public 2-year and public 4-year	3,370	75.6	1.32
Other	730	78.8	1.32
Pennsylvania			
Public 2-year	1,530	90.4	1.14
Public 4-year	2,390	77.2	1.21
Other	2,140	93.7	1.07
Puerto Rico			
Public 2-year and public 4-year	600	54.6	1.92
Other	2,950	61.9	1.61
Rhode Island			
All	4,140	57.7	1.78
South Carolina			
Public 2-year and public 4-year	3,000	74.1	1.26
Other	1,220	79.6	1.28
South Dakota			
Public 2-year	960	61.9	1.61
Public 4-year	2,450	61.7	1.61
Other	780	86.4	1.14
Tennessee			
All	4,990	80.5	1.27
Texas			
Public 2-year	2,950	72.7	1.37
Public 4-year	2,230	88.5	1.17
Other	660	70.0	1.64
Utah			
All	5,450	93.1	1.08
Vermont			
All	5,320	95.5	1.05
Virginia			
All	5,640	89.1	1.13
Washington			
Public 2-year and public 4-year	3,620	75.9	1.31
Other	1,020	87.2	1.32
West Virginia			
Public 2-year and public 4-year	2,060	78.0	1.30
Other	2,300	97.0	1.20
Wisconsin			
All	5,160	81.3	1.29
Wyoming			
Public 2-year and public 4-year	3,330	70.3	1.45
Other	50	100.0	1.00

[1] The weighted response rates were calculated using the student base weight adjusted for multiplicity and unknown eligibility.
NOTE: Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding. For the purposes of this table, Puerto Rico and the District of Columbia are referred to as "states."
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

Table 39. Weight adjustment factors for student nonresponse for undergraduate students, by model predictor variable: 2017–18

Model predictor variable[1]	Number of respondents	Weighted response rate	Average weight adjustment factor (adj5)
Age as of December 31, 2017			
15–23	160,560	76.8	1.32
24–29	38,130	74.3	1.40
30 or older	46,500	74.6	1.43
Sex			
Male	106,500	76.5	1.32
Female	138,690	75.5	1.38
Veteran status			
Yes	9,950	73.3	1.40
No	235,250	76.0	1.35
Total amount of federal Pell Grants awarded to undergraduate students[2]			
$0	161,840	78.5	1.29
$1–$2,959	25,180	73.3	1.44
$2,960–$5,919	32,570	71.8	1.45
$5,920 or more	25,610	68.1	1.53
Institution aid status			
Received	72,730	91.4	1.27
Did not receive	172,460	86.2	1.39
State aid status			
Received	44,210	80.8	1.21
Did not receive	200,990	83.3	1.38
Federal aid status			
Received	131,040	73.3	1.43
Did not receive	114,150	78.7	1.27
Total amount of Direct Loans disbursed to undergraduate students[3]			
$0	158,050	76.4	1.31
$1–Q1	22,710	75.2	1.43
> Q1–median	21,010	74.9	1.41
> Median–Q3	24,800	77.1	1.36
> Q3	18,630	72.4	1.52
Undergraduate full-year enrollment[3]			
1–Q1	57,670	69.3	1.51
> Q1–median	58,870	76.5	1.38
> Median–Q3	63,680	74.6	1.28
> Q3	64,960	78.4	1.25

[1] Not all model predictor variables were retained for every institution stratum due to removal of variables to resolve convergence issues.
[2] Pell Grant categories were defined by quartiles.
[3] Direct Loan and enrollment categories were defined by quartiles within the state and institution stratum listed in table 38.
NOTE: The weighted response rates were calculated using the student base weight adjusted for multiplicity and unknown eligibility. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

Graduate nonresponse model. Before the graduate student nonresponse adjustment, a lower bound was set to trim the extreme low weights going into the adjustment. A lower bound of 0.8 was set for students in public 4-year, non-doctorate-granting, primarily subbaccalaureate institutions, and a lower bound of 1 was set for all other students. For the graduate student nonresponse weight

adjustment (adj5), a lower bound of 1 and an upper bound of 25 were set on the weight adjustment factors coming out of the nonresponse weight adjustment. Table 40 shows the final predictor variables used to determine the graduate student-level nonresponse weight adjustments and the average weight adjustment factors resulting from the graduate student model. Summary statistics of the weight adjustment factors follow:

- minimum: 1.00;
- median: 1.03; and
- maximum: 4.49.

Table 40. Weight adjustment factors for graduate student nonresponse adjustment, by model predictor variable: 2017–18

Model predictor variable	Number of respondents	Weighted response rate[1]	Average weight adjustment factor (adj5)
Total	**21,720**	**90.2**	**1.13**
Control and level of institution			
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	120	66.0	1.55
Public 4-year, non-doctorate-granting, primarily baccalaureate	1,340	87.9	1.17
Public 4-year, doctorate-granting	7,350	91.9	1.09
Private nonprofit 4-year, non-doctorate-granting	2,370	84.8	1.25
Private nonprofit 4-year, doctorate-granting	6,460	88.6	1.13
Private for-profit 4-year	4,080	93.3	1.09
Bureau of Economic Analysis (BEA) region[2]			
New England	1,740	84.7	1.21
Mideast	3,650	85.7	1.16
Great Lakes	2,240	87.1	1.22
Plains	2,020	92.1	1.13
Southeast	6,400	95.6	1.07
Southwest	2,410	97.7	1.04
Rocky Mountains	870	99.3	1.01
Far West	2,250	83.9	1.19
Outlying Areas	140	76.6	1.46
Institution graduate full-year enrollment[3]			
1–772	4,990	84.9	1.20
773–3,422	5,620	87.5	1.12
3,423–8,034	5,340	90.1	1.13
8,035 or more	5,770	93.0	1.07
Student type			
Graduate student	18,810	90.0	1.12
Doctoral–professional practice students	2,910	91.7	1.14
Age as of December 31, 2017			
15–23	2,210	88.0	1.15
24–29	7,200	90.1	1.13
30 or older	12,310	90.9	1.12
Veteran status			
Yes	2,280	90.0	1.15
No	19,440	90.2	1.12
Race/ethnicity			
White	10,750	95.5	1.06
Black	3,090	93.4	1.10
Hispanic	1,890	88.4	1.16
Asian	2,270	96.0	1.05
American Indian or Alaska Native	160	93.7	1.10
Native Hawaiian or other Pacific Islander	180	99.8	1.00
More than one race	400	93.4	1.11
Unknown	2,980	70.6	1.45
Sex			
Male	9,500	90.7	1.13
Female	12,220	89.8	1.13

See notes at end of table.

Table 40. Weight adjustment factors for graduate student nonresponse adjustment, by model predictor variable: 2017–18—Continued

Model predictor variable	Number of respondents	Weighted response rate	Average weight adjustment factor (adj5)
Total amount of unsubsidized Direct Loans disbursed to graduate students[3]			
$0	14,080	90.7	1.12
$1–$10,131	1,840	87.5	1.17
$10,132–$19,087	1,870	93.4	1.09
$19,088–$20,500	2,440	86.9	1.17
$20,501 or more	1,500	90.5	1.11
Institution aid status			
Received	6,550	99.7	1.01
Did not receive	15,170	86.9	1.18
State aid status			
Received	560	79.9	1.23
Did not receive	21,160	90.4	1.12
Federal aid status			
Received	8,780	89.5	1.14
Did not receive	12,940	90.6	1.12
CHAID segments			
Graduate students with White or Asian race/ethnicity in the Far West BEA region	1,140	92.8	1.11
Graduate students with White or Asian race/ethnicity in the Mideast, Great Lakes, or Plains BEA regions	5,130	94.8	1.07
Graduate students with White or Asian race/ethnicity in the Rocky Mountain or Outlying Areas BEA regions	620	99.4	1.01
Graduate students with White or Asian race/ethnicity in the New England BEA region	1,150	89.7	1.14
Graduate students with White or Asian race/ethnicity in the Southeast or Southwest BEA regions	4,990	98.6	1.02
Graduate students with Unknown race/ethnicity who did not receive institution aid	2,010	62.8	1.66
Graduate students with Unknown race/ethnicity who received institution aid	970	99.8	1.00
Graduate students with Native Hawaiian or other Pacific Islander race/ethnicity who did not receive institution aid	90	99.3	1.01
Graduate students with Native Hawaiian or other Pacific Islander race/ethnicity who received institution aid	100	100.0	1.00
Graduate students with More than one race, Black, American Indian, or Alaska Native race/ethnicity in the Far West or Outlying Areas BEA regions	390	84.0	1.18
Graduate students with More than one race, Black, American Indian, or Alaska Native race/ethnicity in the Great Lakes, Rocky Mountains, and Southwest BEA regions	790	98.4	1.03
Graduate students with More than one race, Black, American Indian, or Alaska Native race/ethnicity in the Mideast or Southeast BEA regions	1,940	93.8	1.09
Graduate students with More than one race, Black, American Indian, or Alaska Native race/ethnicity in the New England or Plains BEA regions	530	90.2	1.14

See notes at end of table.

Table 40. Weight adjustment factors for graduate student nonresponse adjustment, by model predictor variable: 2017–18—Continued

Model predictor variables	Number of respondents	Weighted response rate	Average weight adjustment factor (adj5)
Graduate students with Hispanic race/ethnicity who did not receive institution aid	1,310	84.3	1.22
Graduate students with Hispanic race/ethnicity who received institution aid	590	100.0	1.01

[1] The weighted response rates were calculated using the student base weight adjusted for multiplicity and unknown eligibility.
[2] 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, Louisiana, 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.
[3] Enrollment and Direct Loan categories were defined by quartiles.
NOTE: BEA = Bureau of Economic Analysis; CHAID = chi-square automatic interaction detection. 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, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

The resulting nonresponse adjustment factor, denoted adj5, was multiplied by the student base weight and adjusted for multiplicity and unknown eligibility (WT7) to form the nonresponse-adjusted student weight (WT8), WT8 = WT7 * adj5.

Student poststratification adjustment (adj6). To ensure that the weighted student sample adequately represents the student population, the student weights were further adjusted using SUDAAN WTADJUST so that they would sum to known population totals for key characteristics. This adjustment also helped increase the precision of the estimates for these key characteristics and any related characteristics.

Control totals were obtained from FSA data files, the National Association of State Student Grant & Aid Programs (NASSGAP), and IPEDS data.

For undergraduate students, control totals included state-level totals so that state-level totals added up to national totals. The state-level control totals were further split into totals by public 2-year, public 4-year, and other institutions within state when available. Control totals that were unavailable at the state level but were available at the national level were included in the model.

Control totals for undergraduate students were established for the following:

- full-year undergraduate student enrollment, by race/ethnicity; by control and level of institution overall and by state; and by public 2-year, public 4-year, and other institutions within state;
- fall undergraduate student enrollment, by control and level of institution overall and by state and by public 2-year, public 4-year, and other institutions within state;

- number of Pell Grant undergraduate recipients, by control and level of institution overall and by state and by public 2-year, public 4-year, and other institutions within state;
- total amount of Pell Grants awarded to undergraduate students, by control and level of institution overall and by state and by public 2-year, public 4-year, and other institutions within state;
- total amount of PLUS Loan recipients, by control and level of institution overall and by state and by public 2-year, public 4-year, and other institutions within state;
- number of Direct Loan undergraduate student recipients, by subsidized/unsubsidized loan type; by control and level of institution overall and by state; and by public 2-year, public 4-year, and other institutions within state;
- total amount of Direct Loans disbursed to undergraduate students, by subsidized/unsubsidized loan type; by control and level of institution overall and by state; and by public 2-year, public 4-year, and other institutions within state;
- number of state aid undergraduate recipients by aid type, by control and level of institution overall and by state, and by public 2-year and public 4-year institutions within state when available;
- total amount of state aid awarded to undergraduate students by aid type; by control and level of institution overall and by state; and by public 2-year, public 4-year, and other institutions within state;
- total amount of Federal Supplemental Educational Opportunity Grants (SEOGs) disbursed to undergraduate students, by control and level of institution; and
- number of undergraduate students with income equal to $0, by dependency status.

Control totals for graduate students were established for the following:

- fall graduate student enrollment, by control and level of institution;
- full-year graduate student enrollment, by race/ethnicity and by control and level of institution;

- number of unsubsidized Direct Loan recipients among graduate students, by control and level of institution;
- total amount of unsubsidized Direct Loans disbursed to graduate students, by control and level of institution;
- PLUS Loan amounts disbursed to graduate students, by control and level of institution; and
- number of graduate students with income equal to $0, by control and level of institution.

Direct Loan, Pell Grant, PLUS, and SEOG control totals were obtained from FSA. 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 control and level of institution and student level (undergraduate or graduate), and, for undergraduate students, by loan type (subsidized or unsubsidized); state; and public 2-year, public 4-year, and other institutions within state was crucial for poststratification. Before 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.

NASSGAP helped obtain accurate totals of state grant amounts, which were used as control totals.

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

- EF2017A: 2017 Fall Enrollment—includes data on race/ethnicity, sex, attendance status, and level of student for fall 2017;
- EFFY2018: 2017–18 E12—includes 12-month unduplicated head count for 2017–18;
- HD2017: 2017–18 IC-H—includes directory information for 2017–18;

- IC2017: 2017–18 IC—includes data on educational offerings, organization, admissions, services, and athletic associations for 2017–18; and

- IC2017_PY: 2017–18 IC—includes data on student charges by program (vocational programs) for 2017–18.

The HD2017, IC2017, and IC2017_PY files were used in determining which schools were in the NPSAS population of institutions and were also used to create the control and level of institution variable. The EF2017A and EFFY2018 files were used to determine the undergraduate and graduate enrollment totals for fall and the 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 IPEDS counts used for control totals were adjusted for student multiplicity. The counts were modified using the following general formula:

NPSAS control total = (IPEDS enrollment total) * (1 / mean student multiplicity count),

where NPSAS control total, IPEDS enrollment total, and mean student multiplicity count were by undergraduate and graduate student level within each category of control and level of institution.

Additionally, because enrollment counts from IPEDS could include dual-enrolled students (high school students taking courses at postsecondary institutions), the IPEDS counts used for control totals were further adjusted to account for dual-enrolled students who were not eligible for NPSAS. The adjustment of the control totals to account for dual-enrolled students was handled in a similar manner to the multiplicity adjustment.

To determine full-year undergraduate and graduate student enrollment by race/ethnicity, within control and level of institution, the following formula was used for each race/ethnicity category:

NPSAS race/ethnicity category control total = NPSAS full-year control total[32] * Proportion race/ethnicity category,

where Proportion race/ethnicity category = IPEDS race/ethnicity category full-year enrollment total/IPEDS full-year enrollment total.

[32] NPSAS full-year control total = IPEDS full-year enrollment total adjusted for multiplicity and dual enrollment.

NPSAS race/ethnicity category control total, NPSAS full-year control total, Proportion race/ethnicity category, IPEDS race/ethnicity category full-year enrollment total, and IPEDS full-year enrollment total were created by undergraduate and graduate student level within the given control and level of institution.

If convergence problems with the poststratification model were encountered, variables were selected to be collapsed or dropped from the model. The control totals for variables dropped from the poststratification model were used as benchmark totals in the quality control of the data.

As with the nonresponse weight adjustments, different bounds on the poststratification weight adjustments, depending on whether the input weight was classified as high extreme,[33] low extreme,[34] or nonextreme, were used to accomplish the poststratification adjustment, truncation, and smoothing in one step. The option in the WTADJUST procedure for weight trimming and smoothing was used, and the data were inspected for extreme weights by checking for outliers by key domains.

Undergraduate poststratification model. The initial undergraduate poststratification model included variables that were defined by three institution strata (public 2-year, public 4-year, and other institutions) within state for a total of 153 institution strata.[35] For states with convergence issues, the variables were first collapsed for undergraduate students enrolled in public 2-year and public 4-year institutions before further collapsing into state overall variables. Potential representativeness at the public 2-year, public 4-year, and state levels was used to inform decisions of which strata to collapse within state and which institution strata to retain in the model.[36] The institution strata used in the final model included

- thirty potentially state-representative states with variables defined within three institution strata: public 2-year, public 4-year, and other institutions within state;
- twelve public 2-year strata within nonrepresentative states, which were determined to be potentially representative at the public 2-year level;

[33] High extreme weights were identified as those greater than the median weight plus three times the interquartile range for weights.
[34] Low extreme weights were identified as those less than the median weight minus three times the interquartile range for weights.
[35] There were no public 2-year schools in Alaska, Delaware, or the District of Columbia.
[36] Final representativeness at the public 2-year, public 4-year, and state levels is discussed in section 5.4.

- eighteen public 4-year strata within nonrepresentative states, which were determined to be potentially representative at the public 4-year level; and
- thirty-five nonrepresentative institution strata that were collapsed into an "all other sectors" stratum.

Bounds for the undergraduate poststratification were determined at the institution strata level. Before the undergraduate student poststratification adjustment, bounds were set to trim the extreme weights going into the adjustment. To trim small weights, lower bounds were set as needed within institution strata by state. The lower bounds for the 2,856 low extreme weights that were trimmed ranged from 1 to 32.9, with an average lower bound of 5.5. The upper bounds for the 1,755 high extreme weights that were trimmed ranged from 7.4 to 5,000, with an average upper bound of 191.5.

The outcome of the WTADJUST procedure was the poststratification adjustment factor (adj6). For the undergraduate student poststratification weight adjustment, lower bounds and upper bounds at the institution strata level were set as needed on the weight adjustment factors coming out of the poststratification weight adjustment in order to control the magnitude of the poststratified weights. The lower bounds ranged from 0.000001 to 0.9, with an average lower bound of 0.13. The upper bounds ranged from 1.1 to 933.9, with an average upper bound of 23.3. For control and level of institution categories for which no bounds were specified, the lower and upper bound on the adjustment were unbounded.

Table 41 shows state-level control totals and the average poststratification weight adjustment factors by the institution strata used in the final model for undergraduate students.

Table 41. Weight adjustment factors for student poststratification for undergraduate students, by state and institution stratum: 2017–18

State and institution stratum[1]	Control total[2]	Average weight adjustment factor (adj6)
Alabama		
Public 2-year	90,780	1.20
Public 4-year	121,260	1.02
Other	54,640	1.16
Alaska		
Public 4-year	34,360	1.09
Other	2,680	4.79
Arizona		
Public 2-year	193,170	1.08
Public 4-year	100,220	0.92
Other	209,260	1.74
Arkansas		
Public 2-year	39,310	0.96
Public 4-year	69,600	1.07
Other	17,460	1.57
California		
Public 2-year	943,150	0.72
Public 4-year	697,150	0.79
Other	381,250	1.02
Colorado		
Public 2-year	52,540	0.84
Public 4-year	172,360	1.01
Other	87,250	1.80
Connecticut		
Public 2-year	42,360	0.76
District of Columbia		
Public 4-year	6,890	1.48
Other	46,700	0.91
Florida		
Public 2-year	67,650	1.28
Public 4-year	743,820	1.04
Other	295,470	1.20
Georgia		
Public 4-year	262,270	0.92
Hawaii		
Public 2-year	26,390	1.06
Public 4-year	21,300	0.98
Other	11,130	1.34
Idaho		
Public 2-year	23,050	1.29
Public 4-year	37,010	1.02
Illinois		
Public 2-year	378,190	1.22
Public 4-year	118,730	0.95
Other	205,200	1.05
Indiana		
Public 2-year	118,220	0.98
Public 4-year	147,960	0.88
Other	76,170	2.02

See notes at end of table.

Table 41. Weight adjustment factors for student poststratification for undergraduate students, by state and institution stratum: 2017–18—Continued

State and institution stratum	Control total[1]	Average weight adjustment factor (adj6)
Iowa		
Public 2-year	66,260	0.98
Public 4-year	55,340	1.10
Other	83,200	1.15
Kansas		
Public 2-year	84,250	1.07
Public 4-year	65,820	0.93
Kentucky		
Public 2-year	80,220	1.30
Public 4-year	86,710	0.95
Other	38,920	1.26
Louisiana		
Public 2-year	58,230	0.90
Public 4-year	98,830	1.09
Other	38,860	1.05
Maine		
Public 2-year	16,620	0.84
Maryland		
Public 4-year	149,990	1.11
Massachusetts		
Public 2-year	103,920	1.17
Public 4-year	98,000	1.01
Other	188,970	1.01
Michigan		
Public 2-year	171,040	1.12
Public 4-year	235,770	0.92
Other	66,810	1.01
Minnesota		
Public 2-year	109,380	0.92
Public 4-year	92,130	1.02
Other	83,440	0.97
Mississippi		
Public 2-year	77,730	1.12
Public 4-year	59,760	1.06
Missouri		
Public 2-year	68,740	0.75
Public 4-year	112,750	1.06
Other	102,850	1.10
Montana		
Public 4-year	35,650	1.12
Nebraska		
Public 2-year	40,340	0.98
Public 4-year	41,990	0.93
Nevada		
Public 2-year	11,810	0.93
Public 4-year	97,660	1.29
New Hampshire		
Public 4-year	22,960	0.76

See notes at end of table.

Table 41. Weight adjustment factors for student poststratification for undergraduate students, by state and institution stratum: 2017–18—Continued

State and institution stratum	Control total[1]	Average weight adjustment factor (adj6)
New Jersey		
Public 2-year	152,890	0.97
Public 4-year	156,000	1.06
Other	79,420	1.06
New Mexico		
Public 2-year	65,240	1.26
Public 4-year	38,160	0.99
New York		
Public 2-year	302,030	1.09
Public 4-year	357,640	1.09
Other	378,160	1.30
North Carolina		
Public 2-year	104,070	0.59
Public 4-year	161,110	0.90
Other	70,130	0.91
North Dakota		
Public 4-year	34,030	1.19
Ohio		
Public 2-year	177,000	1.04
Public 4-year	214,240	0.90
Other	140,860	0.97
Oklahoma		
Public 2-year	71,380	1.17
Public 4-year	84,740	0.91
Oregon		
Public 2-year	113,130	1.20
Public 4-year	80,110	1.21
Pennsylvania		
Public 2-year	135,670	1.00
Public 4-year	180,050	0.92
Other	257,090	1.06
Rhode Island		
Public 4-year	18,980	0.87
South Carolina		
Public 4-year	81,150	0.96
South Dakota		
Public 2-year	4,970	0.73
Public 4-year	32,890	1.09
Tennessee		
Public 2-year	109,360	1.03
Public 4-year	97,950	0.97
Other	80,290	1.00
Texas		
Public 2-year	549,320	1.33
Public 4-year	450,230	0.84
Other	176,240	1.21
Utah		
Public 2-year	38,900	1.19
Public 4-year	131,710	1.19
Other	150,250	1.60

See notes at end of table.

Table 41. Weight adjustment factors for student poststratification for undergraduate students, by state and institution stratum: 2017–18—Continued

State and institution stratum	Control total[1]	Average weight adjustment factor (adj6)
Vermont		
Public 2-year	7,120	1.11
Public 4-year	16,560	1.52
Other	15,590	1.05
Virginia		
Public 2-year	167,380	1.07
Public 4-year	163,650	0.95
Other	131,210	1.10
Washington		
Public 2-year	6,000	0.18
Public 4-year	298,450	1.21
Other	43,630	1.44
West Virginia		
Public 4-year	55,150	1.16
Wisconsin		
Public 2-year	80,790	1.01
Public 4-year	163,970	1.13
Other	45,640	1.16
Wyoming		
Public 2-year	18,460	1.44
Public 4-year	8,880	0.91
All other institutions[3]	1,449,910	1.31

[1] There were no public 2-year institutions in Alaska, Delaware, or the District of Columbia.
[2] Enrollment control totals are the sum of full-year enrollment across institutions based on 2017–18 Integrated Postsecondary Education Data System enrollment data adjusted for student multiplicity and dual enrollment.
[3] All other institutions include schools for which the model did not include state-level or sector-level control totals: public 2-year, public 4-year, and other institutions in Puerto Rico; public 2-year and other institutions in Georgia, Maryland, Montana, New Hampshire, North Dakota, Rhode Island, South Carolina, and West Virginia; public 4-year and other institutions in Connecticut, Delaware, and Maine; and other institutions in Idaho, Kansas, Mississippi, Nebraska, Nevada, New Mexico, Oklahoma, Oregon, South Dakota, and Wyoming.
NOTE: Sample sizes rounded to the nearest 10. For the purposes of this table, Puerto Rico and the District of Columbia are referred to as "states."
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

There were 1,657 control totals in the final undergraduate poststratification model, which included control totals by public 2-year, public 4-year, and other institutions within state as well as state-level and national-level control totals. For the sake of brevity, table 42 shows only the national-level control totals that were included in the model as well as the average weight adjustment factors resulting from the undergraduate poststratification student model.

Summary statistics of the undergraduate poststratification student weight adjustment factors follow:

- minimum: 0.005;
- median: 0.85; and
- maximum: 1,119.36.

Table 42. Poststratification weight adjustment factors for undergraduate students, by poststratification category: 2017–18

Poststratification category[1]	Control total[2]	Average weight adjustment factor (adj6)
Total full-year undergraduate student enrollment	**16,587,650**	**1.10**
Full-year undergraduate student enrollment by race/ethnicity within control and level of institution		
White non-Hispanic		
Public less-than-2-year	38,250	1.32
Public 2-year	2,590,050	1.00
Public 4-year	3,526,300	0.99
Private nonprofit less-than-4-year	29,590	1.21
Private nonprofit 4-year, non-doctorate-granting	626,240	1.02
Private nonprofit 4-year, doctorate-granting	876,000	0.95
Private for-profit less-than-2-year	93,530	2.11
Private for-profit 2-year	110,350	1.92
Private for-profit 4-year	332,730	1.10
Black non-Hispanic		
Public less-than-2-year	6,810	1.20
Public 2-year	730,460	1.06
Public 4-year	765,960	1.01
Private nonprofit less-than-4-year	24,350	5.57
Private nonprofit 4-year, non-doctorate-granting	145,600	0.89
Private nonprofit 4-year, doctorate-granting	167,130	0.97
Private for-profit less-than-2-year	72,350	2.14
Private for-profit 2-year	88,180	3.93
Private for-profit 4-year	223,480	1.21
Hispanic		
Public less-than-2-year	8,620	1.53
Public 2-year	1,176,260	0.92
Public 4-year	1,202,120	0.92
Private nonprofit less-than-4-year	12,550	0.38
Private nonprofit 4-year, non-doctorate-granting	140,530	0.69
Private nonprofit 4-year, doctorate-granting	227,820	0.76
Private for-profit less-than-2-year	80,640	2.20
Private for-profit 2-year	91,410	5.15
Private for-profit 4-year	141,570	0.83
Other		
Public less-than-2-year	9,100	3.60
Public 2-year	847,480	1.32
Public 4-year	1,292,570	1.22
Private nonprofit less-than-4-year	15,630	12.97
Private nonprofit 4-year, non-doctorate-granting	223,280	2.21
Private nonprofit 4-year, doctorate-granting	386,140	1.51
Private for-profit less-than-2-year	35,470	3.31
Private for-profit 2-year	46,240	9.71
Private for-profit 4-year	202,890	4.01
Fall undergraduate student enrollment by control and level of institution		
Public less-than-2-year	40,860	1.34
Public 2-year	3,908,600	0.90
Public 4-year	5,823,820	0.95
Private nonprofit less-than-4-year	52,830	2.11
Private nonprofit 4-year, non-doctorate-granting	941,640	1.04
Private nonprofit 4-year, doctorate-granting	1,400,320	0.97

See notes at end of table.

Table 42. Poststratification weight adjustment factors for undergraduate students, by poststratification category: 2017–18—Continued

Poststratification category[1]	Control total[2]	Average weight adjustment factor (adj6)
Fall undergraduate student enrollment by control and level of institution—Continued		
Private for-profit less-than-2-year	160,280	1.97
Private for-profit 2-year	205,660	2.88
Private for-profit 4-year	549,110	1.08
Number of Pell Grant undergraduate recipients by control and level of institution		
Public less-than-2-year	19,050	1.67
Public 2-year	2,243,240	1.15
Public 4-year	2,838,990	1.25
Private nonprofit less-than-4-year	56,830	3.62
Private nonprofit 4-year, non-doctorate-granting	498,650	1.13
Private nonprofit 4-year, doctorate-granting	592,630	1.26
Private for-profit less-than-2-year	227,710	3.09
Private for-profit 2-year	246,930	5.17
Private for-profit 4-year	543,760	1.59
Total amount of Pell Grants awarded to undergraduate students by control and level of institution		
Public less-than-2-year	$74,739,480	1.67
Public 2-year	7,772,368,980	1.15
Public 4-year	12,013,592,570	1.25
Private nonprofit less-than-4-year	256,074,460	3.62
Private nonprofit 4-year, non-doctorate-granting	2,107,541,490	1.13
Private nonprofit 4-year, doctorate-granting	2,445,996,300	1.26
Private for-profit less-than-2-year	840,053,910	3.09
Private for-profit 2-year	1,003,058,310	5.17
Private for-profit 4-year	2,106,407,110	1.59
Total amount of PLUS Loans disbursed to undergraduate students by control and level of institution		
Public less-than-2-year	$2,646,430	22.01
Public 2-year	82,867,020	0.45
Public 4-year	6,495,056,920	1.24
Private nonprofit less-than-4-year	32,949,220	4.27
Private nonprofit 4-year, non-doctorate-granting	1,887,284,020	1.12
Private nonprofit 4-year, doctorate-granting	3,365,947,210	1.21
Private for-profit less-than-2-year	129,956,130	2.72
Private for-profit 2-year	212,720,420	5.99
Private for-profit 4-year	401,296,320	2.23
Number of subsidized Direct Loan undergraduate student recipients by control and level of institution		
Public less-than-2-year	8,270	5.50
Public 2-year	720,270	1.01
Public 4-year	2,527,440	1.24
Private nonprofit less-than-4-year	47,770	3.52
Private nonprofit 4-year, non-doctorate-granting	574,500	1.11
Private nonprofit 4-year, doctorate-granting	789,240	1.25
Private for-profit less-than-2-year	164,540	3.05
Private for-profit 2-year	191,140	4.57
Private for-profit 4-year	524,380	1.69

See notes at end of table.

Table 42. Poststratification weight adjustment factors for undergraduate students, by poststratification category: 2017–18—Continued

Poststratification category[1]	Control total[2]	Average weight adjustment factor (adj6)
Total amount of subsidized Direct Loans disbursed to undergraduate students by control and level of institution		
Public less-than-2-year	$25,290,570	5.50
Public 2-year	2,084,569,340	1.01
Public 4-year	9,979,601,190	1.24
Private nonprofit less-than-4-year	170,560,010	3.52
Private nonprofit 4-year, non-doctorate-granting	2,286,451,160	1.11
Private nonprofit 4-year, doctorate-granting	3,217,088,740	1.25
Private for-profit less-than-2-year	474,732,650	3.05
Private for-profit 2-year	658,712,150	4.57
Private for-profit 4-year	1,972,327,210	1.69
Number of unsubsidized Direct Loan undergraduate student recipients by control and level of institution		
Public less-than-2-year	8,960	6.32
Public 2-year	621,340	1.01
Public 4-year	2,470,260	1.25
Private nonprofit less-than-4-year	47,560	3.87
Private nonprofit 4-year, non-doctorate-granting	586,620	1.11
Private nonprofit 4-year, doctorate-granting	792,970	1.27
Private for-profit less-than-2-year	169,170	3.16
Private for-profit 2-year	191,900	4.66
Private for-profit 4-year	532,000	1.77
Total amount of unsubsidized Direct Loans disbursed to undergraduate students by control and level of institution		
Public less-than-2-year	$41,268,420	6.32
Public 2-year	2,200,812,630	1.01
Public 4-year	9,587,803,230	1.25
Private nonprofit less-than-4-year	203,242,760	3.87
Private nonprofit 4-year, non-doctorate-granting	2,265,284,600	1.11
Private nonprofit 4-year, doctorate-granting	3,037,264,970	1.27
Private for-profit less-than-2-year	679,106,730	3.16
Private for-profit 2-year	879,437,380	4.66
Private for-profit 4-year	2,668,931,110	1.77
Total amount of Federal Supplemental Educational Opportunity Grants disbursed to undergraduate students		
Public less-than-2-year and 2-year	$160,370,750	1.25
Public 4-year	361,924,410	1.17
Private nonprofit	364,515,460	1.18
Private for-profit	106,015,940	2.35
Number of undergraduate students with zero income by dependency status		
Dependent	400,150	1.41
Independent	712,820	1.72

[1] There were 1,657 control totals in the final undergraduate poststratification model. For the sake of brevity, the table shows only the national-level control totals that were included in the model.
[2] Enrollment control totals are the sum of enrollment across institutions based on 2017–18 Integrated Postsecondary Education Data System enrollment data adjusted for student multiplicity and dual enrollment.
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, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

Graduate poststratification model. Before the graduate student poststratification adjustment, bounds were set to trim the extreme weights going into the adjustment. To trim small weights, lower bounds were set within each category of control and level of institution; the lower bounds ranged from 1 for students in public 4-year, doctorate-granting institutions to 8.6 for students in private for-profit 4-year institutions. To trim large weights, an upper bound of 20 was set for students in public 4-year, non-doctorate-granting, primarily subbaccalaureate institutions and was set equal to the median incoming weight plus three times the interquartile range by control and level of institution for all other students.

The outcome from the WTADJUST procedure was the poststratification adjustment factor (adj6). For the graduate student poststratification weight adjustment, lower bounds and upper bounds on the weight adjustment factors coming out of the poststratification weight adjustment were set in order to control the magnitude of the poststratified weights. The lower bounds ranged from 0.002 to 0.700, with a median lower bound of 0.002. The upper bounds were set within certain categories of control and level of institution as needed; these upper bounds ranged from 1.00 to 4.00. For all other control and level of institution categories, the upper bound on the adjustment was unbounded.

Table 43 shows the control totals associated with the poststratification weight adjustment and the average weight adjustment factors resulting from the graduate student model. Summary statistics of the graduate poststratification student weight adjustment factors follow:

- minimum: 0.005;
- median: 1.08; and
- maximum: 41.84.

Table 43. Poststratification weight adjustment factors for graduate students, by poststratification category: 2017–18

Poststratification category	Control total[1]	Average weight adjustment factor (adj6)
Total full-year graduate student enrollment	**3,550,840**	**1.24**
Full-year graduate student enrollment, by race, within control and level of institution		
White		
Public		
4-year, non-doctorate-granting, primarily subbaccalaureate	630	1.01
4-year, non-doctorate-granting, primarily baccalaureate	89,370	1.22
4-year, doctorate-granting	802,150	1.16
Private nonprofit		
4-year, non-doctorate-granting	129,560	1.02
4-year, doctorate-granting	634,610	1.05
Private for-profit 4-year	133,430	1.32
Black		
Public		
4-year, non-doctorate-granting, primarily subbaccalaureate	510	6.71
4-year, non-doctorate-granting, primarily baccalaureate	15,090	1.00
4-year, doctorate-granting	133,160	0.96
Private nonprofit		
4-year, non-doctorate-granting	24,910	1.06
4-year, doctorate-granting	135,590	1.10
Private for-profit 4-year	98,760	1.22
Other race		
Public		
4-year, non-doctorate-granting, primarily subbaccalaureate	240	0.85
4-year, non-doctorate-granting, primarily baccalaureate	42,820	1.24
4-year, doctorate-granting	565,560	1.31
Private nonprofit		
4-year, non-doctorate-granting	76,430	1.68
4-year, doctorate-granting	547,770	1.28
Private for-profit 4-year	120,280	2.21
Fall graduate student enrollment by control and level of institution		
Public		
4-year, non-doctorate-granting, primarily subbaccalaureate	890	1.44
4-year, non-doctorate-granting, primarily baccalaureate	96,910	1.03
4-year, doctorate-granting	1,235,210	1.15
Private nonprofit		
4-year, non-doctorate-granting	157,640	1.06
4-year, doctorate-granting	1,047,550	1.06
Private for-profit 4-year	222,520	1.37
Number of unsubsidized Direct Loan graduate student recipients by control and level of institution		
Public		
4-year, non-doctorate-granting, primarily subbaccalaureate	710	3.63
4-year, non-doctorate-granting, primarily baccalaureate	47,280	1.29
4-year, doctorate-granting	523,680	1.28
Private nonprofit		
4-year, non-doctorate-granting	94,640	1.20
4-year, doctorate-granting	556,240	1.30
Private for-profit 4-year	213,030	1.83

See notes at end of table.

Table 43. Poststratification weight adjustment factors for graduate students, by poststratification category: 2017–18—Continued

Poststratification category	Control total[1]	Average weight adjustment factor (adj6)
Total amount of unsubsidized Direct Loans disbursed to graduate students by control and level of institution		
Public		
4-year, non-doctorate-granting, primarily subbaccalaureate	$9,525,320	3.63
4-year, non-doctorate-granting, primarily baccalaureate	629,660,900	1.29
4-year, doctorate-granting	9,760,377,850	1.28
Private nonprofit		
4-year, non-doctorate-granting	1,312,047,240	1.20
4-year, doctorate-granting	11,509,335,950	1.30
Private for-profit 4-year	3,210,943,570	1.83
Total amount of Graduate PLUS Loans disbursed to graduate students by control and level of institution		
Public		
4-year, non-doctorate-granting, primarily baccalaureate	$23,847,440	2.17
4-year, doctorate-granting	2,447,620,950	1.27
Private nonprofit		
4-year, non-doctorate-granting	210,070,210	1.44
4-year, doctorate-granting	6,170,824,280	1.26
Private for-profit 4-year	553,722,590	2.08
Number of graduate students with income equal to $0 by control and level of institution		
Public		
4-year, non-doctorate-granting, primarily subbaccalaureate	140	0.62
4-year, non-doctorate-granting, primarily baccalaureate	7,870	1.68
4-year, doctorate-granting	174,130	2.93
Private nonprofit		
4-year, non-doctorate-granting	18,190	2.16
4-year, doctorate-granting	158,020	2.25
Private for-profit 4-year	23,840	2.99

[1] Enrollment control totals are the sum of enrollment across institutions based on 2017–18 Integrated Postsecondary Education Data System enrollment data adjusted for student multiplicity.
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, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

The resulting poststratification adjustment, denoted adj6, was used to calculate the final analysis weight described in the following section.

Final analysis weight (WTA000). After poststratification was performed, the final analysis weight, denoted WTA000, was computed as the product of the eight weight components described in this section.

Final analysis weight = WTA000 = WT1 * adj1 * adj2 * WT4 * adj3 * adj4 * adj5 * adj6.

All weight components, along with the final analysis weight, are included on the NPSAS:18-AC weights history file found in the NPSAS:18-AC RUF.

Weighted estimates were compared for key variables from the NPSAS:18-AC data with estimates from other sources—such as estimates from NPSAS:16, FSA, and VBA. NPSAS:18-AC estimates were found to be reasonable considering differences in time frame, population, and other factors that would explain differences.

6.2 Nonresponse Bias Analysis

The accuracy of survey statistics is affected by both random and nonrandom errors. Random errors reduce the precision of survey estimates; however, precision can be improved with increased sample size. For example, all sample surveys are subject to random error in the form of sampling error (error that occurs because the survey is based on a sample of population members rather than the population). 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). Although NPSAS:18-AC does not include survey data, nonrandom errors can still result in biased estimates. The NPSAS:18-AC data collection procedures were thoroughly developed and tested to minimize nonrandom errors because these errors are difficult to quantify.

Errors due to nonresponse are often nonrandom and may result in bias; hence, when nonresponse rates are high, an assessment of the potential for study results to be biased due to nonresponse is required under NCES statistical standards. Standard 4-4-1 states, "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: $B(\bar{y}_R) = E(\bar{y}_R) - \pi$. The population mean π can be estimated for characteristics that are observed for both respondents and nonrespondents by 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 in the sampling frame, π can be estimated without sampling error; hence, bias can be estimated as the difference between the respondent mean and the 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 (*RB*) 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 will 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.

Sample studies typically encounter two general types of nonresponse: *unit nonresponse* and *item nonresponse*. For NPSAS:18-AC, there are two sampling units: institutions and students. *Unit nonresponse* occurs when a sampled institution or student does not meet the respondent criteria defined in sections 6.1.1 and 6.1.2. *Item nonresponse* occurs when a student or institution respondent is missing values for one or more items.

The procedures for unit and item nonresponse bias analyses are described in the sections below. The institution-level results are summarized in tables 44 through 47, while student-level results are summarized in tables 48 through 51, and detailed tables of results are provided in appendix K.

6.2.1 Bias Analysis: Institution Level

Institution-level nonresponse bias analyses were conducted overall and within each control and level of institution separately for undergraduate- and graduate-enrolling institutions. Data for the analyses came from the 2017–18 IPEDS files, which contain extensive data for all institutions in the sample. Details of the IPEDS data files used are described in the student poststratification adjustment section in the student weight section (section 6.1.2).

The variables for the institution nonresponse bias analysis included all variables used in the institution nonresponse adjustment listed in section 6.1.1 and additional variables not used in the weight adjustment model. The following variables were used for the institution-level analyses:[37]

- control and level of institution;

[37] Continuous variables were converted to categorical variables using quartiles. Variables related to student population reflect the population of undergraduate students for analyses of undergraduate-enrolling institutions and of graduate students for analyses of graduate-enrolling institutions.

- 2015 Carnegie Basic Classification;
- degree of urbanization (city/suburb/rural);
- region of institution;
- HBCU status (yes/no);
- HSI status (yes/no);[38]
- percentage of undergraduate or graduate students enrolled who were Black, non-Hispanic;
- percentage of undergraduate or graduate students enrolled who were Asian or Pacific Islander, non-Hispanic (categorical);
- percentage of undergraduate or graduate students enrolled who were Hispanic;
- total undergraduate or graduate student enrollment;
- total male undergraduate or graduate student enrollment; and
- total female undergraduate or graduate student enrollment.

The following variables were used only for undergraduate-enrolling institutions:

- state, control, and level of institution;
- percentage of undergraduate students receiving federal grant aid;
- percentage of undergraduate students receiving state/local grant aid;
- percentage of undergraduate students receiving institution grant aid;
- percentage of undergraduate students receiving student loan aid;
- average net price among students receiving grant or scholarship aid;
- percentage of full-time, first-time degree/certificate-seeking undergraduate students who received any grant aid;
- number of full-time, first-time undergraduate students living on campus;

[38] An HIS indicator is no longer available from IPEDS, so an HSI proxy was created following the definition of HSI provided by the U.S. Department of Education (https://www2.ed.gov/programs/idueshsi/definition.html) and using IPEDS Hispanic enrollment data.

- number of full-time, first-time undergraduate students receiving Title IV aid with incomes up to $30,000; and
- average amount of grant and scholarship aid received.

The following variable was used only for graduate-enrolling institutions:

- total office and administrative support employees.

First, nonresponse bias was estimated for each category of each of these variables 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 differed significantly from zero at the 5 percent significance level. In order to evaluate the efficacy of the nonresponse weight adjustment (section 6.1.1), nonresponse bias was reestimated using the adjusted weight and tested for statistical significance. Significance tests were complemented by relative bias and 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 weighted respondent 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. All analyses were carried out using SUDAAN's DESCRIPT and VARGEN procedures (RTI International 2012).

Nonresponse bias analysis for undergraduate-enrolling institutions. The weighted response rate for undergraduate-enrolling institutions was 65 percent and ranged from 13 percent for public less-than-2-year institutions to 90 percent for private nonprofit less-than-4-year institutions (table 29). As shown in table 44, for undergraduate-enrolling institutions overall, the institution nonresponse weighting adjustment reduced but did not eliminate statistically significant bias on the observable characteristics that met reporting requirements. The bias reduction results across control and level of institution were mixed as, unlike previous iterations of NPSAS, the nonresponse adjustment model did not include all levels of control and level of institution. As described in section 6.1.1, the nonresponse adjustment model for undergraduate-enrolling institutions includes public 2-year, public 4-year, and other institutions within the state where the institution was located. Further, as described for graduate-enrolling institutions, institution-level adjustments were conducted with the primary goal of reducing nonresponse bias at the student level.

Table 44. Summary of institution-level nonresponse bias analysis, by control and level of institution, for undergraduate-enrolling institutions: 2017–18

Nonresponse bias statistic[1]	Overall	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
Before nonresponse weight adjustment[2]												
Mean percent relative bias across characteristics	15.98	115.68	5.25	5.49	7.50	3.11	29.43	16.39	19.83	30.53	33.56	30.09
Median percent relative bias across characteristics	12.81	70.73	3.96	3.72	5.49	2.27	22.42	11.59	19.94	22.67	28.06	29.54
Percentage of characteristics with significant bias	22.78	29.73	23.26	6.90	11.54	8.11	#	14.77	7.89	6.49	17.72	16.05
Median effect size	0.12	‡	0.04	0.04	0.09	0.04	‡	0.20	0.20	0.20	0.37	0.42
After nonresponse weight adjustment[3]												
Mean percent relative bias across characteristics	27.97	56.22	4.74	8.55	7.59	3.62	31.73	17.58	14.89	28.62	53.54	40.45
Median percent relative bias across characteristics	9.88	47.00	2.83	8.26	5.70	3.31	25.81	10.48	10.64	19.41	39.51	40.12
Percentage of characteristics with significant bias	8.89	5.41	9.30	8.62	7.69	4.05	11.11	3.41	#	6.49	30.38	23.46
Median effect size	0.08	‡	0.04	0.08	0.10	0.03	‡	0.17	0.09	0.22	0.55	0.53

Rounds to zero.
‡ Reporting standards not met.
[1] Relative bias and effect size were calculated using the weighted differences between respondent and full-sample means. Relative bias was calculated as 100 times the ratio of estimated bias to the weighted full-sample mean. Effect size was 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: "Institution base weight" refers to the institution sampling weight. Variables and characteristics that did not meet reporting standards were excluded from calculation of summary statistics.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

Before weighting, the median percent relative bias across characteristics was 12.81 for institutions overall and ranged from 2.27 to 70.73 for institutions by control and level. After weighting, the median percent relative bias was 9.88 for institutions overall and ranged from 2.83 to 47.00 across control and level of institution. The percentage of characteristics with significant bias before weighting was 22.78 for institutions overall and ranged from zero to 29.73 across control and level of institution. After weighting, the percentage was 8.89 for institutions overall and between zero and 30.38 across control and level of institution. Median effect size was 0.12 for institutions overall before weighting and ranged from 0.04 to 0.42 across control and level of institution. After weighting, median effect size was 0.08 for institutions overall and ranged from 0.03 to 0.55 across control and level of institution. For most categories of control and level, there was not substantial evidence of bias either before or after weight adjustment; however, the median effect sizes were moderate to large for private for-profit institutions. Detailed results can be found in appendix tables K-1 through K-12.

As seen in table 45, the mean absolute difference between respondent means before and after poststratification was 0.12 for undergraduate-enrolling institutions overall and ranged from 0.01 to 0.86 across control and level of institution. The median absolute difference was 0.05 for institutions overall and ranged from 0.01 to 0.67 across control and level of institution. The mean absolute difference between the full-sample and respondent means after poststratification was 1.16 for institutions overall and ranged from 0.83 to 15.19 across control and level of institution. The median absolute difference was 0.57 for institutions overall and ranged from 0.79 to 9.70 across control and level of institution. Detailed results can be found in appendix tables K-13 through K-24.

Table 45. Summary of institution-level differences between means before and after poststratification, by control and level of institution, for undergraduate-enrolling institutions: 2017–18

Summary statistic	Overall	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
Difference between means for respondents before and after poststratification adjustment[1]												
Mean absolute difference across characteristics	0.12	0.19	0.86	0.04	0.04	0.01	0.53	0.21	0.43	0.52	0.21	0.54
Median absolute difference across characteristics	0.05	0.25	0.67	0.03	0.03	0.01	0.19	0.14	0.36	0.23	0.14	0.41
Difference between means for full sample and respondents after poststratification adjustment[2]												
Mean absolute difference across characteristics	1.16	15.19	1.02	2.17	1.59	0.83	5.03	2.50	3.49	4.14	10.59	8.97
Median absolute difference across characteristics	0.57	9.70	0.88	2.04	1.21	0.79	5.56	2.18	1.82	3.52	8.98	6.55

[1] Respondent means before poststratification adjustment were weighted using the institution base weight adjusted for nonresponse. Respondent means after poststratification adjustment were weighted using the institution base weight adjusted for nonresponse and poststratification. Institution base weight refers to the institution sampling weight.
[2] Full-sample means were weighted using the institution base weight, and respondent means were 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.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

Nonresponse bias analysis for graduate-enrolling institutions. The weighted response rate for graduate-enrolling institutions was 75 percent and ranged from 59 percent for private for-profit 4-year institutions to 94 percent for public 4-year, non-doctorate-granting, primarily subbaccalaureate institutions (table 29). As shown in table 46, for graduate-enrolling institutions overall, the percentage of characteristics with significant bias was reduced by the nonresponse weight adjustment from 7.81 percent to 1.56 percent. In addition, the summary statistics overall and across many combinations of control and level of institution were larger after weight adjustment. This can be explained in part by small sample sizes. Another factor is that the institution-level nonresponse weight adjustment included student enrollment counts to reduce bias at the student level, as NPSAS is designed for student-level inferences rather than institution-level inferences (see section 6.1.1). However, the institution-level nonresponse bias analyses use the institution weight without accounting for enrollment counts to evaluate the impact of the institution nonresponse weight adjustment on institution-level analyses. Despite this, there was not substantial evidence of bias either before or after weight adjustment, although the median effect size was moderately large for private for-profit 4-year institutions.

Table 46. Summary of institution-level nonresponse bias analysis, by control and level of institution, for graduate-enrolling institutions: 2017–18

Nonresponse bias statistic[1]	Overall	Public 4-year, non-doctorate-granting, primarily subbac-calaureate	Public 4-year, non-doctorate-granting, primarily bacca-laureate	Public 4-year, doctorate-granting	Private nonprofit 4-year, non-doctorate-granting	Private nonprofit 4-year, doctorate-granting	Private for-profit 4-year
Before nonresponse weight adjustment[2]							
Mean percent relative bias across characteristics	8.11	‡	6.10	2.81	9.98	16.89	21.54
Median percent relative bias across characteristics	6.33	‡	4.22	1.75	5.68	15.71	19.45
Percentage of characteristics with significant bias	7.81	‡	7.50	4.65	6.67	8.70	9.76
Median effect size	0.06	‡	0.04	0.02	0.03	0.17	0.16
After nonresponse weight adjustments[3]							
Mean percent relative bias across characteristics	14.46	‡	8.40	3.99	15.09	11.56	38.76
Median percent relative bias across characteristics	7.86	‡	8.12	3.53	12.42	10.95	31.83
Percentage of characteristics with significant bias	1.56	‡	12.50	6.98	#	#	26.83
Median effect size	0.09	‡	0.08	0.03	0.22	0.11	0.44

Rounds to zero.
‡ Reporting standards not met (fewer than five characteristics had at least five nonrespondents).
[1] Relative bias and effect size were calculated using the weighted differences between respondent and full-sample means. Relative bias was calculated as 100 times the ratio of estimated bias to the weighted full-sample mean. Effect size was calculated as the square root of the sum over categories of the squared differences over full-sample means.
[2] Respondent and full-sample means were weighted using the institution base weight. Institution base weight refers to the institution sampling weight.
[3] Full-sample means were weighted using the institution base weight, and the respondent means were weighted using the institution base weight adjusted for nonresponse.
NOTE: Variables and characteristics that did not meet reporting standards were excluded from calculation of summary statistics.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

The median percent relative bias before the nonresponse weight adjustment was 6.33 for institutions overall and between 1.75 and 19.45 across control and level of institution. After weight adjustment, the median percent bias was 7.86 for institutions overall and 3.53 to 31.83 across control and level of institution. The percentage of characteristics with significant bias before weighting was 7.81 for institutions overall and between 4.65 and 9.76 across control and level of institution. After weighting, the percentage was 1.56 for institutions overall and between 0.00 and 26.83 across control and level of institution. The median effect size before weighting was 0.06 for institutions overall and between 0.02 and 0.17 across control and level of institution, and after weighting, the median effect size was 0.09 for institutions overall and between 0.03 and 0.44 across control and level of institution. Detailed results can be found in appendix tables K-25 through K-31.

As shown in table 47, the mean absolute difference between respondent means before and after poststratification was 1.99 for graduate-enrolling institutions

overall and ranged from 0.17 to 6.52 across control and level of institution. The median absolute difference was 1.35 for institutions overall and ranged from 0.09 to 4.86 across control and level of institution. The mean absolute difference between the full-sample and respondent means after poststratification was 1.63 for institutions overall and ranged from 0.84 to 7.83 across control and level of institution. The median absolute difference was 1.19 for institutions overall and ranged from 0.72 to 7.10 across control and level of institution. Detailed results can be found in appendix tables K-32 through K-38.

Table 47. Summary of institution-level differences between means before and after poststratification, by control and level of institution, for graduate-enrolling institutions: 2017–18

Summary statistic	Overall	Public 4-year, non-doctorate-granting, primarily subbaccalaureate	Public 4-year, non-doctorate-granting, primarily baccalaureate	Public 4-year, doctorate-granting	Private nonprofit 4-year, non-doctorate-granting	Private nonprofit 4-year, doctorate-granting	Private for-profit 4-year
Difference between means for respondents before and after poststratification adjustment[1]							
Mean absolute difference across characteristics	1.99	‡	0.59	0.17	6.52	0.86	0.57
Median absolute difference across characteristics	1.35	‡	0.45	0.09	4.86	0.61	0.46
Difference between means for full sample and respondents after poststratification adjustment[2]							
Mean absolute difference across characteristics	1.63	‡	1.42	0.84	7.59	2.06	7.83
Median absolute difference across characteristics	1.19	‡	1.25	0.72	4.54	1.61	7.10

‡ Reporting standards not met (fewer than five characteristics had at least five nonrespondents).
[1] Respondent means before poststratification adjustment were weighted using the institution base weight adjusted for nonresponse. Respondent means after poststratification adjustment were weighted using the institution base weight adjusted for nonresponse and poststratification. Institution base weight refers to the institution sampling weight.
[2] Full-sample means were weighted using the institution base weight, and respondent means were 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.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

6.2.2 Bias Analysis: Student Level

The same procedures described in section 6.2.1 for institution-level nonresponse bias analyses were used to conduct the student-level nonresponse bias analyses for the sample overall and by control and level of institution, separately for undergraduate and graduate students. In addition, the analyses were conducted by

state and by public 2-year and public 4-year institution sectors within state for undergraduate students.

The variables used for the student-level nonresponse bias analyses include

- control and level of institution;
- region of institution;
- federal aid status (received/did not receive);
- Direct Loan status (received/did not receive);
- total Direct Loan amount received (categorical);
- age as of December 31, 2017 (categorical);
- institution aid status (received/did not receive);
- state aid status (received/did not receive);
- veteran status (yes/no);
- race (eight levels);
- sex (male/female/unknown); and
- institution total undergraduate or graduate enrollment.

The following variables were used only in analyses of undergraduate students:

- state, control, and level of institution;
- major (13 levels);
- degree program (five levels);
- Pell Grant status (yes/no);
- total Pell Grant amount received (categorical);
- percentage of full-time, first-time degree/certificate-seeking undergraduate students who received any grant aid;
- number of full-time, first-time undergraduate students living on campus;
- number of full-time, first-time undergraduate students receiving Title IV aid with incomes up to $30,000; and

- average amount of grant and scholarship aid received (categorical).

The following variables were used only in analyses of graduate students:

- graduate field of study or major (nine levels);
- degree program (seven levels);
- student type (graduate or first professional); and
- total office and administrative support employees.

Nonresponse bias analysis for undergraduate students. The weighted response rate for undergraduate students was 76 percent and ranged from 34 percent for students in private for-profit 2-year institutions to 83 percent for students in private nonprofit 4-year, doctorate-granting institutions (table 36). As seen in table 48, there was not substantial evidence of significant bias in observable characteristics before the nonresponse weight adjustment, and this was generally reduced after weight adjustment in undergraduate students overall and across control and level of institution. The bias reduction results across level and control were mixed because, unlike previous iterations of NPSAS, the nonresponse adjustment model did not include all categories of control of institution. As described in section 6.1, the nonresponse adjustment for undergraduate students was conducted separately within each state and included, as a model variable, institution level and control categorized by public 2-year, public 4-year, and other institutions. Before weight adjustment, the median percent relative bias across characteristics was 9.74 for undergraduate students overall, ranging from 3.84 to 29.21 across control and level of institution. After weight adjustment, the median was 0.00 for undergraduate students overall and ranged from 2.90 to 25.85 across control and level of institution.

Table 48. Summary of student-level nonresponse bias analysis, by control and level of institution, for undergraduate students: 2017–18

Nonresponse bias statistic[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	14.15	25.68	12.97	10.79	11.24	11.42	20.51	8.04	5.58	20.46	38.02	10.46
Median percent relative bias across characteristics	9.74	19.61	8.61	7.60	8.00	5.70	10.56	4.91	3.84	17.09	29.21	7.01
Percentage of characteristics with significant bias	29.18	7.69	45.00	14.10	34.15	48.24	#	18.29	19.23	6.45	25.76	1.47
Median effect size	0.05	0.12	0.02	0.08	0.09	0.07	0.01	0.03	0.01	0.07	0.31	0.07
After nonresponse weight adjustments[3]												
Mean percent relative bias across characteristics	4.92	30.23	12.31	9.86	9.07	9.00	23.56	10.10	6.52	19.71	34.06	11.69
Median percent relative bias across characteristics	#	14.93	4.25	5.82	4.14	2.90	18.55	7.34	3.88	15.26	25.85	6.63
Percentage of characteristics with significant bias	13.30	5.13	33.75	8.97	15.85	17.65	7.89	7.32	17.95	8.06	18.18	1.47
Median effect size	0.02	0.16	0.02	0.03	0.05	0.03	0.08	0.10	0.03	0.09	0.23	0.04

Rounds to zero.
[1] Relative bias and effect size were calculated using the weighted differences between respondent and full-sample means. Relative bias was calculated as 100 times the ratio of estimated bias to the weighted full-sample mean. Effect size was calculated as the square root of the sum over categories of the squared differences over full-sample means.
[2] Respondent and full-sample means were weighted using the student base weight.
[3] Full-sample means were weighted using the student base weight, and the respondent means were weighted using the student base weight adjusted for nonresponse.
NOTE: Student base weight refers to the student sampling weight (final institution weight times student sampling adjustment) adjusted for student multiplicity and unknown eligibility. Variables and characteristics that did not meet reporting standards were excluded from calculation of summary statistics.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

The percentage of characteristics with significant bias before weight adjustment was 29.18 for undergraduate students overall and between 0.00 and 48.24 across control and level of institution. After weight adjustment the percentage was 13.30 for undergraduate students overall and between 1.47 and 33.75 across control and level of institution. The median effect size for undergraduate students before weight adjustment was 0.05 for students overall and between 0.01 and 0.31 across control and level of institution, while after weight adjustment the median effect size was 0.02 for undergraduate students overall and 0.02 to 0.23 across control and level of institution. Detailed results are provided in appendix tables K-39 through K-50.

As seen in table 49, the mean absolute difference between respondent means before and after poststratification was 0.67 for undergraduate students overall and ranged from 1.27 to 13.28 across control and level of institution. The median absolute difference was 0.10 for undergraduate students overall and ranged from 0.53 to 10.41 across control and level of institution. The mean absolute difference between the full-sample and respondent means after poststratification was 0.84 for undergraduate students overall and ranged from 1.70 to 14.10 across control and level of institution. The median absolute difference was 0.11 for undergraduate students overall and ranged from 1.09 to 10.86 across control and level of institution. Detailed results are provided in appendix tables K-51 through K-62.

Table 49. Summary of student-level differences between means before and after poststratification, by control and level of institution, for undergraduate students: 2017–18

Summary statistic	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
Difference between means for respondents before and after poststratification adjustment[1]												
Mean absolute difference across characteristics	0.67	3.32	1.27	1.83	1.29	1.72	13.28	1.58	1.94	6.80	4.72	3.35
Median absolute difference across characteristics	0.10	2.00	0.53	0.90	0.60	0.65	10.41	0.86	0.83	2.93	2.58	1.95
Difference between means for full sample and respondents after poststratification adjustment[2]												
Mean absolute difference across characteristics	0.84	7.91	2.55	2.80	2.04	1.70	14.10	2.25	2.16	10.64	6.54	4.19
Median absolute difference across characteristics	0.11	2.56	1.11	1.80	1.09	0.74	10.86	1.63	1.14	7.85	6.29	2.41

[1] Respondent means before poststratification adjustment were weighted using the student base weight adjusted for nonresponse. Respondent means after poststratification adjustment were weighted using the student base weight adjusted for nonresponse and poststratification.
[2] Full-sample means were weighted using the student base weight, and respondent means were weighted using the student base weight adjusted for nonresponse and poststratification.
NOTE: Student base weight refers to the student sampling weight (final institution weight times student sampling adjustment) adjusted for student multiplicity and unknown eligibility. Characteristics that did not meet reporting standards were excluded from calculation of summary statistics.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

The nonresponse weight adjustments were applied separately by state; therefore, nonresponse bias analyses procedures were repeated for each state and sampling stratum (public 2-year, public 4-year, and other institutions by state). These results are summarized in appendix table K-77.

At the sampling strata level, the median percent relative bias before weight adjustment ranged from 0.09 to 25.96 across strata, and after weight adjustment, it ranged from 0.00 to 25.73. The percentage of characteristics with significant bias before weight adjustment ranged from zero to 76.19 across strata, and after weight adjustment, it ranged from zero to 65.38. Median effect size ranged from 0.00 to 0.34 across strata before weight adjustment. After weight adjustment, it ranged from 0.00 to 0.19. The mean absolute difference between respondent means before and after poststratification ranged from 0.44 to 20.61 across strata; the median absolute difference ranged from 0.43 to 16.22. The mean absolute difference between respondent and full-sample means after poststratification ranged from 1.10 to 21.28 across strata, and the median absolute difference ranged from 0.09 to 19.28.

Nonresponse bias analysis for graduate students. The weighted response rate for graduate students was 90 percent and ranged from 66 percent for students in public 4-year, non-doctorate-granting, primarily subbaccalaureate institutions to 93 percent for students in private for-profit 4-year institutions (table 36). As seen in table 50, there was not substantial evidence of significant bias in observable characteristics before the nonresponse weight adjustment, and this was generally reduced after the weight adjustment in graduate students overall and across control and level of institution. Before weight adjustment, the median percent relative bias across characteristics was 3.43 for graduate students overall, ranging from 2.07 to 4.68 across control and level of institution. After weight adjustment, the median was 0.00 and ranged from 1.02 to 4.95 across control and level of institution. The percentage of characteristics with significant bias before weight adjustment was 32.31 for graduate students overall and between 0.00 and 26.67 across control and level of institution. After weight adjustment, the percentage was 16.92 for graduate students overall and between 11.63 and 18.60 across control and level of institution. The median effect size before weight adjustment for graduate students overall was 0.02 and between 0.00 and 0.03 across control and level of institution. After weight adjustment, the effect size was 0.00 for graduate students overall and 0.00 to 0.03 across control and level of institution. Detailed results are provided in appendix tables K-63 to K-69.

Table 50. Summary of student-level nonresponse bias analysis, by control and level of institution, for graduate students: 2017–18

Nonresponse bias statistic[1]	Overall	Public 4-year, non-doctorate-granting, primarily subbaccalaureate	Public 4-year, non-doctorate-granting, primarily baccalaureate	Public 4-year, doctorate-granting	Private nonprofit 4-year, non-doctorate-granting	Private nonprofit 4-year, doctorate-granting	Private for-profit 4-year
Before nonresponse weight adjustments[2]	.	.	.	.	.	.	.
Mean percent relative bias across characteristics	8.67	‡	7.16	4.48	8.28	7.84	5.38
Median percent relative bias across characteristics	3.43	‡	2.26	2.19	4.68	3.08	2.07
Percentage of characteristics with significant bias	32.31	‡	9.09	25.58	20.00	26.67	#
Median effect size	0.02	‡	0.03	#	0.01	0.01	0.01
After nonresponse weight adjustments[3]	.	.	.	.	.	.	.
Mean percent relative bias across characteristics	5.84	‡	6.27	3.13	8.07	6.54	6.73
Median percent relative bias across characteristics	#	‡	2.03	1.02	4.95	1.82	3.61
Percentage of characteristics with significant bias	16.92	‡	13.64	18.60	15.00	17.78	11.63
Median effect size	#	‡	0.01	#	0.03	0.01	0.02

Rounds to zero.
‡ Reporting standards not met (fewer than five characteristics had at least 30 nonrespondents).
[1] Relative bias and effect size were calculated using the weighted differences between respondent and full-sample means. Relative bias was calculated as 100 times the ratio of estimated bias to the weighted full-sample mean. Effect size was calculated as the square root of the sum over categories of the squared differences over full-sample means.
[2] Respondent and full-sample means were weighted using the student base weight.
[3] Full-sample means were weighted using the student base weight, and the respondent means were weighted using the student base weight adjusted for nonresponse.
NOTE: Student base weight refers to the student sampling weight (final institution weight times student sampling adjustment) adjusted for student multiplicity and unknown eligibility. Variables and characteristics that did not meet reporting standards were excluded from calculation of summary statistics.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

As seen in table 51, the mean absolute difference between respondent means before and after poststratification was 0.94 for graduate students overall and ranged from 1.23 to 2.93 across control and level of institution. The median absolute difference was 0.52 for graduate students overall and ranged from 0.78 to 2.51 across control and level of institution. The mean absolute difference between the full-sample and respondent means after poststratification was 1.30 for graduate students overall and ranged from 2.38 to 8.84 across control and level of institution. The median absolute difference was 0.67 for graduate students overall and ranged from 2.01 to 5.94 across control and level of institution. Detailed results can be found in appendix tables K-70 through K-76.

Table 51. Summary of student-level differences between means before and after poststratification, by control and level of institution, for graduate students: 2017–18

Summary statistics	Overall	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 4-year, non-doctorate-granting	Private nonprofit 4-year, doctorate-granting	Private for-profit 4-year
Difference between means for respondents before and after poststratification adjustment[1]	.	.	.	.	.	.	.
Mean absolute difference across characteristics	0.94	‡	2.56	1.51	1.23	1.23	2.93
Median absolute difference across characteristics	0.52	‡	2.17	1.03	0.78	1.07	2.51
Difference between means for full sample and respondents after poststratification adjustment[2]	.	.	.	.	.	.	.
Mean absolute difference across characteristics	1.30	‡	3.46	1.71	2.60	1.79	3.31
Median absolute difference across characteristics	0.67	‡	3.13	1.02	1.55	1.10	2.44

‡ Reporting standards not met (fewer than five characteristics had at least 30 nonrespondents).
[1] Respondent means before poststratification adjustment were weighted using the student base weight adjusted for nonresponse. Respondent means after poststratification adjustment were weighted using the student base weight adjusted for nonresponse and poststratification. Student base weight refers to the student sampling weight (final institution weight times student sampling adjustment) adjusted for student multiplicity and unknown eligibility.
[2] Full-sample means were weighted using the student base weight, and respondent means were 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.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

6.2.3 Item Nonresponse Bias Analysis

The analyses of the potential bias due to item nonresponse were conducted for the NPSAS:18-AC sample in accordance with NCES statistical standard 4-4-3A, which 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. 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).

In accordance with NCES statistical standard 1-3-5 (Seastrom 2014), item response rates (RRIs) were calculated as the ratio of the number of respondents for whom in-scope data were obtained (I^x for item x) to the number of eligible respondents, which is calculated as the number of unit respondents (I) minus the number of item respondents with a valid skip for item x (V^x):

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

If a unit respondent's eligibility for an item is unknown it is considered an item nonrespondent. To convert refusals, some institutions were allowed to only provide minimal student records data elements. For these cases, the student record items for which data could not be obtained from a secondary source were treated as item nonresponse. A student was defined to be an item respondent for a variable if the student had data for that variable from any source, including logical imputation.[39] Student final (analysis) weights were used for computing item response rates.

For undergraduate student items (appendix table N-4), 60 out of 113 imputed or select items had response rates greater than 95 percent. The undergraduate student item with the lowest overall response rate was *Years completed or planned social studies* (HCYSSOCI), at 22 percent. A total of 44 undergraduate student items had an overall response rate below 85 percent, and an additional 3 items had an item response rate below 85 percent in at least one category of control and level of institution, for a total of 47 variables requiring nonresponse bias analysis.

As shown in appendix table N-6, many graduate student items had high response rates, with 47 out of 77 imputed or select graduate student items having an overall response rate greater than 95 percent. The graduate student item with the lowest overall response rate was *Independent students: Spouse attending college* (SPINCOL), at 47 percent. There were 23 items with overall response rates below 85 percent and an additional 6 items with item response rates below 85 percent in one or more categories of control and level of institution, for a total of 29 graduate student items requiring nonresponse bias analysis.

The procedures used for item nonresponse bias analysis were the same as those used for unit nonresponse bias analysis before nonresponse weight adjustment. That is, using the variables listed in section 6.2.2, nonresponse bias was estimated for each category and tested for significance, and relative bias was computed. Effect size was also computed for each variable. Results of these analyses for each item analyzed varied across items and are summarized in appendix K, table K-78 for undergraduate student items and table K-79 for graduate student items.

Before imputation, for graduate student items overall, the percentage of characteristics with significant bias ranged from 31.4 to 62.90, and median effect sizes ranged from 0.07 to 0.34. Across control and level of institution, the percentage of characteristics with significant bias for graduate items ranged from 0.00 to 66.70, and median effect sizes ranged from 0.03 to 1.37. For

[39] Logical imputation is filling in missing data for cases where values can be deduced with certainty based upon logical or mathematical relationships among observed variables.

undergraduate student items overall, the percentage of characteristics with significant bias ranged from 26.8 to 60.20, and median effect sizes ranged from 0.06 to 0.38. Across control and level of institution, the percentage of characteristics with significant bias for undergraduate items ranged from 0.00 to 66.70, and median effect sizes ranged from 0.05 to 1.01.

Imputation procedures (see section 6.4) were conducted with a goal of reducing or eliminating item nonresponse bias. Although bias after imputation is not directly measurable, estimates were computed 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 survey items, the difference between the weighted preimputation mean and postimputation mean was computed; for categorical survey items, the difference between the weighted preimputation and postimputation means (proportions) was computed for each category. Final analysis weights were used for these comparisons. All differences were tested for statistical significance using *t* tests. For categorical variables, the differences reported in appendix K, tables K-78 and K-79 are size-weighted means of category-level differences[40] and are labeled as significant if any category-level difference is significant. These tests are complemented by effect size calculations. Effect sizes for categorical variables were 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 postimputation standard deviation.

For graduate students overall, statistically significant differences between the preimputation and postimputation means were found for 11 of 29 variables. Effect sizes for those differences ranged from 0.00 to 0.26. About 40 percent of the differences reported by control and level of institution that met reporting requirements were found to be statistically significant, with effect sizes ranging from 0.00 to 0.42.

For undergraduate students overall, statistically significant differences between preimputation and postimputation means were found for 35 of 47 variables. Effect sizes for those differences ranged from 0.00 to 0.30. About 42 percent of the differences reported by control and level of institution that met reporting

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

requirements were found to be statistically significant, with effect sizes ranging up to 1.2.

6.3 Variance Estimation

Every estimate calculated from a probability-based sample study, such as a mean, a percentage, or a regression coefficient, has an associated variance. Hypothesis testing, calculation of confidence intervals, and modeling, which use complex sample data, all require the calculation of variances using appropriate methods that account for the sampling design. Complex sample designs, like those used for NPSAS:18-AC, 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 sample studies may vary from those that would be expected if the sample was a simple random sample and the observations were independent and identically distributed random variables.

When testing hypotheses (e.g., conducting *t* tests, regression analyses) using weighted data from a study such as NPSAS that has a complex design, analysts should use methods that properly estimate variances. Two procedures for estimating variances of statistics from complex sample studies are the Taylor-series linearization procedure and bootstrap replication procedure, which are both available for the NPSAS data files. Stratum, primary sampling unit (PSU), and secondary sampling unit (SSU) identifiers are provided in the data file for use with the Taylor-series method, and bootstrap weights are provided for bootstrap replicate variance estimation.

The analysis strata, PSUs, and SSUs created for the Taylor-series procedure are discussed in section 6.3.1, and section 6.3.2 contains a discussion of the replicate weights created for the bootstrap procedure. Use of software packages for proper variance estimation is discussed in section 6.3.3.

Design effects were also calculated for key variables to assist analysts who do not have access to software that can compute Taylor-series or bootstrap variance estimates. The 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 sample had been selected). It is often used to measure the effects that sample design features have on the precision of 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 6.3.3 and appendix M.

6.3.1 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 sample studies, the Taylor-series procedure requires variance estimation strata and variance estimation PSUs, defined from the sampling strata and PSUs used in the first stage of sampling. For NPSAS:18-AC, separate analysis strata and PSUs were defined for undergraduate and graduate students so that analyses could be conducted for undergraduates and graduates separately since they are on separate files in the RUF; these are available for analyses of any characteristic (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 3, the PSUs were defined as the 2,210 institutions that provided eligible enrollment lists. Undergraduate- and graduate-enrolling institutions were defined as enrolling at least one responding undergraduate or graduate student, respectively. There were 2,170 undergraduate-enrolling institutions and 1,020 graduate-enrolling institutions that provided eligible enrollment lists.

The next step was to sort the PSUs (i.e., the institutions). The PSUs in the census strata (public 2-year and public 4-year) were sorted by state, level of the institution, and then by IPEDS ID. Within the “all other sectors” stratum, the PSUs were sorted by state and certainty (institution probability of selection equal to one) versus noncertainty (institution probability of selection less than one), and then by the selection order and control and level of the institution for the noncertainty institutions and by control and level of the institution and IPEDS ID for the certainty institutions.

Once sorted, some adjacent PSUs/institutions were combined 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, so that analyses could 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 917 analysis strata for undergraduate students and 390 analysis strata for graduate students.

Like PSUs, SSUs can also be used for Taylor-series variance estimation. SSUs were formed by randomly splitting responding students within a PSU into two groups. The RUF 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 that were used when assuming the first-stage units were sampled without replacement and that accounts for the FPC includes the analysis stratum (FANALSTR), analysis PSU (FANALPSU), analysis SSU (FANALSSU), and the count of PSUs in an analysis stratum (PSUCOUNT). Ultimately, FANALSTR equals the institution variance estimation stratum ANALSTR, and FANALPSU equals ANALPSU. FANALSSU was created by randomly dividing the NPSAS:18-AC analysis PSUs into two parts. These variables are by-products of the bootstrap variance estimation weights (described in section 6.3.2), and the justification for using the without-replacement variance formulas follows from the assumptions described in Kott 1988. Some values of the variance estimation strata, PSU, and SSU variables were combined in order to have at least two SSUs in each PSU and at least two PSUs in each stratum.

6.3.2 Bootstrap Replicate Weights

A separate set of 200 bootstrap weights was constructed as was done for NPSAS:04, NPSAS:08, NPSAS:12, and NPSAS:16. The final student weight (WTA000) described in section 6.1 was 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 of the full population, $\hat{\theta}$, by replicating the estimation procedure for each replicate and computing a simple variance of the replicate estimates, as follows:

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

where $\hat{\theta}_b^*$ is the estimate based on the bth replicate weight (where $b = 1$ to the number of replicates) and B is the total number of sets of replicate weights. The standard error for $\hat{\theta}$ can be calculated as the square root of the estimate of variance, var($\hat{\theta}$).

The replication method used was bootstrap with 200 replicates and with an FPC applied. The NPSAS:18-AC sample was selected using a sequential probability minimum replacement method (Chromy 1981). The bootstrap method requires the preparation of an additional set of replicate weights that will only be used for variance estimation. By applying the same student weight adjustment process to the bootstrap replicate weights, the bootstrap method of variance estimation will properly account for the effects of weight adjustments.

Incorporating the FPC factor at the institution stage of sampling recognizes the finite nature of the institution population. Treating the student population within the finite population of institutions as infinite preserves the ability to test statistical hypotheses about the student population with reference to probabilities based on the normal distribution.

In summary, the bootstrap variance estimation strategy for NPSAS satisfies the following requirements:

1. recognition of variance reduction due to stratification at all stages of sampling;
2. recognition of effects of unequal weighting;
3. recognition of possible increased variance due to sample clustering;
4. recognition of effects of weight adjustments on estimates;
5. satisfactory properties for estimating variances of nonlinear statistics and percentages, as well as for linear statistics;
6. ability to apply FPCs at the institution stage of sampling and reflect the reduction in variance due to the high sampling rates and censuses in some first-stage sampling strata; and
7. ability to test hypotheses about students on the basis of normal distribution theory by ignoring the FPCs at the student level of sampling.

Analysis strata and PSUs for bootstrap replicates. The analysis strata, PSUs, and SSUs created for the Taylor series served as the starting point for the replication method. In general, bootstrap strata were defined within design strata to contain two or three PSUs. The strata with three PSUs were needed to handle an odd number of PSUs responding within some natural grouping such as a design stratum. Some additional adjustments in stratum definitions were needed to handle single respondents in the certainty category within a design stratum.

The student sample within each PSU was systematically partitioned into two or more SSUs. The additional portioning of PSUs into subsamples of two units was performed specifically to accommodate a two-stage variance calculation with an FPC at the first stage.

Algorithm for defining bootstrap replicates. Commonly applied bootstrap variance estimation techniques account for requirements 1 through 5 previously outlined; however, to account for 6 and 7, a method adapted from Kott (1988) and Flyer (1987) was applied. The following notation is used in the steps delineated below:

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

$\widehat{N}_h$ = 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 *j* is selected within stratum *h* in institution *i* in the bootstrap sample when bootstrap sampling is at the SSU level; and

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

The process of forming replicates and computing replicate weights follows:

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

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

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 for the undergraduate sample and 200 times for the graduate sample to form 200 replicate samples for each type of student.

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.

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

Rationale for choosing *FPC$_h$*. Before NPSAS:04, schools were treated as being sampled with replacement or, equivalently, assumed to be sampled from a large population of schools. This assumption allowed the variance stratum finite population factor, *FPC$_h$*, to be treated as 1.00. The variance estimate was then based on school variances only, and the within-school component was fully

represented in expectation. The exception was that schools sampled with certainty were treated as strata; this set the finite population factor at the school level to 0, and the variance estimate was based entirely on the within-school variance. This did not account for nonresponse among certainty schools.

A finite population factor less than 1.00 was applied at the school stage of sampling and appropriately reduced the contribution of the school component of variance but also reduced the contribution represented in expectation for the within-school components of variance. It also allowed certainty schools to be treated as a sample when there were nonrespondents. Taylor-series formulas added back a partial contribution from the within-school components of variance to adjust for the reduced expectation. The same result was achieved in expectation by randomly switching between the two contributors to variance when using replication methods.

The choice of FPC for unbiased variance estimation can be based on exact joint probabilities of selection, but those unbiased procedures are difficult to implement and can yield highly unstable variance estimates.[41] This procedure is a reasonable approximation, suggested by Kott (1988), that can be implemented easily and should not suffer from the stability problems encountered with unbiased variance estimation procedures.

The choice of FPC was not completely arbitrary. For adequate approximation, FPC should tend toward 1.00 when the sampling rate is low and toward 0.00 when sampling is approaching the certainty level. For many studies where the sampling rate is low, for example, less than 0.05, the "with replacement" sampling assumption is used to simplify variance estimation with the knowledge that estimates of sampling error will be conservatively high. For many education studies, schools are sampled at a very high rate but below the certainty level. Using the "with replacement" sampling assumption unnecessarily penalizes the estimated precision. The Kott (1988) formula for FPC will be used rather than an arbitrary value of 1.00, which is used by the "with replacement" assumption.

Computing replicate weights. The number of replicate weights was set to 200 to ensure stable variance estimates for a variety of estimates. The student weight adjustments described in section 6.1.2[42] 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 the weight adjustments. For

[41] In the worst cases, unbiased estimation may allow estimates of variance for some samples to be negative in order to be unbiased over all possible samples.
[42] The institution weight adjustments cannot be replicated, due to the bootstrap methodology used.

some of the replicates, the bounds had to be loosened on the nonresponse and poststratification adjustment factors or model variables had to be collapsed 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 worked well for NPSAS:18-AC to achieve model convergence for all replicates and to minimize the effect of different models on the variance estimates.

6.3.3 Software Use for Variance Estimation

Table 52 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 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 may require editing before it is implemented. Additionally, an example of SUDAAN code is provided in appendix L. This example code, along with the code in table 52, can be helpful in writing code in other software packages.

Table 52. Use of analysis weights, replicate weights, and variance estimation strata, primary sampling unit (PSU), secondary sampling unit (SSU), and PSU count variables available from NPSAS:18-AC in selected survey data analysis software: 2018

Analysis weight for estimates	WTA000
Taylor-series variance estimation (with replacement)	
Variance estimation strata 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 analysis procedures	VARMETHOD = JACKKNIFE; 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)
Taylor-series variance estimation (without replacement)	
Variance estimation strata, 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 [pweight=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 bootstrap variance	
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: The survey data analysis software specifications are given for the following versions of the software packages: SUDAAN 11.0.3, Stata 12 and newer, SAS 9.3 and newer, IBM SPSS complex samples 25, and WesVar 4.3 and newer.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC).

6.3.4 Variance Approximation

The design effect (*DEFF*) 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 (*DEFT*) 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:18-AC, result in design effects greater than 1.0. That is, the design-based variance is larger than the simple random sample variance. Appendix M provides design effect estimates for important domains and estimates for undergraduate and graduate students to summarize the effects of stratification, multistage sampling, unequal probabilities of selection, and weight adjustments. These design effects were estimated using SUDAAN and the bootstrap variance estimation procedure described previously and in appendix L.

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 sample studies 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 previously, Taylor-series linearization and replication techniques can be used to compute more precise standard errors for data from complex sample studies. 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:18-AC data without using one of the software packages for analysis of complex sample study data can use the design effect tables in appendix M to make approximate

adjustments to the standard errors of statistics computed with the standard software packages that assume simple random sampling designs. (For details about the use of such software packages, see table 52 and appendix L.)

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.

6.4 Imputation

Missing data were imputed for all variables included in the restricted-use derived files (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),[43] the weighted sequential 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). Graduate, dependent undergraduate, and independent undergraduate students were all imputed separately.

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. Variables and vectors with low levels of missingness were imputed before variables where the rate of missingness was greater. That is, variables with smaller amounts of uncertainty were imputed first, and variables with larger amounts of uncertainty were imputed next. For each variable and vector, imputation classes were identified 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 dataset were available to form the imputation classes. To improve imputation quality, the 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:18-AC, there were five iterations, which improved quality without significantly slowing down the imputation process.

[43] An example of logical imputation follows: 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.

For NPSAS:18-AC, there are two deviations from the standard imputation process. First, because dependent and independent undergraduate financial aid may differ according to state law, each group of undergraduates was imputed separately by state—Hawaii did not require any imputation. Graduate students were imputed as a single group.

Second, the imputation process for income differed from past NPSAS studies because there was no income information from a student survey to supplement income information from FAFSA for FAFSA. FAFSA nonfilers on average have higher income than FAFSA filers. In previous iterations of NPSAS, the student survey has been used to acquire information about income for FAFSA nonfilers. This income information was used in the imputation process to impute income within a specified income range for FAFSA nonfilers with missing income. There was no student survey for NPSAS:18-AC; therefore, the missing data were missing not at random because income information about FAFSA filers was available to impute for FAFSA filers with missing income, but there was no information about FAFSA nonfilers available for imputation because it was missing by definition. Consequently, the usual imputation process would not be appropriate. To generate appropriately imputed values for FAFSA nonfilers with missing income, a modeling approach was used that incorporated FAFSA filing status.

To start, NPSAS:12 and NPSAS:16 income data were cross-classified by several sociodemographic variables and an indicator for FAFSA filing status. Second, the natural log of NPSAS:16 income was modeled as a function of the natural log of NPSAS:12 income. Third, the natural log of NPSAS:16 income was used as the independent variable in the model to predict the natural log of NPSAS:18-AC income. Finally, the cross-classified model estimates of NPSAS:18-AC income were applied to the appropriate individual student-level data with missing income values by transforming the estimates of NPSAS:18-AC income back to the original scale and adding a random component. This process was followed for graduate students, dependent undergraduates, and independent undergraduates separately.

To minimize the potential error due to imputation, quality checks were performed throughout the imputation process. Specifically, the distributions of the observed, imputed, and completed (observed and imputed) data were compared with screen variables for further investigation. For example, the distributions of observed income and imputed income differed because the missing data were primarily for students who did not apply for federal financial aid. Those who did not apply tended to have higher incomes than those who did apply. Consequently, the

imputed income distribution was higher than the observed income distribution. In addition, it was verified that the distributions within imputation classes were similar for the observed and imputed data and concluded that the completed (observed and imputed) distribution for income was reasonable. Item response rates are shown in section 6.2.3, and the observed and imputed distributions for eight key variables are provided in appendix N, tables N-1 through N-3.

6.5 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, including 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:18-AC 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 administrative record items. Perturbation was carried out through specific targeted, but undisclosed, swap rates. Because perturbation of the NPSAS:18-AC 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

Breiman, L., Friedman, J.H., Olshen, R.A., and Stone, C.J. (1984). *Classification and Regression Trees.* Toronto, Canada: Chapman and Hall.

Chromy, J.R. (1981). Variance Estimators for a Sequential Selection Procedure. In D. Krewski, R. Platek, and J.N.K. Rao (Eds.), *Current Topics in Survey Sampling* (pp. 329–347). New York: Academic Press.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cox, B.G. (1980). The Weighted Sequential Hot Deck Imputation Procedure. In *Proceedings of the Section on Survey Research Methods, American Statistical Association*, (pp. 721–726). Alexandria, VA: American Statistical Association.

Flyer, P.A. (1987). Finite Population Correction for Replication Estimates of Variance. In *Proceedings of the Survey Research Methods Section of the American Statistical Association* (pp. 732–736). Alexandria, VA: American Statistical Association.

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.

Iannacchione, V. (1982, February). *Weighted Sequential Hot Deck Imputation Macros*. Paper presented at the Seventh Annual SAS Users Group International Conference, San Francisco.

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). Ottawa, Canada: Statistics Canada.

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: American Statistical Association.

Lumley, T. (2014). *Survey: Analysis of Complex Survey Samples*. R package version 3.30.

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.

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 Language Manual, Volumes I–II, Release 11.* Research Triangle Park, NC: Research Triangle Institute.

Seastrom, M. (2014). *NCES Statistical Standards* (NCES 2014-097). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved February 1, 2021, from https://nces.ed.gov/statprog/2012.

Woodruff, R.S. (1971). A Simple Method for Approximating the Variance of a Complicated Estimate. *Journal of the American Statistical Association, 66*: 411–414.