

2019–20 National Postsecondary Student Aid Study (NPSAS:20)

Data File Documentation

NOVEMBER 2023

Jennifer Wine
Peter Siegel
Harper Haynes
RTI International

Tracy Hunt-White
Project Officer
National Center for Education Statistics

U.S. Department of Education

Miguel A. 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 2023

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-05-CO-0033 with RTI International. Mention of trade names, commercial products, or organizations does not imply endorsement by the U.S. Government.

Suggested Citation

Wine, J., Siegel, P., and Haynes, H. (2024). *2019–20 National Postsecondary Student Aid Study (NPSAS:20) Data File Documentation* (NCES 2024-482). 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 2019–20 National Postsecondary Student Aid Study (NPSAS:20), 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 cross-section of postsecondary students enrolled in Title IV eligible institutions in the 2019–20 academic year in all 50 states, the District of Columbia, and Puerto Rico. In addition, NPSAS:20 is designed to be state representative for undergraduate students overall and in public 2-year and public 4-year institutions. NPSAS:20 is unique among the NPSAS series, as it is the first NPSAS design that combines a student survey component similar to NPSAS:16 and earlier rounds with an administrative data-only collection as done in the 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC). Additionally, like NPSAS:18-AC, NPSAS:20 was designed to produce undergraduate student data that would be representative at the state level.

Sampling Design

The target population includes all students enrolled in Title IV eligible postsecondary institutions during the 2019–20 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:20 were consistent with those in the most recent iterations of NPSAS.

NPSAS:20 used a two-stage sampling design. In the first stage of sampling, a total of 3,110 institutions were selected from three strata: public 2-year institutions, public 4-year institutions, and an “all other sectors” stratum. The 3,110 sample institutions included all public 2-year ($n = 960$) and all public 4-year ($n = 780$) institutions as well as a sample of 1,370 institutions from the “all other sectors” stratum.

For the second stage of sampling, 380,100 students were sampled overall from institution-provided student enrollment lists; of these, 352,740 were undergraduate students. Additionally, because NPSAS:20 serves as the base year for the Beginning Postsecondary Students Longitudinal Study, an emphasis was

placed on sampling first-time beginning students (FTBs). For NPSAS:20, 52,620 of the students sampled were identified as potential FTBs.

After initial undergraduate student sampling, undergraduates were randomly divided into two groups, within student strata and within institution. Consistent with regular NPSAS procedures, one group would be asked to complete the student survey, while the other group would not be offered the student survey. Only administrative data were collected for the second group. From this random selection, 146,150 of the sampled undergraduate students were selected for the student survey. All 27,210 sampled graduate students would be asked to complete a survey, in addition to having student records and administrative data collected.

Institution Data Collection

NPSAS:20 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 3,070 institutions considered eligible for participation in NPSAS:20, a total of 2,190 (71 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. The student records instrument collected student-level data on general student information, enrollment, budget, and financial aid. Institutions provided student records data using the Postsecondary Data Portal.

A total of 1,840 (84 percent) institutions provided student records data, from which student records data were obtained for 74 percent ($n = 282,010$) of eligible student sample members; this group with student records data is hereafter defined as “student records student respondents.” This total included 73 percent of the total undergraduate students in the sample and 84 percent of the graduate students. Some student records student respondents are not included in the final data file if they did not meet the criteria to be considered a study respondent.

Student Survey

The NPSAS:20 student survey included core data elements used in previous NPSAS student surveys. Revisions to previous data elements and the inclusion of new data elements were informed by pretesting results and Technical Review Panel member input. The NPSAS:20 student survey instrument was a multimodal survey designed for web and telephone administration and consisted of eight key

content areas: Enrollment, Free Application for Federal Student Aid (FAFSA), Education Experiences, Financial Aid, Employment, Income and Expenses, and Background including a new COVID module. All graduate students and undergraduate students sampled for the survey were offered a full survey. Later in data collection, nonrespondents were offered an abbreviated version of the survey as a nonresponse conversion technique. Additionally, for the first time, the entire NPSAS:20 survey was translated into Spanish.

A total of 173,360 NPSAS:20 sample members were selected to participate in the student survey, which included 146,150 sampled undergraduate students selected for the student survey and all 27,210 sampled graduate students. NPSAS project staff used a multistep data collection design for locating, tracing, and contacting sample members to encourage survey completion. For sample members who could not be located, NPSAS staff used intensive tracing techniques. About 7 percent of the eligible sample (11,230 cases) required intensive tracing, and NPSAS staff located 70 percent of these cases. Overall, project staff successfully located 135,000 of the 173,360 NPSAS:20 sample members (78 percent) selected for the student survey group.

The full survey averaged 34.5 minutes to complete in English, while the survey averaged 53 minutes to complete in Spanish. Across both languages, full surveys completed on a mobile device were the fastest at 32.6 minutes on average, followed by an average of 34.4 minutes for nonmobile web, and finally, telephone interviews averaged 60.7 minutes to complete.

Of the 173,360 NPSAS:20 sample members selected for the student survey, 97 percent (168,160) were determined to be eligible for the study (see section 2.1.2 for more information about student eligibility criteria). Of these eligible students, 100,410 completed the full or abbreviated student survey (60 percent). One percent of student survey respondents (1,270) opted to complete the survey in Spanish. Ninety-six percent completed by web (mobile and nonmobile) and 4 percent by telephone.

Administrative Records Matching

In addition to student records data gathered from institutions, NPSAS:20 collected student data from administrative sources through data matching. As in prior NPSAS data collections, student data were gathered from the U.S. Department of Education's office of Federal Student Aid through the Central Processing System (CPS) and National Student Loan Data System (NSLDS), along with National Student Clearinghouse (NSC), ACT, College Board, and Veterans Benefits

Administration (VBA). CPS offers information collected on students' FAFSA form; of the 351,620 sample students sent to CPS for matching (93 percent), CPS information was gathered for 213,140 sampled students for the 2019–20 academic year (61 percent). NSLDS offers information on students who, at some point, had received Pell Grant or federal student loan funding. It provides data for the 2019–20 academic year as well as historical data for prior years. Of the 351,360 sample students sent to NSLDS for matching (92 percent), 198,320 students (56 percent) matched to NSLDS loan data, and 199,220 students (57 percent) matched to NSLDS Pell Grant data. NSC provides enrollment and degree records for the 2019–20 academic year. Of the entire sample, a total of 305,500 students (80 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, 95,350 (25 percent) matched to ACT while 67,870 (18 percent) matched to SAT. Lastly, VBA offers information on veterans education benefits. Of the 351,780 sample students sent to VBA for matching (93 percent), 23,030 students (7 percent) matched to VBA.

Data File Processing and Preparation

NPSAS:20 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.

A subset of the variables created in NPSAS:20 will be comparable to NPSAS:18-AC but will not be comparable to administrations of NPSAS prior to 2018. These variables are different primarily due to the lack of student survey data for the students randomly selected into the group that only had administrative data collected.

Another set of variables—“hybrid variables”—are variables that are derived slightly differently for students selected into the student survey and administrative-only groups but have comparable data elements available for both samples. These variables include versions that are applicable only to administrative respondents (the standard NPSAS variables), and versions that are comparable to administrations of NPSAS prior to 2018. For example, because private loan variables can only be created for survey respondents (i.e., not available in administrative data), variables that are loan ratios are not calculated

with private loans in the numerator nor denominator for administrative respondents.

State-Level Representation

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, 34 states were representative of undergraduate students in public 2-year institutions, and 40 states were representative of undergraduate students in public 4-year institutions.

Study Respondents

As a departure from prior traditional NPSAS collections, the NPSAS:20 study has a study respondent definition.

Respondents were defined as follows:

- *Survey respondents:* A survey respondent was defined as any undergraduate or graduate sample member who was determined to be eligible for the study and completed at least the enrollment and FAFSA sections of the student survey.
- *Administrative student respondents:* An administrative student respondent was defined as any undergraduate sample member who was determined to be eligible for the study, was enrolled for at least 1 month (based on student records or administrative sources), and at a minimum, had institution-provided valid key data in student records (i.e., FTB indicator, degree program, class level, high school graduation date/completion program completion date, enrollment in high school/completion program, and date of birth).
- *Study respondents:* A study respondent was defined as any undergraduate sample member who was an administrative student respondent and/or survey respondent. Study respondents who were survey respondents, but not administrative student respondents, had their administrative data imputed.

Weighting and Variance Estimation

Two statistical analysis weights were computed. The first statistical analysis weight was the student survey analysis weight. This was computed separately for

undergraduate and graduate students so that the survey respondents represent the target population. The second statistical analysis weight was the student study weight, computed so that the study respondents represent the population of undergraduates. This weight was designed for analysis of undergraduate student administrative (non-survey) data at the state level and within the public 2-year and public 4-year institution sectors within states for which the sample is sufficient for representation.

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 and students overall and across institution level and control, and for items with response rates 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 are greatly indebted to the students who generously participated in the survey. Their willingness to take the time to share their information and experiences made the 2019–20 National Postsecondary Student Aid Study (NPSAS:20) a success.

We gratefully acknowledge the assistance of the staff members at postsecondary institutions who provided data for the student data collection. We also extend special thanks to members of the Technical Review Panel, who provided their expertise to help shape NPSAS:20. Finally, we are grateful for the content experts who provided invaluable insight in the design of new student survey questions that capture the widespread postsecondary education changes in response to the national reaction to the coronavirus pandemic.

Contents

	PAGE
Executive Summary	iii
Acknowledgments.....	ix
List of Tables	xiii
List of Figures.....	xx
Chapter 1. Overview of NPSAS:20	1
1.1 Background and Purpose	2
1.2 Overview of Study Design.....	3
1.2.1 Study Management Systems and Security.....	4
1.3 Differences Between NPSAS and NPSAS:20	5
1.3.1 State-Level Representation	6
1.3.2 The Coronavirus Pandemic (COVID-19)	7
1.4 Schedule and Products	9
Chapter 2. Sampling Design.....	12
2.1 Respondent Universe	12
2.1.1 Institution Universe.....	12
2.1.2 Student Universe.....	13
2.2 Institution Sample	14
2.3 Student Sample	17
2.4 First-Time Beginning Student Sample.....	30
Chapter 3. Institution Data Collection Design, Outcomes, and Evaluation	34
3.1 Institution Data Collection Design and Systems	34
3.1.1 Institution Contacting System (ICS).....	34
3.1.2 Postsecondary Data Portal (PDP)	34
3.1.3 Student Records Collection Instrument	35
3.2 Institution Contacting, Recruitment, and Student Enrollment List Acquisition.....	37
3.2.1 Institution Contacting and Recruitment.....	37
3.2.2 Student Enrollment List Acquisition	39
3.2.3 Institution Recruitment and Student List Acquisition Outcomes	40
3.3 Student Records Data Collection.....	43
3.3.1 Student Records Collection From Institutions.....	43
3.3.2 Student Records Collection Outcomes	44
3.4 Institution Data Evaluation	50
Chapter 4. Student Survey Design, Data Collection, Outcomes, and Evaluation	52

	PAGE
4.1 Student Survey Design and Systems.....	52
4.1.1 Student Survey Design.....	52
4.1.2 Data Collection Systems.....	58
4.2 Student Survey Data Collection.....	59
4.2.1 Student Survey Data Collection Design	60
4.2.2 Study Website and Help Desk	63
4.2.3 Locating and Contacting Sample Members.....	64
4.2.4 Training of Student Data Collection Staff.....	68
4.2.5 Interviewing	69
4.3 Student Survey Data Collection Outcomes.....	71
4.3.1 Student Locating Results	72
4.3.2 Student Survey Response Rates.....	75
4.3.3 Survey Timing Burden.....	82
4.3.4 Contacts to Sample Members	90
4.3.5 Refusal Conversion.....	94
4.3.6 Responsive Design.....	96
4.3.7 Identification of Potential BPS-Eligible Sample Members	97
4.4 Data Collection Quality Control	98
4.4.1 Interview Monitoring	98
4.4.2 Quality Review Meetings	99
4.4.3 Debriefings.....	99
4.5 Evaluation of Student Survey Items	100
4.5.1 Coder System Usability	100
4.5.2 Help Text	102
4.5.3 Conversion Text.....	104
4.5.4 Item-Level Nonresponse.....	109
Chapter 5. Administrative Records Matching Overview and Outcomes	113
5.1 Administrative Records Matching and Outcomes	113
5.1.1 Central Processing System (CPS).....	113
5.1.2 National Student Loan Data System (NSLDS).....	116
5.1.3 National Student Clearinghouse (NSC).....	118
5.1.4 ACT and SAT	120
5.1.5 Veterans Benefits Administration (VBA).....	121
5.2 Administrative Records Quality Checks.....	122
Chapter 6. Data File Processing and Preparation	124
6.1 Overview of the NPSAS:20 Study Files.....	124
6.2 Post–Data–Collection Editing	128
6.2.1 Student Records	128
6.2.2 Student Survey	131
6.3 Derived Variable Construction	135
6.3.1 Comparison to Previous NPSAS Administrations.....	136
6.3.2 COVID-19 Aid Variables	141

	PAGE
Chapter 7. Weighting, Variance Estimation, and Imputation.....	143
7.1 Analysis Weights	143
7.1.1 Unit Response Rates	147
7.1.2 Institution Weights.....	161
7.1.3 Student Weights.....	186
7.1.4 Weighting Adjustment Summary and Evaluation	243
7.2 Nonresponse Bias Analysis.....	254
7.2.1 Nonresponse Bias Analysis Procedures.....	255
7.2.2 Unit Nonresponse Bias Analysis: Institution Level.....	257
7.2.3 Unit Nonresponse Bias Analysis: Student Level.....	267
7.2.4 Item Nonresponse Bias Analysis	277
7.3 Variance Estimation.....	280
7.3.1 Taylor Series	281
7.3.2 Bootstrap Replicate Weights.....	283
7.3.3 Software Use for Variance Estimation	288
7.3.4 Variance Approximation.....	291
7.4 Imputation	293
7.5 Disclosure Risk Analysis and Avoidance	294
References.....	296
Appendixes	
Appendix A. Acronyms and Abbreviations.....	A-1
Appendix B. Data Collection Outcomes and State Representation Maps.....	B-1
Appendix C. Student Survey Data Elements and English Instrument.....	C-1
Appendix D. Sampling Details	D-1
Appendix E. Student Records Data Elements and Instrument	E-1
Appendix F. Training Agenda for Institution Contactors.....	F-1
Appendix G. NPSAS Endorsements.....	G-1
Appendix H. Institution Data Collection Materials	H-1
Appendix I. Technical Review Panel (TRP) Members	I-1
Appendix J. Student Contacting Materials	J-1
Appendix K. Training Materials for Interviewers	K-1
Appendix L. Derived Variables	L-1
Appendix M. Nonresponse Bias Analysis	M-1
Appendix N. Design Effects	N-1
Appendix O. Weighted Analysis Considerations	O-1
Appendix P. Imputations	P-1

List of Tables

TABLE		PAGE
1.	Schedule of major activities for the NPSAS:20 full-scale survey: 2019–22.....	10
2.	Size of universe and number of institutions sampled, by institution stratum and state: 2019–20.....	15
3.	Number of sampled and eligible institutions, by control and level of institution: 2019–20.....	17
4.	Oversampled and undersampled student groups: 2019–20	20
5.	Number of expected and achieved undergraduate student sample sizes and percentage achieved overall, by institution stratum and state: 2019–20.....	22
6.	Number of expected and achieved sample students and percentage achieved for the student survey, by student sampling type and control and level of institution: 2019–20	25
7.	Expected and achieved student sample sizes and percentage achieved overall, by student stratum: 2019–20.....	26
8.	Expected and achieved student sample sizes and percentage achieved for the survey, by student stratum: 2019–20	27
9.	Initial classification of NPSAS:20 student sample overall, by student type and control and level of institution: 2019–20.....	28
10.	Initial classification of student sample for the survey, by student type and control and level of institution: 2019–20	29
11.	Potential first-time beginning student false-positive rates, by source and institution type: 2019–20	32
12.	Number and percentage of institutions providing enrollment lists and number of sampled students overall and for the survey, by control and level of institution: 2019–20	41
13.	Number and percentage of participating NPSAS:20 institutions, by 2015 Carnegie classification: 2019–20.....	42

TABLE	PAGE
14. Number and percentage of participating institutions, by Historically Black Colleges and Universities (HBCUs) and Hispanic-Serving Institutions (HSIs) classification: 2019–20.....	43
15. Enrollment lists and student records data collection results, by data collection mode and control and level of institution: 2019–20.....	45
16. Student records data collection results, by control and level of institution and student type: 2019–20	46
17. Percentage of institutions participating in the student records data collection, by institution stratum and state: 2019–20	48
18. Student records item-level completion rates, by data collection mode and data element: 2019–20	51
19. Calibration sample design, by experimental condition and data collection period: 2019–20	60
20. Cumulative calibration response rates, by experimental condition and data collection period: 2019–20	61
21. Student locating results, by control and level of institution and student type: 2019–20	72
22. Batch processing match rates, by method of tracing: 2019–20	73
23. Cases requiring intensive tracing, by control and level of institution and student type: 2019–20	74
24. Located and survey response rates of cases requiring intensive tracing procedures: 2019–20.....	75
25. Student survey response rates, by control and level of institution and student type: 2019–20	76
26. Student survey response rates, by data collection phase, control and level of institution, and student type: 2019–20	77
27. Student survey response rates, by survey type, control and level of institution, and student type: 2019–20.....	79
28. Student survey response rates, by mode of administration, control and level of institution, and student type: 2019–20.....	81
29. Number and percentage of NPSAS surveys included and excluded from the timing report: 2019–20	83

TABLE	PAGE
30.	Average time, in minutes, to complete the full and abbreviated NPSAS surveys, by mode of administration and student type: 2019–20.....85
31.	Average time, in minutes, to complete the full NPSAS survey, by student type and survey section: 2019–2087
32.	Forms with the highest average administration time, in seconds, by form and item type: 2019–20.....89
33.	Average number of contacts to sample members, by contact mode, control and level of institution, and student type: 2019–20.....91
34.	Average number of contacts to sample members, by contact mode, response status, and mode of administration: 2019–2093
35.	Refusal and refusal conversion rates, by control and level of institution and student type: 2019–2095
36.	First-time beginning student status determination, by student type: 2019–20.....98
37.	Percentage of responses coded in the survey, by mode of administration and coding system: 2019–20.....102
38.	Forms with at least 1 percent help text access rate, by mode of administration: 2019–20103
39.	Conversion rates for critical forms, by mode of administration: 2019–20.....106
40.	Conversion rates for critical forms with a hidden “Don’t know” response option, by mode of administration: 2019–20108
41.	Item nonresponse for items with 10 percent or more of data missing, by mode of administration: 2019–20.....110
42.	Results of Central Processing System (CPS) matching for 2019–20, by control and level of institution and student type: 2019–20.....115
43.	Results of National Student Loan Data System (NSLDS) loan and Pell Grant matching, by control and level of institution and student type: 2019–20117
44.	Results of National Student Clearinghouse matching, by control and level of institution and student type: 2019–20119
45.	Results of ACT and SAT matching, by control and level of institution and student type: 2019–20121

TABLE	PAGE
46.	Results of Veterans Benefits Administration (VBA) matching, by control and level of institution and student type: 2019–20.....122
47.	Restricted-use NPSAS:20 files, by file name, description, and file path: 2019–20126
48.	Description of missing data codes for student records: 2019–20130
49.	Description of missing data codes for student survey: 2019–20132
50.	Percentage of recoded values, by coding system: 2019–20.....134
51.	Summary of upcoding rates, by mode of administration and coding system: 2019–20135
52.	Description of missing data codes for derived variables: 2019–20 ...136
53.	NPSAS:20 analysis weights: 2019–20.....144
54.	Components of final analysis weights: 2019–20146
55.	Institution response rates for undergraduate- and graduate-enrolling institutions, by collection type and by control and level of institution: 2019–20.....149
56.	Institution response rates for undergraduate-enrolling study institutions, by state and institution stratum: 2019–20150
57.	Student response rates for undergraduate and graduate students, by collection type and by control and level of institution: 2019–20.....156
58.	Student response rates for undergraduate study students, by state and institution stratum: 2019–20157
59.	Candidate predictor variables for institution nonresponse adjustment models: 2019–20165
60.	Weight adjustment factors for institution nonresponse adjustment for undergraduate-enrolling survey institutions, by model predictor variable: 2019–20166
61.	Weight adjustment factors for institution nonresponse adjustment for graduate-enrolling institutions, by model predictor variable: 2019–20.....171
62.	Weight adjustment factors for institution nonresponse adjustment for undergraduate-enrolling study institutions, by state and institution stratum: 2019–20174

TABLE	PAGE
63.	Weight adjustment factors for institution poststratification adjustment for undergraduate-enrolling survey institutions, by poststratification category: 2019–20.....180
64.	Weight adjustment factors for institution poststratification adjustment for graduate-enrolling institutions, by control and level of institution by size: 2019–20.....181
65.	Weight adjustment factors for institution poststratification adjustment for undergraduate-enrolling study institutions, by state and institution stratum: 2019–20.....181
66.	Weight adjustment factors for unknown student eligibility status adjustment for undergraduate and graduate students, by control and level of institution: 2019–20190
67.	Candidate predictor variables for student nonresponse adjustment models: 2019–20.....192
68.	Weight adjustment factors for undergraduate survey student not located nonresponse adjustment, by model predictor variable: 2019–20.....194
69.	Weight adjustment factors for undergraduate survey student refusal nonresponse adjustment, by model predictor variable: 2019–20.....200
70.	Weight adjustment factors for undergraduate survey student other nonresponse adjustment, by model predictor variable: 2019–20.....205
71.	Weight adjustment factors for graduate student not located nonresponse adjustment, by model predictor variable: 2019–20.....211
72.	Weight adjustment factors for graduate student other nonresponse adjustment, by model predictor variable: 2019–20.....215
73.	Weight adjustment factors for undergraduate student study nonresponse adjustment, by state and institution stratum: 2019–20.....219
74.	Weight adjustment factors for undergraduate student study nonresponse adjustment, by model predictor variable: 2019–20.....223
75.	Weight adjustment factors for undergraduate survey student poststratification, by poststratification category: 2019–20.....228

TABLE	PAGE
76.	Weight adjustment factors for graduate student poststratification, by poststratification category: 2019–20.....232
77.	Weight adjustment factors for undergraduate student study poststratification, by state and institution stratum: 2019–20235
78.	Weight adjustment factors for undergraduate student study poststratification, by national poststratification category: 2019–20.....239
79.	Institution weight distribution and unequal weighting effects for undergraduate-enrolling survey institutions, by control and level of institution: 2019–20.....244
80.	Institution weight distribution and unequal weighting effects for graduate-enrolling institutions, by control and level of institution: 2019–20.....244
81.	Institution weight distribution and unequal weighting effects for undergraduate-enrolling study institutions, by control and level of institution: 2019–20.....245
82.	Undergraduate student survey respondents weight distribution and unequal weighting effects, by control and level of institution: 2019–20.....250
83.	Graduate student respondents weight distribution and unequal weighting effects, by control and level of institution: 2019–20250
84.	Undergraduate student study respondents weight distribution and unequal weighting effects, by control and level of institution: 2019–20.....251
85.	Variables used in institution-level unit nonresponse bias analysis: 2019–20.....258
86.	Summary of institution-level survey nonresponse bias analysis for undergraduate-enrolling institutions, by control and level of institution: 2019–20260
87.	Summary of institution-level survey analysis after poststratification for undergraduate-enrolling institutions, by control and level of institution: 2019–20261
88.	Summary of institution-level study nonresponse bias analysis for undergraduate-enrolling institutions, by control and level of institution: 2019–20263

TABLE	PAGE
89.	Summary of institution-level study analysis after poststratification for undergraduate-enrolling institutions, by control and level of institution: 2019–20264
90.	Summary of institution-level nonresponse bias analysis for graduate-enrolling institutions, by control and level of institution: 2019–20.....266
91.	Summary of institution-level analysis after poststratification for graduate-enrolling institutions, by control and level of institution: 2019–20.....267
92.	Variables used in student-level nonresponse bias analyses: 2019–20.....268
93.	Summary of undergraduate student-level survey nonresponse bias analysis, by control and level of institution: 2019–20.....270
94.	Summary of undergraduate student-level survey analysis after poststratification, by control and level of institution: 2019–20271
95.	Summary of undergraduate student-level study nonresponse bias analysis, by control and level of institution: 2019–20273
96.	Summary of undergraduate student-level study analysis after poststratification, by control and level of institution: 2019–20274
97.	Summary of graduate student-level nonresponse bias analysis, by control and level of institution: 2019–20276
98.	Summary of graduate student-level analysis after poststratification, by control and level of institution: 2019–20277
99.	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:20 in selected survey data analysis software: 2019–20.....289

List of Figures

FIGURE	PAGE
1. Postsecondary Data Portal home page: 2019–20.....	35
2. Postsecondary Data Portal task menu: 2019–20.....	40
3. Home page for NPSAS:20 study website: 2019–20.....	63
4. NPSAS:20 sample member locating and contacting activities.....	65
5. Data collection phases: 2019–20	70
6. NPSAS:20 student survey locating and response results: 2019–20.....	71
7. Distribution of student survey respondents, by mode of administration: 2019–20	82
8. Receiver operating characteristic (ROC) curve for overall institution response propensity for undergraduate-enrolling survey institutions: 2019–20.....	247
9. Receiver operating characteristic (ROC) curve for overall institution response propensity for graduate-enrolling institutions: 2019–20.....	248
10. Receiver operating characteristic (ROC) curve for overall institution response propensity for undergraduate-enrolling study institutions 2019–20.....	249
11. Receiver operating characteristic (ROC) curve for overall undergraduate survey student response propensity: 2019–20	252
12. Receiver operating characteristic (ROC) curve for overall graduate student response propensity: 2019–20	253
13. Receiver operating characteristic (ROC) curve for overall undergraduate study student response propensity: 2019–20	254

Chapter 1. Overview of NPSAS:20

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

This report describes the design, methods, and results of the 2019–20 National Postsecondary Student Aid Study (NPSAS:20), conducted by the U.S. Department of Education’s National Center for Education Statistics (NCES). The following legislation authorizes this and previous cycles of NPSAS, as well as the two longitudinal studies derived from NPSAS—the Beginning Postsecondary Students Longitudinal Study (BPS) and the Baccalaureate and Beyond Longitudinal Study (B&B):

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

Since 1987, NPSAS has provided current information on financial aid programs, and typically been fielded every 3 to 4 years, with the exception of the 2017–18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18 AC),¹ which was the first NPSAS to be conducted at the midway point in a typical 4-year NPSAS cycle. The regularity of NPSAS administration makes it possible to examine the impact of changes in federal policy concerning higher education over time. With respect to federal student aid, eligibility criteria change, grant and loan amounts fluctuate, and the balance between various aid options can

¹ For more information about NPSAS:18-AC, see <https://nces.ed.gov/pubs2022/2022477.pdf>.

shift dramatically. A recurring study such as NPSAS is essential to understanding those changes, particularly as they affect how students and families pay for college.

NPSAS:20 is a nationally representative cross-section of postsecondary students enrolled in Title IV eligible institutions in the 2019–20 academic year in all 50 states, the District of Columbia, and Puerto Rico. NPSAS:20 differs from previous iterations of NPSAS because it is the first study to combine regular NPSAS data collection with the administrative data-only collection as done in NPSAS:18-AC. Additionally, like NPSAS:18-AC, NPSAS:20 was designed to produce undergraduate student data that would be representative at the state level.

Chapter 1 of this report provides an overview of the background and purpose of NPSAS, as well as the study design, schedule, and products. Chapter 2 describes the sampling design and the steps NPSAS statisticians used to select the institution and student samples. Chapter 3 describes the design, outcomes, and evaluation activities associated with institution data collection. Chapter 4 provides details on the student survey design, data collection, outcomes, and evaluations. Chapter 5 contains information on the student administrative records matching activities and outcomes. Chapter 6 contains a description of NPSAS:20 study files, along with data file processing and editing after data collection. Chapter 7 provides information on weighting, imputation, bias analysis, and variance estimation. Appendix A provides a list of acronyms and abbreviations used throughout the report.

Tables and figures throughout this report present relevant analyses from the full-scale study. Unless otherwise indicated, a probability level of .05 was used for all tests of significance conducted for NPSAS:20 evaluations. Because of rounding, row and column entries in tables may not sum to their respective totals, and reported percentages may differ somewhat from those that would result from the rounded numbers. Rounding is used to ensure the confidentiality of study participants.

1.1 Background and Purpose

NPSAS is a comprehensive, nationwide study of how students and their families pay for postsecondary education. It features a nationally representative sample of undergraduate and graduate students enrolled in Title IV eligible postsecondary education institutions in the United States. The institution sampling frame

includes public institutions and private institutions (both for-profit and nonprofit) and spans less-than-2-year institutions to 4-year colleges and universities.

NPSAS collections traditionally serve as the base-year data collection for one of two longitudinal studies, BPS and B&B. NPSAS:20 is the base-year data collection of the 2020 cohort for the 2020/22 Beginning Postsecondary Students Longitudinal Study (BPS:20/22), a study of first-time beginning students (FTBs) conducted 3 years (BPS:20/22) and 6 years (BPS:20/25) after they begin postsecondary education. Subsets of questions in the NPSAS:20 student survey focus on aspects of the experience of FTBs in their first year of postsecondary education, including student debt and education experiences, to support longitudinal analyses of student choices related to major, persistence, and completion.

1.2 Overview of Study Design

The data collected for NPSAS:20 come from three sources: (1) postsecondary institutions, (2) a student survey, and (3) administrative data records. Only a subset of sampled undergraduate students were offered the student survey. All graduate students were offered the survey. To facilitate selection of a nationally representative sample, the target population included all students enrolled in Title IV eligible postsecondary institutions during the 2019–20 academic year in the 50 states, the District of Columbia, and Puerto Rico. A sample was generated from this target population in order to be nationally representative for undergraduate and graduate students as well as representative at the state level for undergraduate students.

NPSAS staff contacted institutions to request student-level information on enrollment and financial aid. Many of the required student financial aid data elements requested from institutions were also available for verification from the U.S. Department of Education’s office of Federal Student Aid (FSA) through the Central Processing System (CPS) and the National Student Loan Data System (NSLDS). CPS gathers information from the Free Application for Federal Student Aid (FAFSA) to determine federal aid eligibility. NSLDS contains student-level data on Pell Grants and federal student loans. NPSAS staff obtained these data through file matching with both CPS and NSLDS data to reduce the data collection burden on sampled institutions and sampled students. Early in the institution data collection process, NPSAS staff performed checks to verify that an institution was open and eligible for the study, and confirmed participation in Title IV financial aid programs, although explicit confirmation by the institution

was not always possible. After NPSAS staff sampled students from institution-provided enrollment lists, student records were collected from the institutions, and data were collected from students using a mobile-enhanced web-based student survey.

To supplement the institution and student survey data collections, NPSAS staff gathered additional data for the NPSAS:20 student sample from a variety of administrative data sources. These included the previously mentioned data obtained from CPS and NSLDS, as well as from the National Student Clearinghouse (NSC), ACT, the College Board, and the Veterans Benefits Administration (VBA).

1.2.1 Study Management Systems and Security

The design of NPSAS:20 required the collection and management of personally identifiable information (PII) at multiple points throughout the course of the study. PII was required when updating student enrollment status at the institutions, tracing sample members, and collecting data. These data were also managed during data processing and data product preparation activities. This section describes the study management systems used to securely collect, transmit, and process NPSAS:20 data.

On a nightly basis, NPSAS staff downloaded student record and student survey data from NCES web servers to the Enhanced Security Network (ESN) via a secure web service, a process that meets the requirements for a Federal Information Processing Standards moderate system.² Once data were in the ESN, staff cleaned the data and performed quality analysis. The websites used for the NPSAS:20 data collection resided on NCES's Secure Sockets Layer (SSL)-certified servers with a secure data connection. SSL protocols are used to encrypt data transmitted over the Internet, and all parts of the websites that collect student data are password protected. The forms that gather data on these websites require session cookies to run in accordance with the U.S. Department of Education's privacy policy for the use of cookies.

The Integrated Management System (IMS) is an RTI International project management tool designed to give NPSAS staff ready access to a repository of reports and other project information and deliverables. The IMS provides a web-

² Federal Information Processing Standards (FIPS) are standards and guidelines for federal computer systems that are developed by the National Institute of Standards and Technology in accordance with the Federal Information Security Management Act and approved by the U.S. Secretary of Commerce. For more information about FIPS, see <https://csrc.nist.gov/publications/fips>.

based utility with which to track major components of the study, including receipt control, survey management and implementation, and case management. The IMS combines management reports, research materials, and archives, and makes them accessible via a website protected by log-in/password security and SSL encryption. The IMS website, accessible to NCES and NPSAS staff, serves as a repository for up-to-date information about study operations. The website houses the document archive, the directory of the data library, and daily reports generated from the receipt control and case management modules of the IMS. No PII is accessible from the IMS.

For further information regarding study management systems and security procedures, refer to section 3.1 for institution data collection systems, section 4.1.2 for student data collection systems, and section 5.1 for administrative data matching.

1.3 Differences Between NPSAS and NPSAS:20

Traditionally, NPSAS includes administrative file matching, student records collected from postsecondary institutions, and a student survey. NPSAS:20 is the first study to combine regular NPSAS data collection with the administrative data collection only per the NPSAS:18-AC design. NPSAS:18-AC was the first NPSAS solely reliant on administrative data sources, designed to provide researchers and policymakers with the ability to create national and state-level undergraduate estimates for all 50 states, the District of Columbia, and Puerto Rico.³ The inclusion of an administrative collection in NPSAS:20 increases the student sample size, which enhances the precision of estimates and offers more statistical power.

The remainder of this section describes additional differences between a traditional NPSAS and NPSAS:20, particularly the design of NPSAS:20 to be representative at the national and state levels. Finally, this section includes discussion of the impact the coronavirus pandemic had on data collection and survey design.

³ NPSAS:18-AC was designed to represent all 50 states, the District of Columbia, and Puerto Rico for undergraduate students. However, the final data are representative of 30 states (including the District of Columbia and Puerto Rico). For more information about state-level representativeness of NPSAS:18-AC, see appendix C in the NPSAS:18-AC data file documentation: https://nces.ed.gov/pubs2022/2022477_1.pdf.

1.3.1 State-Level Representation

As a departure from prior NPSAS administrations with a student survey component, NPSAS:20 was designed to produce undergraduate student data that would be both state representative and nationally representative. As mentioned earlier, students were sampled from institutions in all 50 states, the District of Columbia, and Puerto Rico. This sample was designed to provide state-level estimates that describe undergraduate students for public 2-year and public 4-year institutions as well as overall. The sample is also nationally representative for undergraduate and graduate students. Other NPSAS administrations have not typically been designed to produce representative data at the state level, with the exception⁴ of NPSAS:18-AC, which collected data from administrative sources only. NPSAS:20 complements NPSAS:18-AC by also including the nationally representative student survey component in a NPSAS with state representativeness. NPSAS:20 provides researchers, policymakers, and analysts a unique and rich data source for studying state-level undergraduate educational issues that prior NPSAS administrations may not have been suited to examine.

Many of the important questions that NPSAS:20 aims to address are the same as in the past years. College cost increases, net price levels, and increases in student loan debt continue to be central issues. The NPSAS:20 data can be used to address policy issues related to changes in federal financial aid programs resulting from the anticipated reauthorization of the Higher Education Act. Moreover, NPSAS:20 is well suited for state policymakers and researchers to use these state-representative data to address policy questions relevant to undergraduate students in their state. Estimates at the state level are valuable because they enable examination of state financial aid policies, in-state retention of students, and the intersection of state and federal financial aid. Because these estimates exist in both NPSAS:20 and NPSAS:18-AC, researchers and policymakers can also assess how these issues change over time.

NPSAS:20 was designed to represent all 50 states, the District of Columbia, and Puerto Rico for undergraduate students. However, the final data are not representative of all states but are representative as follows:

- 30 states representative for undergraduate students overall: Arizona, Arkansas, California, Colorado, Florida, Georgia, Hawaii, Illinois, Iowa,

⁴ Two prior NPSAS administrations did include state representativeness, although not for all 50 states, the District of Columbia, and Puerto Rico. The 2003–04 National Postsecondary Student Aid Study (NPSAS:04) was representative of 12 states, and the 2007–08 National Postsecondary Student Aid Study (NPSAS:08) was representative of 6 states.

Kentucky, Maine, Maryland, Massachusetts, Michigan, Minnesota, Nebraska, Nevada, New Jersey, New Mexico, New York, North Carolina, Ohio, Oklahoma, Pennsylvania, South Carolina, Tennessee, Texas, Virginia, Washington, and West Virginia;

- 34 states representative of undergraduate students in public 2-year institutions: Alabama, Arizona, Arkansas, California, Colorado, Connecticut, Florida, Georgia, Hawaii, Illinois, Indiana, Iowa, Kentucky, Maine, Maryland, Massachusetts, Michigan, Minnesota, Mississippi, Nebraska, New Jersey, New Mexico, New York, North Carolina, Ohio, Oklahoma, Pennsylvania, South Carolina, Tennessee, Texas, Vermont, Virginia, Washington, and West Virginia; and
- 40 states representative of undergraduate students in public 4-year institutions: Arizona, Arkansas, California, Colorado, Florida, Georgia, Hawaii, Illinois, Iowa, Kansas, Kentucky, Louisiana, Maine, Maryland, Massachusetts, Michigan, Minnesota, Mississippi, Nebraska, Nevada, New Hampshire, New Jersey, New Mexico, New York, North Carolina, Ohio, Oklahoma, Oregon, Pennsylvania, Puerto Rico, South Carolina, South Dakota, Tennessee, Texas, Utah, Virginia, Washington, Wisconsin, West Virginia, and Wyoming.

Refer to chapter 2 for sampling details on the national and state-representative nature of NPSAS:20, chapter 7 for the criteria for determining state representativeness, and appendix B for state-level institution participation rates and representation maps.

1.3.2 The Coronavirus Pandemic (COVID-19)

NPSAS:20 student surveys started in March 2020, just as the coronavirus pandemic began affecting students' educational experiences. This section describes the changes made to NPSAS:20 data collection protocols and survey instrument design because of the coronavirus pandemic.

Notably, the coronavirus pandemic impacted response rates by sampled institutions and students. Institutions experienced delays in providing the enrollment lists needed for sampling, forcing an extension of the overall data collection schedule. The NPSAS:20 list collection period, which began in January 2020, was scheduled to end in July 2020 but was extended through September 2020 to accommodate institution disruptions caused by the coronavirus pandemic. Institution contacting materials were revised to acknowledge the coronavirus pandemic and its effects on postsecondary staff and students.

Students' lives were upended in the spring of 2020 as postsecondary institutions made a sudden shift to virtual enrollment in the middle of the spring 2020 semester. In addition to their classes being affected, many students were either required to move out of campus-owned housing or encouraged to move out of off-campus housing and return to their permanent residences. As a result, response rates for student-level data collection lagged expectations. Data from past cross-sectional and longitudinal studies were used along with the NPSAS:20 response rates to identify student subgroups of interest with low final response rate estimates that would increase nonresponse bias and decrease precision. These groups were potential FTBs,⁵ undergraduate students from private for-profit institutions,⁶ and undergraduates who did not file a FAFSA (i.e., FAFSA nonfilers).⁷ To encourage participation, these three student subgroups were offered an additional \$10 incentive, and additional contacting protocols (e.g., tracing, text messaging, e-mail, and mailings) were implemented.

Items about the coronavirus pandemic were added to the student records and student survey instruments to collect information about the impacts this event had on postsecondary education between January 1 and June 30, 2020. Revisions to both instruments were implemented in April 2020, and the coronavirus pandemic items were administered through the end of data collection in January 2021.

The student records instrument was revised to provide updated instructions for reporting student records data for students impacted by the coronavirus pandemic; these included a new response option for reporting special financial aid awards

⁵ FTBs are a critical subgroup of the NPSAS:20 sample and are oversampled accordingly. FTBs in the 2019–20 NPSAS academic year form the basis for the two follow-ups that make up BPS. Given the shorter than anticipated data collection period for students who attended institutions that submitted lists later than anticipated, an increased incentive offer maximized the participation of potential FTBs.

⁶ Historically, private for-profit institutions have been less likely to submit lists, and NPSAS:20 is no different. Further complicating matters is that private for-profit institutions are generally sampled later in the process given the institutions' continuous enrollment status. As a result of these two issues, the number of students sampled from private for-profit institutions has been low and the students sampled near the end of the sampling period would not have had adequate time in data collection for sufficient follow-up to maximize response for this group. Because private for-profit institutions are an important sector in higher education, NPSAS needs adequate numbers of students from these institutions to be representative, with minimal nonresponse bias and more accurate estimates, of this unique sector in higher education.

⁷ Approximately 67 percent of student survey respondents file the FAFSA, which provides high-quality data on important topics not otherwise obtained (e.g., student earnings, family income for dependent students). Imputation is difficult for these data given that it is not a random group that does not file the FAFSA—students from both very wealthy and very poor families tend not to file. As a result, the student survey is a vital source of information for students who do not fill out the FAFSA.

related to the coronavirus pandemic and two new items for collecting coronavirus pandemic-related refunds of tuition and fees or room and board.

To collect student-level information on the impact of the coronavirus pandemic, survey items were added to collect data on enrollment and employment disruptions, food and housing experiences, and resources and information provided by the institution. Because these items were based on effects seen early in the pandemic, they likely do not reflect all possible student experiences. In addition to these new topics, the student survey instrument was revised to distinguish full tuition refunds from the institution due solely to the coronavirus pandemic from all other full tuition refund reasons. This is because full tuition refunds due to enrollment changes unrelated to the pandemic are a component of study eligibility determination, as it indicates the student was not enrolled during the 2019–20 academic year.

When these items were added to the NPSAS:20 student survey, many students and institutions were in a period of flux and uncertainty because of the pandemic. Changes to how students were accessing and paying for education began in March 2020. Institutions closed or moved to online instruction; new rules about how student aid could be used were released (Office of Postsecondary Education 2020); and Congress enacted the Coronavirus Aid, Relief, and Economic Security (CARES) Act. The CARES Act suspended federal student loan payments, paused collections on defaulted loans, and waived interest on federal loans through January 2022. These developments and guidance from a wide range of education experts were the basis for the coronavirus pandemic items in the NPSAS:20 student survey (refer to appendix C for the student survey instrument).

1.4 Schedule and Products

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

Table 1. Schedule of major activities for the NPSAS:20 full-scale survey: 2019–22

NPSAS:20 activity	Start date	End date
Contact institutions to request student enrollment lists	Oct. 16, 2019	Sept. 30, 2020
Collect student enrollment lists	Jan. 16, 2020	Sept. 30, 2020
Select student sample	Jan. 21, 2020	Jul. 20, 2020
Collect student survey self-administered web-based data	Mar. 3, 2020	Jan. 31, 2021
Conduct telephone interviews with students	Mar. 3, 2020	Jan. 31, 2021
Collect student data from institution records	Mar. 10, 2020	Jan. 31, 2021
Conduct nonresponse conversion efforts	Jul. 22, 2020	Mar. 19, 2021
Process data, construct data files	Jan. 13, 2020	Sept. 28, 2022
Prepare/update reports	Jan. 13, 2020	Nov. 13, 2023

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

The following reports will be available on the NCES website at

<https://nces.ed.gov/surveys/npsas/>:

- *2019–20 National Postsecondary Student Aid Study (NPSAS:20): First Look at the Impact of the Coronavirus (COVID-19) Pandemic on Undergraduate Student Enrollment, Housing, and Finances (Preliminary Data)*; and
- *2019–20 National Postsecondary Student Aid Study (NPSAS:20): First Look at Student Financial Aid Estimates for 2019–20.*

Microlevel data files, associated codebooks, and data file documentation for NPSAS:20 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 data analysis tools in the DataLab application, found at <https://nces.ed.gov/datalab>, to review and analyze the derived files for NPSAS:20. DataLab includes undergraduate and graduate files that allow users to generate national estimates for student survey respondents. There is also an additional undergraduate file to allow users to generate state-level estimates for undergraduate students based on the administrative sample of NPSAS:20. In DataLab, which includes the interactive PowerStats data analysis tool, data for survey respondents using the survey weight are available under the headings for “National Postsecondary Student Aid Study, Undergraduate” and “National Postsecondary Student Aid Study, Graduate,” while data for study respondents using the study weight is available under the heading for “National Postsecondary

Student Aid Study-Administrative Collection.” Within DataLab, PowerStats can produce summary statistics and complex tables, as well as estimate regression models. It permits analysis without disclosing microlevel data to the user and suppresses or flags any estimates that fail to meet NCES reporting standards. DataLab also contains the Tables Library, which houses thousands of published analysis tables sortable by topic, publication, and source.

Chapter 2. Sampling Design

This chapter describes the target population, sampling design, and sampling methods for NPSAS:20. 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, statisticians from NCES, and representatives of other federal agencies.⁸

2.1 Respondent Universe

NPSAS:20 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:20, 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:20 institution (first-stage) sampling frame included all levels (less-than-2-year, 2-year, and 4-year) and control classifications (public, private nonprofit, and private for-profit) of Title IV eligible postsecondary institutions in the United States, including the District of Columbia and Puerto Rico. To be eligible for NPSAS:20, 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;⁹
- offered courses that were open to more than the employees or members of the company or group (e.g., union) that administered the institution;

⁸ See appendix I for a complete list of TRP participants.

⁹ Clock hours reflect the actual hours of class attendance. Title IV regulations require clock hour measurement for Title IV if (1) the institution's accrediting agency requires it; (2) the institution must measure student progress in clock hours to receive federal or state approval or licensure to offer the program; or (3) completion of clock hours is a requirement for graduates to apply for licensure or the authorization to practice the occupation that the student is intending to practice.

- been located in the 50 states, the District of Columbia, or Puerto Rico;
- not been a U.S. service academy institution; and
- signed the Title IV participation agreement with the U.S. Department of Education.¹⁰

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:20 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 could be made with data from previous NPSAS cycles (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 the 2003–04 National Postsecondary Student Aid Study (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:20 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:20, the target population consisted of, and eligible students are defined as,

¹⁰ 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.

students who were enrolled at any time between July 1, 2019, and June 30, 2020,¹¹ at eligible postsecondary institutions in the United States and who were

- enrolled in either (1) an academic program; (2) at least one course for credit that could be applied toward fulfilling the requirements for an academic degree; (3) exclusively noncredit remedial coursework but determined by the institution to be eligible for Title IV aid; or (4) an occupational or vocational program that required at least 3 months or 300 clock hours of instruction to receive a degree, certificate, or other type of formal award;
- not currently enrolled in high school; and
- not enrolled solely in a high school completion program.

2.2 Institution Sample

The NPSAS:20 institution frame was constructed from the IPEDS 2018–19 Institutional Characteristics Header (IC-H), 2018–19 Institutional Characteristics (IC), 2017–18 12-month Enrollment (E12), and 2017 Fall Enrollment (EF) files. Web searches were performed to find online articles to identify and then exclude institutions that were still in IPEDS but were no longer eligible for NPSAS:20.

The institution strata were based on sectors (combinations of institution level and control) within each of the 52 states and territories, for a total of 152 sampling strata ($52 \times 3 - 4$).¹²

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.

¹¹ So as to not delay data collection, enrollment lists covered the period of July 1, 2019, through April 30, 2020, for institutions with specific enrollment dates and July 1, 2019, through March 31, 2020, for institutions with continuous enrollment. The dates of March 31 and April 30 were selected to include virtually all students enrolled prior to the summer term. Any lack of coverage resulting from the truncated enrollment periods is accounted for by the poststratification weight adjustment, which is described in chapter 7.

¹² Four states had no public 2-year institutions during sampling (Alaska, Delaware, the District of Columbia, and Nevada).

Table 2. Size of universe and number of institutions sampled, by institution stratum and state: 2019–20

State	All sectors		Institution stratum ¹					
			Public 2-year		Public 4-year		All other sectors	
	Size of universe ²	Sample size	Size of universe ²	Sample size	Size of universe ²	Sample size	Size of universe ²	Sample size
All states	6,230	3,110	960	960	780	780	4,490	1,370
Alabama	80	70	20	20	10	10	40	30
Alaska	10	10	†	†	#	#	10	10
Arizona	110	60	20	20	10	10	80	30
Arkansas	80	60	20	20	10	10	50	30
California	640	190	110	110	50	50	480	30
Colorado	100	60	10	10	20	20	70	30
Connecticut	70	50	10	10	10	10	50	30
Delaware	20	20	†	†	#	#	10	10
District of Columbia	20	20	†	†	#	#	20	20
Florida	330	100	30	30	40	40	260	30
Georgia	150	80	20	20	30	30	90	30
Hawaii	20	20	10	10	#	#	10	10
Idaho	40	40	#	#	#	#	30	30
Illinois	240	90	50	50	10	10	180	30
Indiana	110	50	#	#	20	20	90	30
Iowa	80	50	20	20	10	10	60	30
Kansas	80	60	30	30	10	10	40	30
Kentucky	90	50	20	20	10	10	70	30
Louisiana	110	60	20	20	20	20	80	30
Maine	40	40	10	10	10	10	20	20
Maryland	80	60	20	20	10	10	50	30
Massachusetts	160	60	20	20	10	10	130	30
Michigan	160	80	20	20	20	20	120	30
Minnesota	100	70	30	30	10	10	50	30
Mississippi	60	60	20	20	10	10	30	30
Missouri	150	60	20	20	10	10	120	30
Montana	30	30	10	10	10	10	10	10
Nebraska	50	50	10	10	10	10	30	30
Nevada	40	40	†	†	10	10	30	30
New Hampshire	40	40	10	10	10	10	30	30
New Jersey	160	60	20	20	10	10	130	30
New Mexico	50	50	20	20	10	10	20	20
New York	430	110	40	40	40	40	350	30
North Carolina	170	110	60	60	20	20	90	30
North Dakota	30	30	10	10	10	10	10	10
Ohio	280	100	30	30	40	40	210	30
Oklahoma	110	70	20	20	20	20	70	30
Oregon	80	60	20	20	10	10	50	30
Pennsylvania	340	90	20	20	50	50	270	30
Puerto Rico	130	50	10	10	10	10	110	30
Rhode Island	20	20	#	#	#	#	20	20
South Carolina	90	60	20	20	10	10	60	30
South Dakota	30	30	10	10	10	10	20	20
Tennessee	150	80	40	40	10	10	100	30
Texas	390	140	60	60	50	50	280	30
Utah	70	40	#	#	10	10	60	30
Vermont	20	20	#	#	#	#	20	20
Virginia	150	70	20	20	20	20	110	30
Washington	100	70	10	10	40	40	60	30
West Virginia	70	60	10	10	10	10	50	30
Wisconsin	100	60	20	20	20	20	60	30
Wyoming	10	10	10	10	#	#	#	#

† Not applicable.

Rounds to zero.

¹ Institution stratum reflects institution categorization as determined from the 2018–19 Integrated Postsecondary Education Data System (IPEDS) files.² Based on the 2018–19 IPEDS file.

NOTE: For the purposes of this table, Puerto Rico and the District of Columbia are referred to as “states.” Sample sizes rounded to the nearest 10. Detail may not sum to totals because of rounding.

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

Overall, 3,110 institutions were selected, including a census of all public 2-year and all public 4-year institutions and a sample of 1,370 institutions (about 31 percent of eligible institutions) from the “all other sectors” stratum. Within the “all other sectors” stratum, the following criteria were used to determine institution sample sizes:

- in the 17 states with 36 or fewer institutions in the “all other sectors” stratum,¹³ a census of all institutions was selected; and
- in the 35 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 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 stratification¹⁴ was accomplished by sorting the sampling frame by the following classifications:

1. institution control and level;
2. Historically Black Colleges and Universities (HBCUs) status;
3. Hispanic-Serving Institutions (HSIs) status;¹⁵
4. Carnegie classifications of postsecondary institutions; and
5. the institution measure of size.

This implicit stratification helped ensure that the sample was approximately proportional to the population for these measures.

¹³ Although 36 was set as the cutoff for a census, no states with a census of institutions selected in the “all other sectors” stratum had more than 32 institutions.

¹⁴ *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.

¹⁵ 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/edueshsi/definition.html>) and using IPEDS Hispanic enrollment data.

Table 3 shows the counts of sampled and eligible institutions by control and level of institution. Of the 3,110 sampled institutions, 3,070 met the eligibility requirements.

Table 3. Number of sampled and eligible institutions, by control and level of institution: 2019–20

Control and level of institution ¹	Sampled institutions	Eligible institutions
All institutions	3,110	3,070
Control of institution		
Public	1,770	1,760
Private nonprofit	820	810
Private for-profit	510	500
Level of institution		
Less-than-2-year	270	260
2-year	1,150	1,140
4-year, non-doctorate-granting	880	870
4-year, doctorate-granting	810	800
Control and level of institution		
Public less-than-2-year	40	40
Public 2-year	960	960
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	160	150
Public 4-year, non-doctorate-granting, primarily baccalaureate	230	230
Public 4-year, doctorate-granting	390	380
Private nonprofit less-than-4-year	30	30
Private nonprofit 4-year, non-doctorate-granting	390	390
Private nonprofit 4-year, doctorate-granting	390	390
Private for-profit less-than-2-year	230	220
Private for-profit 2-year	160	160
Private for-profit 4-year	120	120

¹ Control and level of institution are based on data from the sampling frame, which was formed from the Integrated Postsecondary Education Data System (IPEDS) 2018–19 Institutional Characteristics Header, 2018–19 Institutional Characteristics, 2017–18 12-month Enrollment, and 2017 Fall Enrollment.

NOTE: Sample sizes rounded to the nearest 10. Detail may not sum to totals because of rounding.

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

2.3 Student Sample

Each sampled institution verified as NPSAS eligible was asked to provide a complete enrollment list of students who satisfied all NPSAS eligibility conditions. However, students in high school or a high school completion program were included on the enrollment lists even though these students are not eligible for NPSAS. Although these students were excluded from sampling, counts of students in high school or a high school completion program were needed later for the weighting poststratification adjustment, as described in chapter 7.

As described above, the NPSAS target population includes students enrolled in the academic year covering July 1, 2019, through June 30, 2020, but since NPSAS:2004, the enrollment lists have contained students enrolled through April 30, excluding students first enrolled in May or June. This allows lists to be collected earlier and data collection to end earlier. The following four assumptions were made to support the decision to exclude students first enrolled in May/June: (1) only a small percentage of students are missed, (2) weighting can account for this minimal lack of coverage, (3) estimates are not substantively affected, and (4) students enrolled for the first time in May/June 2020 have similar characteristics to those enrolled for the first time in May/June 2019. During NPSAS:16 data collection, enrollment list data were analyzed to test these assumptions, using the month first enrolled, ignoring year, provided on the lists. The analysis upheld the four assumptions.

Many institutions know their enrolled students prior to April 30 and provide lists in February, March, or April. However, continuous enrollment institutions, including many of the for-profit institutions, typically cannot provide enrollment lists until mid-May, at the earliest, given that the lists include students enrolled through April 30. This results in students from these institutions having less time in data collection and potentially lower survey response rates than other students. The endpoint of enrollment was changed from April 30 to March 31 so as to receive the enrollment lists for institutions with continuous enrollment earlier, allowing more time for student data collection. Similar to the analysis described above, NPSAS:16 enrollment list data from for-profit institutions were analyzed, using the month first enrolled, ignoring year, provided on the lists. The analysis indicated that representation of the target population is not significantly harmed by excluding students who enroll in continuous enrollment institutions in April for the first time during the academic year. The assumptions mentioned above for excluding students on the list who are first enrolled in May or June are also true for excluding students on the list in continuous enrollment institutions.

The enrollment lists included information needed to identify students for matching to administrative records, classifying students to create sampling strata, computing weight adjustments, and conducting nonresponse bias analysis. 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);

- date of birth (DOB);
- FTB indicator;
- 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.);
- high school/completion program completion date (month and year);
- enrollment in high school/completion program;
- date of first enrollment (at the postsecondary level);
- Classification of Instructional Programs (CIP) code or major;
- contact information (local and permanent street address and phone number and school and home e-mail address);
- veteran status;¹⁶
- ethnicity;
- race; and
- sex.

The 11 student sampling strata were the following:

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

¹⁶ For sampling, veterans were identified by institutions on the student enrollment lists, when available, and students who receive veterans benefits were identified through matching to VBA (see chapter 5 for more information). The majority of the veterans identified during the sampling process are those who receive veterans benefits, and therefore a subset of all students who are veterans.

9. doctoral–research/scholarship/other students in other programs;
10. doctoral–professional practice students; and
11. 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. Because of large numbers of master’s degree and doctoral students in education and business programs, these students were sampled at rates lower than their natural occurrence. Otherwise, their sample size would be too large, and the sample size of other master’s degree and doctoral students would be too small to draw inferences about their experiences. Undergraduate students who were veterans were not oversampled within each state because that would have required too large a total sample size. Table 4 displays the oversampled and undersampled student groups.

Table 4. Oversampled and undersampled student groups: 2019–20

Oversampled student groups	Undersampled student groups
Undergraduate students who are potential first-time beginning students	Master’s degree students in education and business programs
Graduate students who are veterans	Doctoral–research/scholarship/other students in education and business programs
Master’s degree students in STEM programs	
Doctoral–research/scholarship/other students in STEM programs	
Master’s degree students enrolled in for-profit institutions	

NOTE: STEM = science, technology, engineering, and mathematics.

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

NPSAS:20 was designed to sample a total of 400,000 students; 150,000 would be asked to complete a survey, and 250,000 would not be asked to complete a survey. Student records and administrative data were collected for all sampled students. Of the 150,000 students who would be asked to complete a survey, 125,000 were to be undergraduates. As described above, the undergraduate student sample was designed to be both nationally and state representative for public 2-year and public 4-year institutions as well as overall. All 375,000 undergraduate sample students would be included in the state-representative sample. As part of setting the NPSAS:20 sample sizes, a target sample of 55,390 FTBs was determined to provide enough surveyed and confirmed FTBs for the BPS:20 cohort. See section 2.4 for more information about the FTB sample.

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 they were veterans and whether they received federal aid. Within each graduate student stratum, individuals were sorted by whether 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 estimates of undergraduate veterans and financial aid. Greater detail on VBA and NSLDS matching can be found in chapter 5.

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 D). In certain instances, sampling rates were modified as follows:

- Student sampling rates were increased for each institution to yield at least 30 students, if possible, to ensure sufficient yield for variance estimation.
- Student sampling rates were decreased, with exceptions, if an institution sample size was greater than 600 students.¹⁷
- Student sampling rates were adjusted higher or lower based on expected yield calculations for institutions where the sample had not yet been selected.

These adjustments to the initial sampling rates resulted in some additional variability in the student sampling rates and increased survey design effects (variance inflation, see section 7.3). Table 5 shows the expected and achieved¹⁸ undergraduate student sample sizes and percent achieved overall for the full sample, by institution stratum and state.

¹⁷ Because of their large enrollments or few institutions in a state, 44 institutions had a student sample size greater than 600.

¹⁸ Achieved sample sizes are the actual number of students sampled.

Table 5. Number of expected and achieved undergraduate student sample sizes and percentage achieved overall, by institution stratum and state: 2019–20

State	All sectors			Institution stratum ¹								
	Number expected	Number achieved	Percent achieved ²	Public 2-year			Public 4-year			All other sectors		
				Number expected	Number achieved	Percent achieved ²	Number expected	Number achieved	Percent achieved ²	Number expected	Number achieved	Percent achieved ²
All states	375,020	352,740	94.1	151,580	136,100	89.8	135,850	127,890	94.1	87,590	88,760	101.3
Alabama	7,550	7,050	93.5	3,240	3,440	106.2	2,790	2,340	84.0	1,520	1,270	83.5
Alaska	1,930	1,790	92.8	†	†	†	1,420	1,300	91.9	510	490	95.4
Arizona	7,110	7,110	99.9	3,280	2,990	91.1	1,450	1,650	113.3	2,380	2,470	103.8
Arkansas	5,960	5,090	85.4	2,660	2,340	88.2	1,850	1,630	88.3	1,450	1,110	76.7
California	24,000	24,940	103.9	15,050	14,420	95.8	7,470	6,890	92.2	1,470	3,640	246.5
Colorado	5,990	5,770	96.4	1,430	1,410	98.5	2,920	3,090	106.1	1,640	1,270	77.3
Connecticut	6,760	6,130	90.7	2,370	2,840	120.1	1,580	840	53.5	2,810	2,440	86.9
Delaware	2,880	2,460	85.5	†	†	†	1,420	1,630	114.4	1,450	830	57.3
District of Columbia	3,150	3,250	103.2	†	†	†	1,160	1,390	119.6	1,980	1,860	93.5
Florida	12,340	13,770	111.5	3,600	2,840	78.9	6,860	6,870	100.1	1,880	4,060	215.2
Georgia	9,800	10,630	108.5	3,700	3,720	100.6	4,100	4,830	117.7	2,000	2,090	104.3
Hawaii	4,060	4,150	102.3	1,410	1,650	116.9	1,400	1,470	105.1	1,250	1,030	82.5
Idaho	5,670	4,100	72.3	1,630	910	55.7	1,430	800	56.1	2,610	2,390	91.5
Illinois	10,330	10,010	97.0	6,720	6,190	92.2	1,820	2,020	111.3	1,800	1,800	100.2
Indiana	6,420	6,950	108.1	2,610	2,610	100.0	2,280	2,610	114.3	1,530	1,730	112.8
Iowa	5,120	3,850	75.2	2,180	1,550	71.1	1,450	1,140	79.0	1,490	1,160	77.5
Kansas	6,740	4,400	65.3	3,650	2,090	57.2	1,440	1,370	95.1	1,650	940	57.1
Kentucky	5,840	4,860	83.1	2,900	1,810	62.5	1,450	1,390	96.2	1,490	1,650	110.5
Louisiana	6,420	6,130	95.5	2,470	2,290	92.8	2,460	2,530	102.8	1,490	1,310	88.0
Maine	5,230	4,750	90.9	1,390	910	65.8	1,550	1,580	102.1	2,300	2,260	98.4
Maryland	6,640	5,890	88.8	3,130	2,760	88.1	2,180	2,060	94.4	1,330	1,080	81.2
Massachusetts	7,210	7,460	103.5	2,300	2,260	98.2	2,140	2,250	104.9	2,770	2,960	106.8
Michigan	9,350	7,880	84.3	3,970	3,390	85.5	3,750	3,200	85.4	1,640	1,290	79.1
Minnesota	8,760	9,080	103.7	4,630	5,110	110.4	2,390	2,280	95.3	1,730	1,690	97.2
Mississippi	4,410	3,700	84.0	1,620	1,560	96.8	1,440	1,290	89.4	1,360	850	63.0
Missouri	7,210	5,710	79.2	2,450	1,810	73.8	2,770	1,950	70.5	1,990	1,950	98.0
Montana	3,800	3,170	83.4	1,530	1,170	76.2	1,420	1,400	98.9	850	600	70.5

See notes at end of table.

Table 5. Number of expected and achieved undergraduate student sample sizes and percentage achieved overall, by institution stratum and state: 2019–20—Continued

State	All sectors			Institution stratum ¹								
				Public 2-year			Public 4-year			All other sectors		
	Number expected	Number achieved	Percent achieved ²	Number expected	Number achieved	Percent achieved ²	Number expected	Number achieved	Percent achieved ²	Number expected	Number achieved	Percent achieved ²
Nebraska	4,330	3,680	84.8	1,430	940	65.5	1,430	1,060	74.1	1,470	1,680	114.0
Nevada	3,110	3,080	99.0	†	†	†	1,450	1,610	111.0	1,660	1,470	88.5
New Hampshire	4,850	4,130	85.3	1,380	1,040	75.3	1,390	1,460	104.5	2,070	1,630	78.9
New Jersey	7,730	8,990	116.2	3,370	3,360	99.7	2,690	2,710	100.8	1,670	2,910	174.6
New Mexico	4,780	4,350	91.0	2,530	2,190	86.5	1,430	1,400	98.0	820	760	93.0
New York	15,540	16,390	105.5	5,620	5,410	96.3	8,270	8,490	102.7	1,640	2,490	151.1
North Carolina	12,840	12,640	98.4	8,500	8,110	95.5	2,520	2,530	100.1	1,820	1,990	109.5
North Dakota	3,580	1,850	51.5	1,290	410	31.8	1,420	740	52.2	870	690	79.7
Ohio	11,910	11,580	97.2	4,840	4,690	96.9	5,460	5,300	97.1	1,610	1,590	98.5
Oklahoma	8,260	7,160	86.7	3,370	2,930	86.9	3,140	2,850	90.5	1,740	1,390	79.7
Oregon	5,430	4,580	84.2	2,570	1,870	73.0	1,450	1,460	100.6	1,420	1,250	87.8
Pennsylvania	11,340	11,160	98.5	2,450	2,490	101.9	6,770	6,290	93.0	2,120	2,380	112.0
Puerto Rico	5,250	6,560	124.9	980	1,560	159.1	1,790	1,830	102.4	2,480	3,170	127.7
Rhode Island	4,490	3,080	68.5	1,370	1,650	120.0	1,390	#	#	1,730	1,430	82.8
South Carolina	6,980	5,640	80.9	3,190	1,920	60.3	1,950	1,880	96.5	1,840	1,840	100.1
South Dakota	4,100	3,010	73.4	1,270	840	65.8	1,420	1,390	97.8	1,410	780	55.6
Tennessee	8,950	9,130	101.9	5,840	5,880	100.7	1,600	1,520	95.5	1,520	1,720	113.4
Texas	18,970	18,080	95.3	8,980	7,740	86.2	8,660	8,190	94.5	1,330	2,150	161.9
Utah	6,130	5,500	89.8	1,430	600	41.7	1,450	1,390	95.2	3,250	3,520	108.5
Vermont	4,030	3,490	86.5	1,290	1,280	99.1	1,380	760	55.5	1,360	1,440	105.8
Virginia	7,230	6,470	89.5	3,060	2,980	97.5	2,530	2,600	102.8	1,640	890	54.4
Washington	8,520	8,120	95.3	1,430	1,490	104.1	5,720	5,030	87.9	1,370	1,600	116.8
West Virginia	6,350	5,870	92.4	1,400	1,160	83.2	2,220	2,240	101.0	2,740	2,470	90.1
Wisconsin	6,860	5,400	78.8	2,680	2,110	78.6	2,640	2,050	77.9	1,540	1,240	80.8
Wyoming	2,820	2,750	97.3	1,400	1,360	96.9	1,310	1,300	99.6	110	80	75.6

† Not applicable.

Rounds to zero.

¹ Institution stratum reflects institutional categorization as determined from the 2018–19 Integrated Postsecondary Education Data System (IPEDS) files.² Percentage reported reflects the ratio of “achieved” to “expected.”

NOTE: Achieved sample sizes are the actual number of students sampled. 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

After undergraduate students were initially sampled, they were randomly divided into two groups—within student strata and within institution. One group included students who would receive the survey, and the other group would not receive the survey. The proportion of students in each group within strata within an institution was determined such that the overall sample sizes of 125,000 undergraduates for the student survey, including 55,390 FTBs, would be achieved. Both groups had student records and administrative data collected.¹⁹ All sampled graduate students would be asked to complete a survey, in addition to having student records and administrative data collected.

Tables 6 through 8 show expected and achieved sample sizes. Differences in expected and achieved sample sizes can be attributed to the number of students on the enrollment lists varying from the numbers expected or sampling rates being adjusted during sampling. Table 6 shows the expected and achieved student sample sizes and percent achieved for the survey, by student type and control and level of institution. Student survey sampling rates were increased, and the data collection period was extended to help achieve the target of at least 100,000 students responding to the survey. These changes resulted in the achieved student survey sample size being larger than expected.

¹⁹ Thirty institutions did not provide sufficient or accurate data to contact students for the survey, so no students selected from these institutions were surveyed. Administrative data were collected for undergraduate students selected from these institutions. Graduate students selected from these institutions were treated as nonrespondents for weighting, as described in chapter 7.

Table 6. Number of expected and achieved sample students and percentage achieved for the student survey, by student sampling type and control and level of institution: 2019–20

Control and level of institution	All students ¹			Student type ²								
	Number expected ³	Number achieved	Percent achieved ⁴	Potential undergraduate FTB			Other undergraduate			Graduate		
	Number expected ³	Number achieved	Percent achieved ⁴	Number expected ³	Number achieved	Percent achieved ⁴	Number expected ³	Number achieved	Percent achieved ⁴	Number expected ³	Number achieved	Percent achieved ⁴
All institutions	150,000	173,360	115.6	55,390	52,620	95.0	69,610	93,540	134.4	25,000	27,210	108.8
Public less-than-2-year	1,500	1,580	105.4	950	890	93.7	550	690	125.7	†	†	†
Public 2-year	43,320	53,300	123.0	17,840	19,600	109.8	25,480	33,700	132.2	†	10	†
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	8,580	9,300	108.3	4,230	4,360	103.0	4,270	4,870	113.9	80	70	94.9
Public 4-year, non-doctorate-granting, primarily baccalaureate	9,110	10,530	115.7	3,180	2,950	93.0	4,100	5,730	139.7	1,830	1,850	101.0
Public 4-year, doctorate-granting	24,860	37,320	150.1	6,280	7,380	117.5	11,860	20,430	172.3	6,720	9,510	141.5
Private nonprofit less-than-4-year	1,990	1,380	69.2	1,240	690	55.8	750	680	91.5	†	†	†
Private nonprofit 4-year, non-doctorate-granting	10,140	11,990	118.3	2,440	2,780	113.9	5,080	7,030	138.5	2,620	2,180	83.3
Private nonprofit 4-year, doctorate-granting	15,090	19,880	131.8	3,410	3,660	107.1	5,190	7,520	144.9	6,490	8,710	134.2
Private for-profit less-than-2-year	5,610	5,440	97.0	3,050	1,960	64.3	2,550	3,480	136.1	†	†	†
Private for-profit 2-year	8,960	7,750	86.4	5,390	4,240	78.7	3,570	3,510	98.2	†	†	†
Private for-profit 4-year	20,840	14,890	71.4	7,380	4,110	55.7	6,210	5,910	95.2	7,260	4,870	67.1

† Not applicable.

¹ "All students" refers to the students sampled for the student survey.² Some institution classifications of student type on the enrollment lists (e.g., undergraduate or 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.³ Based on sample allocation and Integrated Postsecondary Education Data System (IPEDS) 2017–18 12-month Enrollment counts.⁴ Percentage reported reflects the ratio of "achieved" to "expected."

NOTE: FTB = first-time beginning student. Achieved sample sizes are the actual number of students sampled. 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

Table 7 shows the expected and achieved student sample sizes and percentage achieved overall, by student stratum. The achieved sample size of 380,100 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, and fewer enrollment lists were received than expected. Despite adjusting sampling rates, as described above, the sample size was less than targeted. Overall, more graduate students were selected into the sample than planned because institutions identified more graduate students than expected (for further details about sample allocation, see appendix D).

Table 7. Expected and achieved student sample sizes and percentage achieved overall, by student stratum: 2019–20

Student stratum ²	Student sample size ¹		
	Number expected ³	Number achieved	Percent achieved ⁴
Total	400,020	380,100	95.0
Undergraduate students ⁵	375,020	352,740	94.1
Graduate students	25,000	27,360	109.4
Veterans	3,550	3,260	91.9
Master's degree students in STEM programs	3,180	2,840	89.4
Master's degree students in education or business programs	3,180	4,790	150.3
Master's degree students in other programs	4,160	2,840	68.2
Doctoral–research/scholarship/other students in STEM programs	2,850	2,950	103.6
Doctoral–research/scholarship/other students in education or business programs	2,050	2,460	120.2
Doctoral–research/scholarship/other students in other programs	2,480	2,380	95.7
Doctoral–professional practice students	2,810	3,920	139.7
Other graduate students ⁶	730	1,920	261.8

¹ Student sample size refers to the full sample, which includes the students sampled for the student survey and the students sampled only for administrative data collection.

² Some institution classifications of student type on the enrollment lists (e.g., undergraduate or 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.

³ Based on sample allocation and Integrated Postsecondary Education Data System (IPEDS) 2017–18 12-month Enrollment.

⁴ Percentage reported reflects the ratio of “achieved” to “expected.”

⁵ Undergraduate students include both first-time beginning students and other undergraduate students because they were combined for determining number of expected sample students for the overall sample.

⁶ 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. Achieved sample sizes are the actual number of students sampled. 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

Table 8 shows the expected and achieved student sample sizes and percentage achieved for the survey, by student stratum. The achieved sample size of 173,360 was higher than originally targeted. As described above, differences between the number of expected versus the number of achieved student sample sizes were the result of changes to sampling rates during sampling and differences in expected number of students reported by institutions.

Table 8. Expected and achieved student sample sizes and percentage achieved for the survey, by student stratum: 2019–20

Student stratum ²	Student sample size ¹		
	Number expected ³	Number achieved	Percent achieved ⁴
Total	150,000	173,360	115.6
Undergraduate students	125,000	146,150	116.9
Potential first-time beginning students	55,390	52,620	95.0
Other undergraduates	69,610	93,540	134.4
Graduate students	25,000	27,210	108.8
Veterans	3,550	3,250	91.5
Master's degree students in STEM programs	3,180	2,810	88.4
Master's degree students in education or business programs	3,180	4,780	150.0
Master's degree students in other programs	4,160	2,810	67.5
Doctoral–research/scholarship/other students in STEM programs	2,850	2,950	103.4
Doctoral–research/scholarship/other students in education or business programs	2,050	2,460	120.0
Doctoral–research/scholarship/other students in other programs	2,480	2,360	95.2
Doctoral–professional practice students	2,810	3,890	138.6
Other graduate students ⁵	730	1,900	259.1

¹ Student sample size refers to the students sampled for the student survey.

² Some institution classifications of student type on the enrollment lists (e.g., undergraduate or 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.

³ Based on sample allocation and Integrated Postsecondary Education Data System (IPEDS) 2017–18 12-month Enrollment.

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. Achieved sample sizes are the actual number of students sampled.

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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

Table 9 provides the initial classification of the overall student sample, by student type and control and level of institution. Table 10 provides the initial classification of the student sample for the survey, by student type and control and level of institution.

Table 9. Initial classification of NPSAS:20 student sample overall, by student type and control and level of institution: 2019–20

Control and level of institution	Total sample ¹		Student type ²			
			Undergraduate		Graduate	
	Number	Percent	Number	Percent	Number	Percent
All institutions	380,100	100.0	352,740	100.0	27,360	100.0
Control of institution						
Public	277,470	73.0	266,000	75.4	11,470	41.9
Private nonprofit	69,330	18.2	58,320	16.5	11,010	40.3
Private for-profit	33,300	8.8	28,430	8.1	4,870	17.8
Level of institution						
Less-than-2-year	9,360	2.5	9,360	2.7	†	†
2-year	146,530	38.6	146,520	41.5	10	#
4-year, non-doctorate-granting	78,010	20.5	73,090	20.7	4,920	18.0
4-year, doctorate-granting	146,200	38.5	123,780	35.1	22,430	82.0
Control and level of institution						
Public less-than-2-year	2,010	0.5	2,010	0.6	†	†
Public 2-year	136,110	35.8	136,100	38.6	10	#
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	18,220	4.8	18,140	5.1	70	0.3
Public 4-year, non-doctorate-granting, primarily baccalaureate	23,250	6.1	21,400	6.1	1,850	6.7
Public 4-year, doctorate-granting	97,880	25.8	88,340	25.0	9,540	34.9
Private nonprofit less-than-4-year	1,740	0.5	1,740	0.5	†	†
Private nonprofit 4-year, non-doctorate-granting	27,320	7.2	25,130	7.1	2,190	8.0
Private nonprofit 4-year, doctorate-granting	40,280	10.6	31,450	8.9	8,830	32.3
Private for-profit less-than-2-year	7,280	1.9	7,280	2.1	†	†
Private for-profit 2-year	8,750	2.3	8,750	2.5	†	†
Private for-profit 4-year	17,270	4.5	12,390	3.5	4,870	17.8

† Not applicable.

Rounds to zero.

¹ The student sample was drawn from 2,190 eligible institutions that provided enrollment lists.² Some institution classifications of student type on the enrollment lists (e.g., undergraduate or 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.

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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

Table 10. Initial classification of student sample for the survey, by student type and control and level of institution: 2019–20

Control and level of institution	Total sample ¹		Student type ²					
			Potential undergraduate FTB		Other undergraduate		Graduate	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
All institutions	173,360	100.0	52,620	100.0	93,540	100.0	27,210	100.0
Control of institution								
Public	112,040	64.6	35,180	66.9	65,410	69.9	11,440	42.1
Private nonprofit	33,250	19.2	7,130	13.5	15,240	16.3	10,890	40.0
Private for-profit	28,070	16.2	10,310	19.6	12,890	13.8	4,870	17.9
Level of institution								
Less-than-2-year	7,070	4.1	2,890	5.5	4,180	4.5	†	†
2-year	62,380	36.0	24,500	46.6	37,870	40.5	10	#
4-year, non-doctorate-granting	40,510	23.4	13,310	25.3	22,280	23.8	4,920	18.1
4-year, doctorate-granting	63,410	36.6	11,930	22.7	29,210	31.2	22,270	81.9
Control and level of institution								
Public less-than-2-year	1,580	0.9	890	1.7	690	0.7	†	†
Public 2-year	53,300	30.7	19,600	37.2	33,700	36.0	10	#
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	9,300	5.4	4,360	8.3	4,870	5.2	70	0.3
Public 4-year, non-doctorate-granting, primarily baccalaureate	10,530	6.1	2,950	5.6	5,730	6.1	1,850	6.8
Public 4-year, doctorate-granting	37,320	21.5	7,380	14.0	20,430	21.8	9,510	35.0
Private nonprofit less-than-4-year	1,380	0.8	690	1.3	680	0.7	†	†
Private nonprofit 4-year, non-doctorate-granting	11,990	6.9	2,780	5.3	7,030	7.5	2,180	8.0
Private nonprofit 4-year, doctorate-granting	19,880	11.5	3,660	6.9	7,520	8.0	8,710	32.0
Private for-profit less-than-2-year	5,440	3.1	1,960	3.7	3,480	3.7	†	†
Private for-profit 2-year	7,750	4.5	4,240	8.1	3,510	3.7	†	†
Private for-profit 4-year	14,890	8.6	4,110	7.8	5,910	6.3	4,870	17.9

† Not applicable.

Rounds to zero.

¹ The student sample was drawn from 2,190 eligible institutions that provided enrollment lists.² Some institution classifications of student type on the enrollment lists (e.g., undergraduate or 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.

NOTE: FTB = first-time beginning student. 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

Multiplicity adjustments in the sample weighting (described in more detail in chapter 7) accounted for students who had more than one chance of being selected because they attended multiple institutions during the 2019–20 academic year. To eliminate duplication of students sampled from multiple institutions, the SSNs of those selected from an institution were compared with SSNs of students who had already been selected from other institutions.

2.4 First-Time Beginning Student Sample

NPSAS:20 serves as the base year for the BPS:20 cohort and includes a nationally representative sample of FTBs. Correct classification of FTBs is important given that high rates of misclassification (i.e., false positives and false negatives²⁰) can result in (1) excessive cohort loss with too few eligible sample members to sustain the longitudinal study, (2) excessive cost to “replenish” the sample with little value added, and (3) inefficient sample design (excessive oversampling of “potential” FTBs) to compensate for anticipated misclassification error.

Participating institutions and several administrative data sources provided data to aid in properly classifying FTBs. Key data provided by the institutions included an FTB indicator, degree program, class level, high school graduation date/completion program completion date, enrollment in high school/completion program, and DOB.

Once a list of potential FTBs was created from institution-provided data, administrative data sources, including NSLDS, CPS, and NSC, provided data that were useful in identifying false positives prior to sampling, although these data could not identify with certainty if a student was a false negative. Overall, matching to all sources identified about 20 percent of listed FTBs as false positives.

Prior to sampling, all students listed as potential FTBs were matched to NSLDS records to determine if any had a federal financial aid history predating the NPSAS academic year (i.e., earlier than July 1, 2019). Students with data showing aid disbursements from prior years were excluded from the sampling frame of FTBs. Simultaneously, all potential FTBs were matched to CPS to identify students who indicated on their FAFSA that they had previously attended college. Similarly, students identified as having attended college previously were excluded

²⁰ A false-positive FTB is a student sampled as an FTB but who turns out to not be an FTB, based on the student survey or administrative data. A false-negative FTB is a student not sampled as an FTB but who is confirmed to be an FTB based on the student survey.

from the sampling frame of FTBs. After matching to both NSLDS and CPS, a subset of the potential FTBs was matched to NSC for further narrowing of FTBs based on the presence of evidence of earlier enrollment.²¹

In addition to the steps listed above taken to ensure an adequate BPS cohort, FTB false-positive rates from NPSAS:12 were considered when setting FTB sampling rates.

As shown in table 11, matching to NSLDS identified about 13 percent of cases as false positives, while matching to CPS identified about 11 percent as false positives. CPS also identified many of the false positives identified by NSLDS. Private nonprofit less-than-4-year and private for-profit institutions had a high percentage of false positives.

Of the 890,800 potential FTBs NPSAS staff sent to NSC, about 6 percent were false positives. The NSC matching appeared most effective among 2-year institutions, where 6 percent of FTBs at public 2-year and 7 percent of FTBs at private for-profit 2-year institutions were false positives.

²¹ There is a “charge per case matched” for NSC matching, so a subset of potential FTBs were matched to NSC, including those who were over the age of 18 in the public 2-year and for-profit sectors because these sectors had high false-positive rates in NPSAS:12 and have large NPSAS:20 sample sizes.

Table 11. Potential first-time beginning student false-positive rates, by source and institution type: 2019–20

Control and level of institution	Total			Source								
				NSLDS			CPS			NSC		
	Sent for matching	False positives	Percent false positive	Sent for matching	False positives	Percent false positive	Sent for matching	False positives	Percent false positive	Sent for matching	False positives	Percent false positive
Total	2,232,660	438,400	19.6	2,232,660	281,860	12.6	2,232,660	238,570	10.7	890,800	54,110	6.1
Public less-than-2-year	3,180	600	18.9	3,180	500	15.6	3,180	150	4.8	†	†	†
Public 2-year	803,890	241,870	30.1	803,890	155,300	19.3	803,890	115,310	14.3	803,890	50,120	6.2
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	200,660	39,290	19.6	200,660	28,860	14.4	200,660	24,920	12.4	†	†	†
Public 4-year, non-doctorate-granting, primarily baccalaureate	118,570	11,220	9.5	118,570	5,130	4.3	118,570	9,330	7.9	†	†	†
Public 4-year, doctorate-granting	711,520	67,810	9.5	711,520	31,940	4.5	711,520	52,910	7.4	†	†	†
Private nonprofit less-than-4-year	2,050	680	33.1	2,050	630	30.8	2,050	250	12.3	†	†	†
Private nonprofit 4-year, non-doctorate-granting	98,840	12,460	12.6	98,840	8,490	8.6	98,840	9,170	9.3	†	†	†
Private nonprofit 4-year, doctorate-granting	207,030	22,520	10.9	207,030	16,640	8.0	207,030	9,120	4.4	†	†	†
Private for-profit less-than-2-year	17,760	11,070	62.3	17,760	9,870	55.6	17,760	3,000	16.9	17,760	760	4.3
Private for-profit 2-year	27,090	15,910	58.7	27,090	13,240	48.9	27,090	5,800	21.4	27,090	1,750	6.5
Private for-profit 4-year	42,060	14,970	35.6	42,060	11,270	26.8	42,060	8,610	20.5	42,060	1,480	3.5

† Not applicable.

NOTE: CPS = Central Processing System; NSC = National Student Clearinghouse; NSLDS = National Student Loan Data System. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.

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

When an institution did not identify FTBs, a proxy was used to identify potential FTBs. The proxy was determined using NPSAS:20 data from early enrollment lists and student survey responses, based on a combination of variables: date of first enrollment, degree program, no prior aid receipt, DOB, high school graduation date/completion program completion date, and class level.

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

This chapter describes the design, implementation, and outcomes of institution data collection for NPSAS:20. 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:20 was conducted using several systems created to facilitate and manage institution data collection. The Institution Contacting System (ICS) was used to manage communications with institutions staff. The Postsecondary Data Portal (PDP) website served as the data collection system for all data collected from institutions. The student records collection instrument, part of the PDP, was used to collect student-level data from institutions.

3.1.1 *Institution Contacting System (ICS)*

The ICS was used to track institution participation, maintain a record of all outbound and incoming communications with institutions, and manage the workload of institution contacting staff. ICS reporting functions were used to monitor progress of institution recruitment, enrollment list collection, and student records collection overall and by institution sector.

3.1.2 *Postsecondary Data Portal (PDP)*

NPSAS:20 institution data collection was conducted using the PDP (see figure 1). The PDP provides staff from participating institutions with access to instructional materials and video tutorials on use of the PDP; contact information for the help desk, RTI staff, and NCES staff; frequently asked questions; and a secure platform for providing requested data.

Prior to NPSAS:20 data collection, the PDP was updated to enhance usability and data quality. The task menu was updated to clearly indicate the status of each step in the data request. The “Register Your Institution” page was updated to guide institutions through the process of registering to participate in NPSAS:20. This page was also updated to collect information about institutions’ academic calendars for the period of July 1, 2019, to June 30, 2020—information that was previously collected in the student records collection instrument, described in the next section (3.1.3).

Figure 1. Postsecondary Data Portal home page: 2019–20

PDP POSTSECONDARY DATA PORTAL

NATIONAL CENTER FOR EDUCATION STATISTICS | INSTITUTE OF EDUCATION SCIENCES

HOME ABOUT FAQs CONTACT CONFIDENTIALITY

Get started

USER ID:

User ID

CONTINUE

Forgot ID

NPSAS B&B BPS PETS HSLS-09 QuickStats PowerStats TrendStats

NCES is authorized to conduct the 2019–20 National Postsecondary Student Aid Study (NPSAS:20) by the Education Sciences Reform Act of 2002 (ESRA 2002, 20 U.S.C. §9543) and the Higher Education Opportunity Act of 2008 (HEOA 2008, 20 U.S.C. §1015). The data are being collected for NCES by RTI International, a U.S.-based nonprofit research organization. All of the information you provide may be used only for statistical purposes and may not be disclosed, or used, in identifiable form for any other purpose except as required by law (20 U.S.C. §9573 and 6 U.S.C. §151).

The Family Educational Rights and Privacy Act of 1974 (FERPA, 20 U.S.C. §1232g) allows for the release of institution record information to the Secretary of Education or her agent without prior consent of survey members (34 CFR §§ 99.31[a][9][III] and 99.35).

According to the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number. The valid OMB control number for this voluntary information collection is 1850-0666. The time required to complete this information collection is estimated to average the number of minutes or hours listed below per response, including the time to review instructions, search existing data sources, gather the data needed, and complete and review the information collection. If you have any comments concerning the accuracy of the time estimate, suggestions for improving this information collection, or any comments or concerns regarding the status of your individual submission of these data, please write directly to: Postsecondary Data Portal studies, National Center for Education Statistics, PCP, 550 12th St., SW, Room 4007, Washington, DC 20202.

NPSAS:20 OMB Clearance No: 1850-0666 Expiration Date: 07/31/2022

NPSAS:20 List Collection: 5 hours

NPSAS:20 Student Records Collection: 30 hours

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

3.1.3 Student Records Collection Instrument

The NPSAS:20 student records instrument, part of the PDP website described in section 3.1.2, collected student-level data for the 2019–20 academic year. The

instrument was organized into four sections based on the content of the data request:

1. General Student Information (e.g., demographics, contact information, and other student characteristics);
2. Enrollment (e.g., students' degree program, major[s], class level, and enrollment intensity in the 2019–20 academic year);
3. Budget (i.e., estimated costs of attendance); and
4. Financial Aid (e.g., student financial aid received by sampled students for the 2019–20 academic year, including federal, state, institution, private and other government, and other awards).

To minimize burden on participating institutions, NPSAS:20 student records data elements were largely consistent with the data elements collected for NPSAS:18-AC. All changes made since NPSAS:18-AC were intended to increase quality of the data collected and clarify item definitions for participating institutions. For example, text-entry fields for the tuition and financial aid items were prefilled with the specific institution term(s) for which data were requested. A new item was added to the Budget section to collect whether “full-year” budgets included costs for summer sessions.

During data collection, the student records instrument was revised based on policies enacted in response to the coronavirus pandemic. Wording for multiple data items was revised to provide institutions with reporting guidance for students whose programs of study were disrupted by the coronavirus pandemic, including students' eligibility for NPSAS:20, program or degree, tuition, enrollment status, and financial aid.

Appendix E includes a complete list of data items in each section of the student records instrument.

Institutions were offered three modes to submit student records data:

1. In Excel mode, institution staff downloaded a preformatted Excel spreadsheet template from the PDP, completed the template offline by hand-keying or pasting data, and then uploaded the completed template to the PDP.
2. In comma-separated values (CSV) mode, institution staff downloaded CSV file layout specifications from the PDP, prepared .csv formatted data files offline, and then uploaded the completed files to the PDP.

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

Institutions could choose to use any combination of these modes to provide data and could change modes at any time. A fourth mode (referred to hereafter as “alternate mode”) was offered to institutions on a case-by-case basis. Alternate mode allowed institutions to upload data in any file format and without data validation error checking.

3.2 Institution Contacting, Recruitment, and Student Enrollment List Acquisition

Institution data collection began with contacting sampled institutions to request their participation in the study. Institutions’ chief administrators were asked to confirm or designate a campus coordinator (CC) to act as a primary point of contact for the submission of student enrollment lists and subsequent data collection activities. Institution contacting, recruitment, and student enrollment list acquisition activities are described in the following sections.

3.2.1 *Institution Contacting and Recruitment*

Sampled institutions received their first communication of the data collection in October 2019. Sixteen people were trained as institution contactors to interface with institution staff throughout the study. In contrast to prior data collections, the institution contactors for NPSAS:20 were located remotely throughout the United States rather than in a single call center. These contactors were organized by geographic region, and each was assigned a set of institutions located in that region. Institution contactors’ geographic proximity and local knowledge enabled them to establish deeper relationships with institution staff and to provide enhanced, personalized support throughout data collection.

The initial training for institution contactors was conducted remotely using videoconference technology. Appendix F contains the training agenda for institution contactors. This session included an overview of NPSAS:20, training on data collection tracking tools, guidance on communicating and cultivating relationships with CCs, and instruction on assisting with data collection and submission using the PDP.

Prior to contacting, data collection staff reviewed sampled institutions to ensure they were open and to verify contact information. Institutions that had closed or

merged with other institutions were further reviewed to determine whether they were ineligible for the study.

To encourage participation and reinforce legitimacy of the study, a list of postsecondary organizations and associations that endorsed NPSAS:20 was posted on the PDP (appendix G). Additionally, in the text of each letter to institution staff, one organization with relevance to that institution was called out specifically as having endorsed NPSAS (e.g., the Association of Public and Land-grant Universities and the National Association of Independent Colleges and Universities).

Once institution contactors verified institution eligibility and contact information, a packet of information about the study (appendix H) was sent to the chief administrator of each institution.²² This included a cover letter printed on NCES letterhead providing information about the data collection process for NPSAS:20, instructions for accessing the PDP website, and a request that the chief administrator designate or confirm a CC via the PDP. A brochure was also included that summarized NPSAS:20 background, objectives, and confidentiality information.

Because NPSAS:20 began relatively soon after the conclusion of NPSAS:18-AC and the BPS:12 transcript and student records collections, there were CCs already in place at some institutions that were part of those collections. If a CC could be identified from prior studies or from another legitimate source, the letter named that person as the selected CC, indicated that they would be contacted for the study, and requested that the chief administrator log in to the PDP to change this selection if they preferred to designate a different CC. For institutions without a previously identified CC, the chief administrator was simply asked to log in and designate one.

Institution contactors began follow-up communication to chief administrators' offices soon after the packets prompting CC designation were mailed. If chief administrators were unable or unwilling to log in to the website to designate a coordinator, they could provide that information over the telephone.

After a CC was designated and verified by institution contacting staff, that person received a cover letter describing the study with information on how to access the

²² The chief administrator communication and other official request packets for NPSAS:20 were sent by express shipping. However, due to impacts of the coronavirus pandemic, including closure of fulfillment facilities and closures at many institutions, beginning in mid-March 2020, all communication originally intended to be mailed was instead sent by e-mail.

PDP, a brochure, an overview guide to the NPSAS:20 data collection process, and instructions for completing registration. Institution contactors soon followed up to prompt and assist CCs with completing registration. During registration, CCs entered or confirmed basic details about their institution's terms, including start and end dates or whether they offered continuous enrollment, information that would later be used during student records data collection. Deadlines for submission of the student enrollment lists were generated during registration and were customized based on the institution's term structure and the date of registration.

Some institution systems with common CCs or centralized record keeping were identified using information from prior studies, IPEDS reporting data, or information learned during data collection. For these groups of institutions, CCs were given the option of reporting for constituent institutions individually or at the system level. This served to reduce burden and streamline data collection for these institutions.

3.2.2 Student Enrollment List Acquisition

Upon completion of registration, each CC received a letter formally requesting the student enrollment list along with instructions for preparing the list (see appendix H). Institutions were asked to include on the enrollment list all students enrolled at any time between July 1, 2019, and April 30, 2020.²³ CCs could also access instructions for preparing the student list on the PDP, and institution contactors further supported each CC by telephone and e-mail as necessary, assisting them with the process, answering questions, and reminding them of upcoming deadlines. Contactors included a CC's PDP username and a link to the PDP in e-mail reminders, allowing the CC to easily navigate to the site. After logging in, the CC could view their progress through data collection with a task menu (see figure 2) and identify any actions needed from them.

²³ The NPSAS:20 target population consisted of all eligible students enrolled at any time between July 1, 2019, and June 30, 2020. However, most institutions provided enrollment lists that covered the period of July 1, 2019, through April 30, 2020. The date of April 30 was selected to include virtually all students enrolled before the summer term without delaying data collection. See section 2.3 for more information.

Figure 2. Postsecondary Data Portal task menu: 2019–20

POSTSECONDARY DATA PORTAL TASK MENU

Welcome back! Please complete the steps below.

	Register Your Institution	Complete
	Provide Your Student Enrollment List	Not started
	Provide Student Record Data	Not yet available
	Archive Notes or Other Documentation	Not started

ANNOUNCEMENTS

WHAT TO EXPECT
Review NPSAS:20 Timeline

NPSAS:20 FAQs
Questions About NPSAS:20

CONTACT MATERIALS
Download Mailing Packets

RESOURCES
View Instructions and Video Tutorials

CHANGE MY PASSWORD

MANAGE PDP USERS

CONTACT HELP DESK

1-855-500-1441

PortalHelp@rti.org

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

Most enrollment lists were submitted using the secure upload interface on the PDP. If an institution encountered technical problems during upload, its enrollment list could be sent by e-mail as an encrypted file. To ensure confidentiality, physical lists were not permitted.

Enrollment lists were reviewed for quality and completeness prior to selection of the student sample. These checks included verifying that the format of the list was readable, that key data needed for sampling (e.g., degree program) were provided, and that student counts fell within a predetermined range of the IPEDS counts. Contactors followed up with institution staff to resolve issues that were identified with lists during quality checks.

3.2.3 Institution Recruitment and Student List Acquisition Outcomes

Of the total sample of 3,070 eligible institutions, 71 percent provided student enrollment lists. Thirty institutions did not provide sufficient or accurate data to contact students for the survey (e.g., including too many missing items or obvious

errors such as mismatched e-mail addresses), so no students selected from these institutions were surveyed. These students were, however, included in the administrative sample for student records collection. Graduate students selected from these institutions were treated as nonrespondents for weighting, as described in section 7.1. Institutions that were reluctant to participate in NPSAS:20 but provide enrollment data to NSC were offered the option of sending the files they send to NSC. Twenty institutions provided their NSC data as their enrollment list. In total, 2,190 institutions provided enrollment lists in NPSAS:20 (see table 12).

The percentage of institutions providing enrollment lists across control and level of institution ranged from 41 percent for private for-profit less-than-2-year institutions to 88 percent for public 4-year, doctorate-granting institutions.

Table 12. Number and percentage of institutions providing enrollment lists and number of sampled students overall and for the survey, by control and level of institution: 2019–20

Control and level of institution	Eligible institutions	Institutions providing lists			Sampled students	
		Number	Unweighted percent	Weighted percent ¹	Overall	Student survey
All institutions	3,070	2,190	71.4	63.6	380,100	173,360
Control of institution						
Public	1,760	1,370	78.2	77.3	277,470	112,040
Private nonprofit	810	590	72.0	69.9	69,330	33,250
Private for-profit	500	230	46.4	44.9	33,300	28,070
Level of institution						
Less-than-2-year	260	110	43.1	44.9	9,360	7,070
2-year	1,140	810	70.8	64.3	146,530	62,380
4-year, non-doctorate-granting	870	630	72.0	71.2	78,010	40,510
4-year, doctorate-granting	800	640	80.8	78.4	146,200	63,410
Control and level of institution						
Public less-than-2-year	40	20	55.6	68.5	2,010	1,580
Public 2-year	960	720	75.0	75.2	136,110	53,300
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	150	120	76.6	75.4	18,220	9,300
Public 4-year, non-doctorate-granting, primarily baccalaureate	230	180	80.3	81.4	23,250	10,530
Public 4-year, doctorate-granting	380	340	87.5	87.1	97,880	37,320
Private nonprofit less-than-4-year	30	10	42.4	58.8	1,740	1,380
Private nonprofit 4-year, non-doctorate-granting	390	280	71.7	70.9	27,320	11,990
Private nonprofit 4-year, doctorate-granting	390	290	74.9	73.4	40,280	19,880
Private for-profit less-than-2-year	220	90	41.0	38.9	7,280	5,440
Private for-profit 2-year	160	80	50.6	49.7	8,750	7,750
Private for-profit 4-year	120	60	50.8	61.9	17,270	14,890

¹ The weight described in this column is a base weight.

NOTE: Control and level of institution are based on data from the time of institution sampling. 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

As previously described, enrollment lists were evaluated for accuracy and completeness in several ways, including comparing institution-provided data to IPEDS data for the same institutions. Institution contacting staff contacted institutions that had submitted student counts with discrepancies to resolve the issues. During quality checks, 71 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 2015 Carnegie classification categories (table 13). Of the 2,190 institutions that provided enrollment lists in NPSAS:20, 230 did not have a Carnegie classification. Of those with a known Carnegie classification, the number that provided enrollment lists ranged from a high of 250 institutions classified as Master's Colleges & Universities: Larger Programs to four classification categories with numbers that round to zero.

Table 13. Number and percentage of participating NPSAS:20 institutions, by 2015 Carnegie classification: 2019–20

2015 Carnegie classification	Number	Percent
All institutions	2,190	100.0
Not classified	230	10.6
Associate's Colleges: High Transfer-High Traditional	120	5.7
Associate's Colleges: High Transfer-Mixed Traditional/Nontraditional	100	4.6
Associate's Colleges: High Transfer-High Nontraditional	60	2.8
Associate's Colleges: Mixed Transfer/Career & Technical-High Traditional	80	3.6
Associate's Colleges: Mixed Transfer/Career & Technical-Mixed Traditional/Nontraditional	80	3.4
Associate's Colleges: Mixed Transfer/Career & Technical-High Nontraditional	90	4.2
Associate's Colleges: High Career & Technical-High Traditional	30	1.5
Associate's Colleges: High Career & Technical-Mixed Traditional/Nontraditional	50	2.5
Associate's Colleges: High Career & Technical-High Nontraditional	70	3.3
Special Focus Two-Year: Health Professions	40	1.6
Special Focus Two-Year: Technical Professions	10	0.4
Special Focus Two-Year: Arts & Design	#	0.0
Special Focus Two-Year: Other Fields	#	0.2
Baccalaureate/Associate's Colleges: Associate's Dominant	50	2.3
Doctoral Universities: Highest Research Activity	100	4.7
Doctoral Universities: Higher Research Activity	80	3.8
Doctoral Universities: Moderate Research Activity	70	3.1
Master's Colleges & Universities: Larger Programs	250	11.3
Master's Colleges & Universities: Medium Programs	140	6.2
Master's Colleges & Universities: Small Programs	70	3.2
Baccalaureate Colleges: Arts & Sciences Focus	110	5.2
Baccalaureate Colleges: Diverse Fields	140	6.2
Baccalaureate/Associate's Colleges: Mixed Baccalaureate/Associate's	50	2.1
Special Focus Four-Year: Faith-Related Institutions	30	1.3
Special Focus Four-Year: Medical Schools & Centers	30	1.5
Special Focus Four-Year: Other Health Professions Schools	40	2.0
Special Focus Four-Year: Engineering Schools	#	0.2
Special Focus Four-Year: Other Technology-Related Schools	#	0.0
Special Focus Four-Year: Business & Management Schools	10	0.3
Special Focus Four-Year: Arts, Music & Design Schools	20	0.7
Special Focus Four-Year: Law Schools	10	0.3
Special Focus Four-Year: Other Special Focus Institutions	10	0.3
Tribal Colleges	20	0.9

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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

Table 14 shows the number and percentage of HBCUs and HSIs that provided enrollment lists, also used for implicit stratification, in NPSAS:20. Seventy HBCUs and 380 HSIs provided enrollment lists in NPSAS:20.

Table 14. Number and percentage of participating institutions, by Historically Black Colleges and Universities (HBCUs) and Hispanic-Serving Institutions (HSIs) classification: 2019–20

Institution classification	Number	Percent
All institutions	2,190	100
HBCU	70	3.2
HSI	380	17.5
Not HBCU or HSI	1,740	79.3

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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

3.3 Student Records Data Collection

Once institutions sent student enrollment lists, the student sample was created as detailed in section 2.3, and collection of institution record data for sample members began. These data were used to identify both students and institutions as student records respondents. “Student records student respondents” include eligible sample members who have valid data for key aid items from student records.²⁴ “Student records institution respondents” include institutions whose student records data include at least one student records student respondent. The final count of student respondents included student survey respondents and student records respondents who had at least 1 month of enrollment confirmed in any data source, including administrative data sources detailed in chapter 5. The following sections describe student records collection and outcomes.

3.3.1 Student Records Collection From Institutions

After an institution’s student sample was selected, the CC received information on the student records collection process, including a cover letter and a Student Records Handbook (see appendix H) providing guidance for completing the request, such as detailed instructions for providing data in each of the available modes. The PDP contained a customized list of the institution’s sampled students along with tutorial videos, help text, file upload templates and specifications, and other features and resources to help institutions during the process. Institution

²⁴ Some student records student respondents are not included in the final data file if they did not meet the criteria to be considered a study respondent. See section 7.1.1.2 for more information about study respondents.

contactors and other help-desk project staff were also available to assist if institution staff had questions or encountered problems.

3.3.2 Student Records Collection Outcomes

Of the 2,190 institutions with sampled students, 1,840 (84 percent) provided student records data. Most institutions that provided student records data opted for Excel mode (39 percent), 21 percent uploaded a CSV file, 3 percent used web mode, 2 percent chose an alternate mode, and 20 percent of institutions used two or more modes. Table 15 shows student records data collection results by collection mode and control and level of institution. Student records data were obtained from institutions for 74 percent of eligible sample members (see table 16). This total included 73 percent of the total undergraduate students in the sample and 84 percent of the graduate students.

**Table 15. Enrollment lists and student records data collection results, by data collection mode and control and level of institution:
2019–20**

Control and level of institution	Provided enrollment lists	Provided student records		Data collection mode									
				Web		Excel		CSV		Alternate mode ¹		Two or more modes	
		Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Total	2,190	1,840	84.0	50	2.5	850	38.6	460	21.2	30	1.6	440	20.1
Control of institution													
Public	1,370	1,160	84.1	20	1.6	450	32.8	360	26.1	30	2.3	290	21.3
Private nonprofit	590	500	86.0	20	2.6	300	51.1	80	13.8	#	0.3	110	18.1
Private for-profit	230	180	78.4	20	7.4	100	42.0	20	10.4	#	0.4	40	18.2
Level of institution													
Less-than-2-year	110	90	75.2	10	10.6	50	46.9	#	2.7	#	0.0	20	15.0
2-year	810	640	79.4	20	3.0	270	33.4	180	22.5	10	0.6	160	19.9
4-year non-doctorate-granting	630	540	85.6	10	1.8	280	45.0	110	17.4	20	2.7	120	18.7
4-year doctorate-granting	640	580	89.7	10	1.1	240	37.6	170	26.4	10	1.9	150	22.7

Rounds to zero.

¹ Alternate mode allowed institutions to provide data in a format of their choice. Because there were no specific data requirements or layout specifications in alternate mode, it was enabled only as needed in specific circumstances.

NOTE: CSV = comma-separated values. Control and level of institution are based on data from the time of institution sampling. 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

Table 16. Student records data collection results, by control and level of institution and student type: 2019–20

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	2,190	1,840	84.0	380,100	282,010	74.2
Control of institution						
Public	1,370	1,160	84.1	277,470	208,130	75.0
Private nonprofit	590	500	86.0	69,330	53,380	77.0
Private for-profit	230	180	78.4	33,300	20,500	61.6
Level of institution						
Less-than-2-year	110	90	75.2	9,360	5,290	56.5
2-year	810	640	79.4	146,530	98,120	67.0
4-year, non-doctorate-granting	630	540	85.6	78,010	57,440	73.6
4-year, doctorate-granting	640	580	89.7	146,200	121,170	82.9
Control and level of institution						
Public less-than-2-year	20	20	85.0	2,010	1,680	83.4
Public 2-year	720	570	79.4	136,110	92,890	68.2
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	120	110	89.8	18,220	14,140	77.6
Public 4-year, non-doctorate-granting, primarily baccalaureate	180	160	85.2	23,250	16,760	72.1
Public 4-year, doctorate-granting	340	310	91.3	97,880	82,660	84.4
Private nonprofit less-than-4-year	10	10	92.9	1,740	1,200	69.0
Private nonprofit 4-year, non-doctorate-granting	280	230	83.9	27,320	20,480	75.0
Private nonprofit 4-year, doctorate-granting	290	260	87.7	40,280	31,700	78.7
Private for-profit less-than-2-year	90	70	72.5	7,280	3,550	48.8
Private for-profit 2-year	80	60	77.2	8,750	4,080	46.7
Private for-profit 4-year	60	50	88.5	17,270	12,870	74.5
Student type						
Total undergraduate	†	†	†	351,150	257,720	73.4
Potential FTB	†	†	†	62,230	45,710	73.5
Other undergraduate	†	†	†	288,920	212,010	73.4
Graduate	†	†	†	28,950	24,290	83.9

† Not applicable.

¹ Total sample-eligible students sampled from 2,190 institution enrollment lists.

NOTE: FTB = first-time beginning student. Control and level of institution are based on data from the time of institution sampling. 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

Table 17 shows the percentage of institutions that provided an enrollment list and were student records institution respondents, by state and institution stratum. Institutions that provided an enrollment list and were student records institution respondents were considered participating institutions for the student records data collection. Appendix B contains maps of enrollment list participation rates and student records participation rates for each state overall and by institution stratum.

Table 17. Percentage of institutions participating in the student records data collection, by institution stratum and state: 2019–20

State	All sectors		Institution stratum ¹					
	Number eligible	Percent participated	Public 2-year		Public 4-year		Other sectors	
			Number eligible	Percent participated	Number eligible	Percent participated	Number eligible	Percent participated
All states	3,070	59.9	960	59.6	770	74.2	1,330	52.1
Alabama	70	61.8	20	79.2	10	71.4	30	43.3
Alaska	10	66.7	†	†	#	75.0	10	60.0
Arizona	60	62.7	20	65.0	10	100.0	30	50.0
Arkansas	60	50.0	20	77.3	10	63.6	30	24.1
California	190	52.4	110	49.5	50	62.0	30	46.7
Colorado	60	66.1	10	90.9	20	89.5	30	41.4
Connecticut	50	52.8	10	64.3	10	50.0	30	48.3
Delaware	20	41.2	†	†	#	66.7	10	35.7
District of Columbia	20	57.1	†	†	#	50.0	20	57.9
Florida	100	59.8	30	33.3	40	76.2	30	63.3
Georgia	80	67.5	20	34.8	30	100.0	30	63.3
Hawaii	20	68.2	10	100.0	#	100.0	10	41.7
Idaho	40	45.9	#	25.0	#	75.0	30	44.8
Illinois	90	66.7	50	60.4	10	91.7	30	66.7
Indiana	50	63.0	#	100.0	20	53.3	30	66.7
Iowa	50	61.2	20	68.8	#	100.0	30	53.3
Kansas	60	44.4	30	20.0	10	75.0	30	56.7
Kentucky	50	68.5	20	56.3	10	75.0	30	73.3
Louisiana	60	54.2	10	42.9	20	64.7	30	53.6
Maine	40	68.6	10	57.1	10	62.5	20	75.0
Maryland	60	54.2	20	43.8	10	92.3	30	43.3
Massachusetts	60	76.3	20	81.3	10	100.0	30	62.1
Michigan	80	61.3	20	70.8	20	81.8	30	37.9
Minnesota	70	78.4	30	93.8	10	100.0	30	53.3
Mississippi	50	40.4	20	66.7	10	62.5	30	20.7
Missouri	60	45.0	20	35.3	10	53.8	30	46.7
Montana	30	40.0	10	40.0	10	42.9	10	38.5

See notes at end of table.

Table 17. Percentage of institutions participating in the student records data collection, by institution stratum and state: 2019–20
—Continued

State	All sectors		Institution stratum ¹					
	Number eligible	Percent participated	Public 2-year		Public 4-year		Other sectors	
			Number eligible	Percent participated	Number eligible	Percent participated	Number eligible	Percent participated
Nebraska	40	69.8	10	88.9	10	42.9	30	70.4
Nevada	40	47.4	†	†	10	100.0	30	35.5
New Hampshire	40	43.2	10	14.3	10	83.3	20	41.7
New Jersey	60	62.9	20	63.2	10	84.6	30	53.3
New Mexico	50	63.0	20	78.9	10	55.6	20	50.0
New York	110	62.7	40	51.4	40	72.1	30	63.3
North Carolina	110	76.2	60	77.6	20	76.5	30	73.3
North Dakota	30	28.6	10	20.0	10	33.3	10	28.6
Ohio	100	80.2	30	83.3	40	88.9	30	66.7
Oklahoma	70	55.1	20	54.5	20	70.6	30	46.7
Oregon	60	51.8	20	52.9	10	66.7	30	46.7
Pennsylvania	90	72.0	20	66.7	50	86.7	30	53.3
Puerto Rico	50	61.2	10	0.0	10	71.4	30	66.7
Rhode Island	20	33.3	#	0.0	#	0.0	20	38.9
South Carolina	60	58.7	20	25.0	10	84.6	30	70.0
South Dakota	30	53.8	10	60.0	10	71.4	10	42.9
Tennessee	80	68.4	40	76.9	10	70.0	30	56.7
Texas	140	67.9	60	65.0	50	79.6	30	53.6
Utah	40	55.0	#	33.3	10	71.4	30	53.3
Vermont	20	52.4	#	100.0	#	50.0	10	50.0
Virginia	70	56.3	20	58.3	20	76.5	30	43.3
Washington	70	56.2	10	42.9	40	44.4	30	73.3
West Virginia	60	41.8	10	41.7	10	76.9	30	26.7
Wisconsin	60	55.7	20	56.3	20	66.7	30	50.0
Wyoming	10	60.0	10	57.1	#	100.0	#	50.0

† Not applicable.

Rounds to zero.

¹ Institution stratum reflects institutional categorization as determined from the 2018–19 Integrated Postsecondary Education Data System (IPEDS) files.

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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

3.4 Institution Data Evaluation

Enrollment lists and student records data were reviewed for quality and completeness continually throughout institution data collection. Automated data quality programs assessed the missingness of critical data elements, invalid or out-of-range responses, and whether submissions included sample members' PII. The automated data review programs produced reports summarizing the quality of the submitted data and a pass or fail result. For submissions with a fail result, the data were reviewed and corrected whenever possible. Institution contactors followed up with the institution about any data problems that could not be resolved. Institutions would then be asked to provide updated data, if necessary, or indicate that the data should be used as submitted. The most common data problems that institutions were contacted for were missing critical data elements.

Table 18 shows completeness of key student records data elements by data collection mode—Excel, CSV, web, alternate mode, or two or more modes—for the 282,010 student records student respondents. Overall, these data elements have high rates of completeness. Critical data elements for determining students' eligibility for financial aid, such as degree program and class level (99.7 percent and 99.9 percent, respectively), were among the items with the highest completion levels. Items with the highest completion rates also include ethnicity (99.9 percent) and sex (99.3 percent). Of the key data elements listed in table 18, the items with the lowest completion rates were citizenship status (70.1 percent) and race (79 percent).

The data elements requested for NPSAS:20 were available at a large majority of institutions. Variation in item-level completion rates typically arose from differences in institutions' student data management systems, which varied in data format, level of specificity, and difficulty of retrieving the requested data.

Table 18. Student records item-level completion rates, by data collection mode and data element: 2019–20

Data element	Total		Data collection mode									
			Web		CSV		Excel		Alternate mode ¹		Two or more modes	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Total student records student respondents	282,010	100.0	2,990	1.1	101,910	36.2	102,940	35.8	5,900	2.4	68,270	24.6
Student characteristics												
Sex	280,080	99.3	2,990	100.0	101,390	99.5	102,229	99.3	5,310	89.9	68,160	99.8
Marital status	248,500	88.1	2,980	99.7	98,950	97.1	88,620	86.1	5,120	86.7	52,820	77.4
Citizenship	201,960	70.1	2,650	88.5	80,530	79.0	66,800	64.9	3,890	66.0	48,090	70.4
High school completion type	253,960	89.9	2,980	99.7	98,060	96.2	85,040	82.6	5,110	86.5	62,780	92.0
Race	224,770	79.1	1,810	60.6	81,092	79.6	81,920	79.6	5,150	87.1	54,800	80.3
Ethnicity	281,970	100.0	2,990	100.0	101,910	100.0	102,910	100.0	5,900	100.0	68,250	100.0
Enrollment												
Degree program	281,190	99.7	2,990	100.0	101,340	99.4	102,730	99.8	5,900	100.0	68,230	99.9
Student class level	281,920	99.9	2,990	100.0	101,890	100.0	102,870	99.9	5,900	100.0	68,270	100.0
Residency for tuition purposes	263,050	92.6	2,990	99.9	99,520	97.7	91,000	88.4	5,700	96.5	63,850	93.5
Total tuition and fees charged	257,920	90.8	2,990	99.9	98,760	96.9	89,290	86.7	4,630	78.5	62,260	91.2
Budget												
Tuition and fees	262,890	92.5	2,990	99.9	95,420	93.6	95,670	93.0	4,250	72.0	64,540	94.5
Financial aid												
Any aid received ²	278,020	97.5	2,990	100.0	101,600	99.7	100,960	98.1	5,900	100.0	66,560	97.5

¹ Alternate mode allowed institutions to provide data in a format of their choice. Because there were no specific data requirements or layout specifications in alternate mode, it was enabled only as needed in specific circumstances.

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 nonmissing responses, including responses of “Unknown,” are counted as complete. 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

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

This chapter provides an overview of the NPSAS:20 survey design, data collection, outcomes, and evaluation.

4.1 Student Survey Design and Systems

The NPSAS:20 full-scale student survey collected information on students enrolled in postsecondary education during the 2019–20 academic year, with a special focus on how they financed their education. The survey was administered between March 2020 and January 2021 and included many data elements from prior NPSAS administrations, such as enrollment and employment history, to allow for trend comparisons among cohorts. The NPSAS:20 student survey also serves as the base-year data collection for the BPS cohort. Data elements pertaining to FTB experiences, such as persistence, attainment, and social and academic activities, were also included in the full-scale survey. This section provides an overview of the NPSAS:20 full-scale survey design process, an introduction to survey features and functionalities, and descriptions of the data elements included in each key content area of the survey.

4.1.1 Student Survey Design

The data elements included in the NPSAS:20 survey were determined with consideration of past NPSAS and BPS studies, while also incorporating new and modified data elements. Common NPSAS items related to enrollment history, financial aid sources, employment history, income and expenses, and demographic characteristics were retained. Revisions to previous data elements

and the inclusion of new data elements were informed by pretesting results and TRP member input.²⁵ For a list of TRP members, see appendix I.

Historically, NPSAS has collected data needed to determine dependency status in the academic year of interest. However, in prior NPSAS administrations, the survey content measuring dependency differed slightly from the FAFSA.²⁶ To improve data quality across sources, the NPSAS:20 survey was restructured to group all survey items used to determine dependency status into one new section that aligned directly with the 2019–20 FAFSA and served to create FAFSA-comparable data for respondents who did not apply for federal aid.

Data elements pertaining to food security and an additional indicator of homelessness were also added to the NPSAS:20 survey. These items were identified and developed with input from the U.S. Department of Agriculture (USDA) and the Hope Center.²⁷ Additionally, as a response to the worldwide coronavirus pandemic, several items capturing respondent experiences and institution response to the coronavirus pandemic were added to the survey after data collection began.

Per TRP recommendations, new FTB-only data elements were added to enhance measures of student knowledge and perceptions of the costs and benefits of education to support analysis of student choices, persistence, and attainment. These data elements included perceptions of teacher effectiveness, as well as frequency and types of communication and interactions with family and friends, faculty members, and other students.

The survey consisted of seven key content areas: enrollment, FAFSA, education experiences, financial aid, employment, income and expenses, and background. For a complete list of data elements in each key content area, see appendix C. A brief overview of each area follows.

²⁵ Readers may refer to the pretesting components available from the OMB attachments: https://www.reginfo.gov/public/do/PRAViewIC?ref_nbr=201704-1850-007&icID=234606.

Results from the pretesting are available in appendix G of the OMB documentation: https://www.reginfo.gov/public/do/PRAViewDocument?ref_nbr=202011-1850-004.

²⁶ Per the FAFSA, a student who does not meet any of the criteria for an independent student is determined to be dependent. An independent student is one of the following: at least 24 years old, married, a graduate or professional student, a veteran, a member of the armed forces, an orphan, a ward of the court, someone with legal dependents other than a spouse, an emancipated minor, or someone who is homeless or at risk of becoming homeless.

²⁷ For more information about the Hope Center for College, Community, and Justice at Temple University, see <https://hope4college.com/>.

Enrollment confirmed eligibility to participate in NPSAS and identified members of the BPS cohort. The student survey collected extensive information on enrollment at the sampled institution (referred to hereafter as the NPSAS institution) as well as any additional postsecondary institutions attended during the 2019–20 academic year, expected degree completion, and future enrollment.

The **FAFSA** section collected information that aligned with the 2019–20 FAFSA such as DOB; marital status; respondent income; military status; information about dependent children and other dependents; and parent information including parental income, marital status (current, and as of age 18), education, and occupation. For respondents with a spouse, information on spousal income and education level was collected. Additionally, the section collected indicators of homelessness, siblings who attended college first, and receipt of untaxed benefits.

Education experiences collected information on high school experiences including grade point average (GPA); Advanced Placement, International Baccalaureate, and college-level coursetaking; taking of the SAT and ACT exams; and mathematics courses taken in high school. The section also captured information on postsecondary experiences such as major, GPA, remedial education, online coursework, and experiences studying abroad. Finally, the BPS cohort was administered a series of questions regarding academic and social experiences of college students. These items included interactions with family and friends, faculty, and other students; social and academic groups and clubs participated in; services used and the importance of those services; transfer plans; and expected occupation and wages.

Financial aid captured information about sources of aid used in the 2019–20 academic year, such as loans, grants or scholarships, employer assistance, veterans benefits, etc. NPSAS collects information about student borrowing behavior including information on the amounts borrowed in private loans for undergraduate and/or graduate education in the 2019–20 academic year and the total amount borrowed in private loans for undergraduate and/or graduate education. For students who had assistantships, fellowships, or traineeships, information about award amounts and type of placement was collected. Additionally, information about awareness and use of income-driven repayment plans and loan forgiveness programs and information about emergency aid were collected.

Employment items collected information about all work for pay during the 2019–20 academic year. This included assistantships, traineeships, work-study, fellowships with a work component, full-time and part-time employment, paid internships, and

self-employment. Information about school jobs (i.e., assistantships, traineeships, work-study, and fellowships with a work component) was collected first, including dates employed, hours worked, amount earned, whether the job was on or off campus, and if the job was related to the respondent's major. Once all information on all school jobs was collected, respondents reported on all other employment in the 2019–20 academic year. For the BPS cohort, additional questions about employment, such as characteristics of jobs while enrolled and potential alternative activities to enrollment in postsecondary education, were included in this section.

The **income and expenses** section of the NPSAS:20 survey collected information including monthly day care costs for respondents who had dependent children, financial support for education and living expenses from family and friends, number of credit cards and balances, residence while attending school, mortgage and rent amount for those who did not live on campus, and measures of broad financial concept knowledge.

Background items obtained information about student demographic characteristics such as U.S. citizenship, immigration status, race, ethnicity, sex, sexual orientation, and gender identity. An additional homelessness indicator and items from the USDA 10-item food security module were also included in the background section of the NPSAS:20 survey. Finally, during data collection, items pertaining to the national postsecondary education changes in response to the coronavirus pandemic were added to the survey.²⁸ Information collected included enrollment and employment disruptions, food and housing experiences, and resources and information provided by the NPSAS institution.

4.1.1.1 *Survey mode administration*

The NPSAS:20 survey instrument was a multimodal survey designed for web and telephone administration. Self-administered web surveys could be completed on a mobile device (e.g., smartphone or tablet), referred to hereafter as “web mobile” mode, and on a nonmobile device, referred to hereafter as “web nonmobile” mode.²⁹ Respondents advanced through the survey according to skip logic based on information provided by administrative records or earlier in the NPSAS:20 survey. The survey consisted of forms, which are survey screens that typically contained a single question, but sometimes multiple questions, organized by content area; form-specific help text; the response options to each question; and

²⁸ Refer to the COVID-19 First Look publication for more information about how the coronavirus pandemic impacted postsecondary students at <https://nces.ed.gov/pubs2021/2021456.pdf>.

²⁹ Two elements of instrument paradata identified web mobile mode: the rendering of the instrument (e.g., how the instrument displayed on a browser for a respondent) and parsing the browser agent string to obtain information about device type, browser, and touchscreen capability.

navigation buttons. To minimize mode effects, specific methodological features were programmed to provide web nonmobile and mobile respondents with the assistance normally provided by a trained telephone interviewer:

- help text on every form to define key terms and clarify question intent;
- pop-up messages to correct responses that were out of range or in an incorrect format;
- conversion text to encourage responses to critical questions when left unanswered; and
- pop-up messages prompting sample members to provide a response when they left three consecutive forms blank.

Help text was provided at the form level and was written specifically for the question(s) on screen. Section 4.5.2 provides analysis of help text usage in the survey. To reduce nonresponse on critical items (e.g., annual income), conversion text appeared when a respondent did not provide a response. The conversion text was written to communicate the importance of the form and encouraged the respondent to reconsider providing a response. For additional information on conversion text, see section 4.5.3. The instrument also included coder forms, which were custom forms in the survey that were linked to an underlying database of standardized responses to the question on the form. For more information on coders included in the NPSAS:20 survey, see section 4.1.1.2.

In addition to these features, for the first time, the entire NPSAS:20 survey was translated into Spanish.³⁰ The survey initially launched using the browser's preferred language and featured a toggle button that allowed respondents or telephone interviewers (when warranted) to switch between English and Spanish on every form.

4.1.1.2 Coding systems

The NPSAS:20 full-scale survey used coding systems, or “coders”—custom forms that enabled the conversion of respondent text entries into standardized responses in real time during the survey. Predictive text string coders were used for the following survey items: high school attended, postsecondary institutions attended, major/field of study, occupations, and country for respondents who studied abroad. For each coder, respondents entered their response as a text string into a text box. As respondents typed, a keyword search of the underlying database returned a list of possible matches for selection, displayed in a drop-

³⁰ The Spanish NPSAS:20 student survey is available upon request by contacting nces.info@ed.gov.

down menu. For respondents completing the student survey in Spanish, the coders were converted to a simple text-box entry form and the coder functionality was disabled since not all underlying databases were available in Spanish.³¹

See section 6.2.2.2 for an explanation of how codes were assigned to text strings that were not successfully coded in the instrument. The following are brief descriptions of the individual coding systems, sometimes used on multiple forms, and the underlying databases for each coding system:

- The **high school coder** contained all schools from the Private School Universe Survey for private high schools (<https://nces.ed.gov/surveys/pss>), and the Common Core of Data for public high schools (<https://nces.ed.gov/ccd>). For schools not identified within the high school coder, the coder retained the entered school name and asked respondents to supply the school type, district or county name, and the highest and lowest grade levels at the school.
- The **postsecondary institution coder** was linked to the complete set of institutions contained in the 2017–18 IPEDS, developed by NCES (<https://nces.ed.gov/ipeds>). Any postsecondary institution other than the NPSAS institution attended in the 2019–20 academic year was coded using this coder. For institutions not listed in the database, the coder saved any entered institution name and prompted respondents to provide the control (e.g., public, private nonprofit, or private for-profit) and level (e.g., 4-year or 2-year) of the institution, as well as the city and state in which the institution was located.
- The **major/field of study coder** used the 2020 CIP taxonomy, developed by NCES (<https://nces.ed.gov/ipeds/cipcode>). For any major or field of study not found in the CIP database, the coder saved the entered major and asked respondents to select a general area of study and a specific discipline within that area.
- The **study abroad coder** used the 2014 International Organization for Standardization 3166 Country Codes System (https://www.iso.org/iso/home/standards/country_codes.htm).
- The **occupation coder** linked respondents' occupation titles to version 24.0 of the Occupational Information Network (O*NET) database (<https://onetonline.org>), which uses the 2018 Standard Occupational Classification taxonomy (<https://www.bls.gov/soc/home.htm>). For any

³¹ The Spanish text strings were collected for each coder and upcoded by project staff at a later time.

occupations not listed in the database, the coder saved the entered occupation and asked respondents to provide a general area, specific area, and finally a detailed classification for the occupations.

4.1.2 Data Collection Systems

A study as complex as NPSAS requires a robust set of systems to address the unique features of the project. These tailored data collection systems were used to develop the survey instrument, contact sample members (including phone, e-mail, mail, and text reminders), report data collection progress, and evaluate interviewer performance.

The NPSAS:20 survey instrument was developed using a proprietary web-based system called Hatteras. Hatteras allows multiple users to interact with the survey during specification, review, and testing. All instrument design specifications were stored in a Structured Query Language database via the survey editor interface. When published to the web server, the survey forms were dynamically rendered, reflecting the specified content of each form, question routing, valid ranges, and if applicable, any previous responses entered by the respondent. The survey's appearance was automatically adjusted to fit the screen size of the respondent's computer, mobile phone, or other device. Both self-administered web surveys (nonmobile and mobile) and telephone interviews used the same Hatteras survey instrument to collect data.

The proprietary case management system used by telephone interviewers, the computer-assisted telephone interviewing Case Management System (CATI-CMS), managed all sample member locating information and all activity related to outbound and inbound calls. Any contact updates, including new telephone numbers, were added to CATI-CMS as they were identified via interviewer contacts, batch tracing services, or intensive tracing methods. The CATI-CMS worked seamlessly beside the tracing system to account for each case, whether it was ready for outbound calls or required additional tracing efforts (see section 4.2.3 for more information on tracing operations).

The CATI-CMS was used to assign cases to interviewers by prior contact status (e.g., cases that had been recently contacted or had never been contacted), best day and time to call, and previously scheduled appointments. Sample members who had previously refused to participate were placed into a separate queue to be contacted by telephone interviewers specifically trained in refusal conversion techniques designed to encourage sample members to complete the survey after a refusal. The system also automatically ranked cases to call, prioritizing sample

members most likely to respond. Additionally, NPSAS:20 staff members prioritized other cases, as needed, that could be accessed separately by a subset of telephone interviewers. For example, sample members who were eligible for the shortened version of the survey (see section 4.2.5) could be assigned to this priority dialing queue shortly after becoming eligible for the shorter survey to maximize our contact efforts and encourage participation.

Project staff used a proprietary suite of data collection applications called Symphony to coordinate e-mail and text message reminders to sample members. The system allowed staff to deliver personalized reminders by incorporating the latest contacting and survey progress updates. Data collection activities were monitored via real-time reports as well as daily reporting of survey completion, response, timing, and trend analysis. Project staff used a proprietary Adaptive Total Design dashboard to check quality indicators from a variety of data sources during survey data collection. This included response rates, case dispositions, completes by wave and mode, and interviewer performance reports, allowing for prompt decisionmaking during data collection. Additionally, staff used Tableau, a visual analytics platform to create need-specific dashboards, charts, and worksheets for analyzing specific data collection performance, such as response by contacting mode (mail, e-mail, text, and phone) and tracing activities.

A proprietary Quality Evaluation System (qUEST) facilitated interviewer performance monitoring. Protocols for evaluating interviewer performance were used in real time while an interview was being performed or through recordings after the survey had been administered. QUEST supported all phases of telephone-interviewer quality monitoring including selecting interviews, observing interviewers' work, evaluating interviewer performance, providing feedback, and analyzing performance data across interviewers to identify cross-cutting instrument or performance issues.

4.2 Student Survey Data Collection

The primary mode for NPSAS:20 student data collection was a web survey available through the study website. Sample members also had the option of completing the survey with an interviewer trained in computer-assisted telephone interviewing (CATI) methods. Regardless of completion mode, effective data collection required several processes and systems. The NPSAS:20 data collection study website and help desk offered information and support to sample members. Before and during data collection, sample members were located, using batch and intensive tracing processes. Interviewers completed extensive training on

interviewing processes and protocols. Additionally, staff were trained on locating, tracing, and contacting procedures to ensure efficiency and consistency.

4.2.1 Student Survey Data Collection Design

Calibration sample. In lieu of a field test study, project staff conducted limited testing with a subset of the full-scale student sample (a calibration sample). The goal was to inform the NPSAS:20 full-scale design regarding incentive structure and nonresponse follow-up strategies. Table 19 provides specifications of the design, which included two phases and three incentive experimental conditions. Phase 1 of the design tested the baseline incentive (i.e., a combination of prepaid and promised), while phase 2 tested the nonresponse follow-up incentive (i.e., a combination of prepaid and promised vs. double promised incentive vs. control condition).

Table 19. Calibration sample design, by experimental condition and data collection period: 2019–20

Data collection period	Group 1 <i>n</i> = 2,030	Group 2 <i>n</i> = 2,030	Group 3 (Control) <i>n</i> = 2,030
Phase 1: Initial offer	\$2 prepaid + \$30 promised	\$2 prepaid + \$15 promised	\$0 prepaid + \$30 promised
Phase 2: Nonresponse follow-up	\$10 prepaid (via PayPal or check) + \$20 promised	\$30 promised	\$30 promised

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

A calibration sample of 6,200 students from 60 institutions was selected from fall enrollment lists in December 2019 and January 2020. The experimental period for the calibration sample was 10 weeks, with phase 1 taking the first 8 weeks of data collection that began March 2020. Because the calibration sample was part of the main sample and not treated as a field test sample, phase 2 for the calibration cases did not stop after the first 2 weeks of nonresponse follow-up but continued until phase capacity (i.e., the best outcome that a set of design features in a particular phase can produce) was reached based on stability of the cumulative completion rate, increase in refusal and noncontact rates, and CATI hours per complete. However, analyses of the experimental data were based only on the first 10 weeks of data collection for those cases.

Analyses of the calibration data included response rate and sample representativeness comparisons. Table 20 presents the cumulative calibration response rates by experimental condition and phase.

Table 20. Cumulative calibration response rates, by experimental condition and data collection period: 2019–20

Data collection period	Group 1 <i>n</i> = 2,030	Group 2 <i>n</i> = 2,030	Group 3 (Control) <i>n</i> = 2,030
Phase 1: Initial offer ¹	54.5%	44.7%	53.7%
Phase 2: Nonresponse follow-up ²	60.3%	54.4%	57.9%

¹ For phase 1, the baseline incentive was offered for the first 8 weeks.

² For phase 2, the nonresponse follow-up incentives were offered after 8 weeks.

NOTE: Results exclude ineligible cases. Partial interviews are considered nonrespondents for analytic purposes. Sample sizes rounded to the nearest 10.

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

Comparison of Group 1 and Group 3 phase 1 response rates (American Association for Public Opinion Research RR1³² = 54.5 percent and 53.7 percent, respectively) allowed assessment of offering a \$2 *prepaid incentive*. A two-tailed *z* test yielded no statistically significant differences in response rates between the two groups at the end of phase 1 ($z = -0.52, p = .60$). A possible explanation for this unexpected finding is that the start of the NPSAS:20 calibration data collection coincided with the start of the coronavirus pandemic; many institutions closed their campuses shortly after the mailing of the initial survey invitation containing the \$2 prepaid incentive. As a result of students moving off campus and finding alternative housing, mailings may not have reached them on time.

As a direct test of the initial promised incentive amount (\$30 vs. \$15), Group 1 and Group 2 response rates were compared during phase 1 (i.e., end of week 8). Group 1 yielded a significantly higher response rate (54.5 percent) than Group 2 (44.7 percent), based on a two-tailed *z* test ($z = 6.25, p < .001$), suggesting that front-loading the promised incentive might be the preferred approach for shorter data collections.

Comparison of the overall phase 1 and 2 response rates (i.e., at the end of week 10) across groups showed significantly higher response rates for Groups 1 and 3 (60.3 percent [$z = 3.82, p < .001$] and 57.9 percent [$z = 2.25, p < .05$], respectively) relative to Group 2 (54.4 percent), despite doubling the Group 2 promised incentive. This finding further supports the strategy of front-loading the incentive. Further, there was no statistically significant difference in phase 2 response rates between Groups 1 and 3, despite the introduction of a \$10 prepaid incentive in Group 1 ($z = 1.57, p = .12$). Due to the coronavirus pandemic and

³² Unless noted otherwise, all response rates reported refer to the response rate 1 (RR1) as defined by the standards of the American Association for Public Opinion Research (2016). The RR1 is the number of complete interviews (excluding partial interviews) divided by the number of complete and partial interviews plus all noninterviews (excluding confirmed ineligible).

increased student mobility at the time, the \$10 prepaid incentive was limited to PayPal and may explain the nonsignificant results.

Comparison of key demographic characteristics (e.g., age, gender, race, and ethnicity) across groups also did not yield statistically significant differences, suggesting no differential nonresponse bias across the three experimental conditions.

Main data collection design. Given the lack of evidence for an advantage using the \$2 prepaid incentive, and no statistically significant differences between experimental Groups 1 and 3, the incentive design for Group 3 (i.e., \$30 promised incentive from the start) was implemented for the NPSAS:20 main data collection.

The NPSAS:20 data collection involved a four-phase design: (1) initial survey offer, informed by the calibration sample; (2) nonresponse follow-up, informed by the calibration sample; (3) abbreviated 15-minute survey; and (4) 5-minute survey as a last effort to gather information on nonrespondents. The duration of each phase varied across sample release waves and was determined based on phase capacity (monitored via various dashboard indicators for marginal returns). As in previous NPSAS data collections, sample members were offered an abbreviated survey once phase 2 capacity was reached. The abbreviated survey took 21 minutes on average and offered the same promised incentive as phase 2. It included questions from the enrollment section (e.g., NPSAS and BPS eligibility, NPSAS degree type, and attendance dates), FAFSA section (e.g., homelessness, age, marital status, military status, income, number of dependents, and parents' income and education), education experiences section (e.g., academic and social experiences), financial aid section (e.g., types of aid used in 2019–20 academic year), employment section (e.g., employer name, hours worked, and pay), and background section (e.g., demographics, food insecurity, and COVID-19 experiences). Response rates by phase of data collection, institution level, and student type are presented in section 4.3.2.

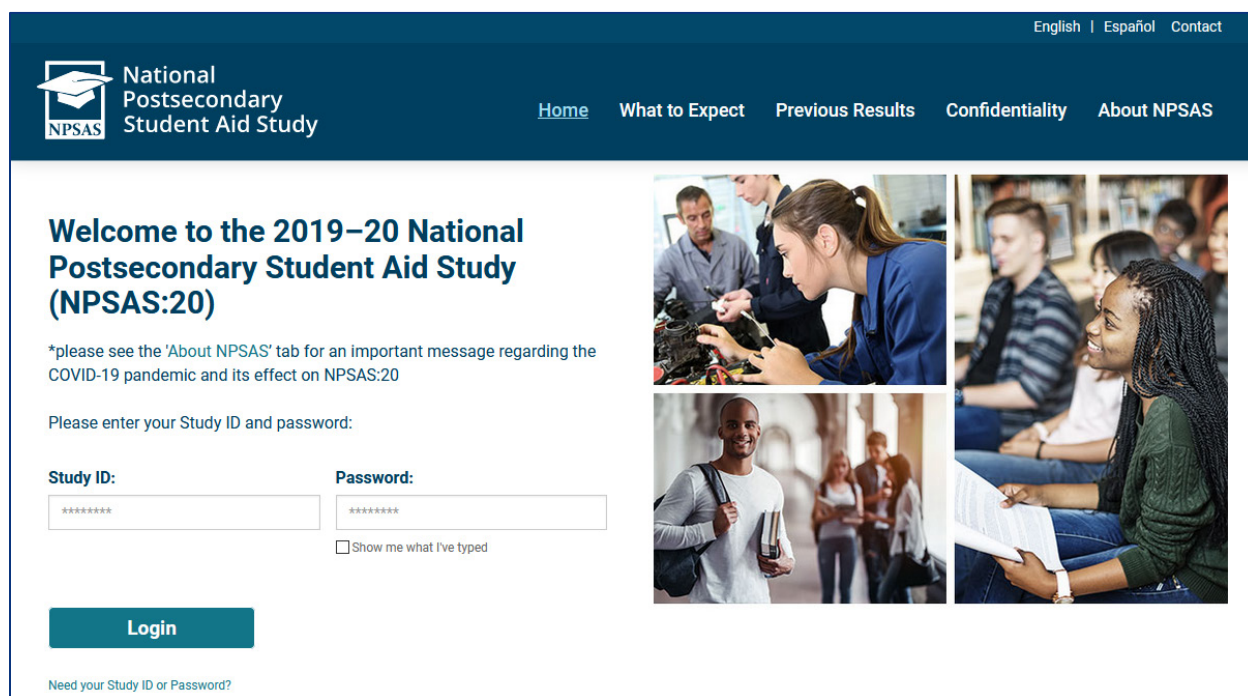
The purpose of phase 4 was to assess nonresponse bias and, depending on response rate, develop a nonresponse weighting adjustment. Phase 4 data collection began approximately a month after the main data collection ended (in late February 2021) and lasted through late March 2021. Phase 4 completes were not counted as respondents in the response rate calculations (phase 4 cases are not included in the final data files) but only used for nonresponse bias assessment. The 5-minute survey was offered to all remaining nonrespondents, with a promised incentive of \$5. The survey included a subset of items from the

abbreviated survey, mostly questions from the FAFSA section. Of the remaining 65,620 nonrespondents at the end of NPSAS:20 data collection, 890 completed the 5-minute survey.

4.2.2 Study Website and Help Desk

Communications with sample members included a link to the NPSAS:20 website, which provided general information about the study, details about the study sponsor, how the data would be used, answers to frequently asked questions, information security, and selected findings from previous NPSAS rounds. The website also included contact information for the study help desk and project staff at RTI and links to the main NCES and RTI websites. Sample members could log in to the secure section of the website to update their contact information and to complete the survey. Figure 3 shows a screenshot of the NPSAS:20 website home page.

Figure 3. Home page for NPSAS:20 study website: 2019–20



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

The website, designed according to NCES web policies, protects all collected data using secure log-ins composed of a study ID and password unique to each sample member, and SSL technology for data encryption during transmission and at rest.

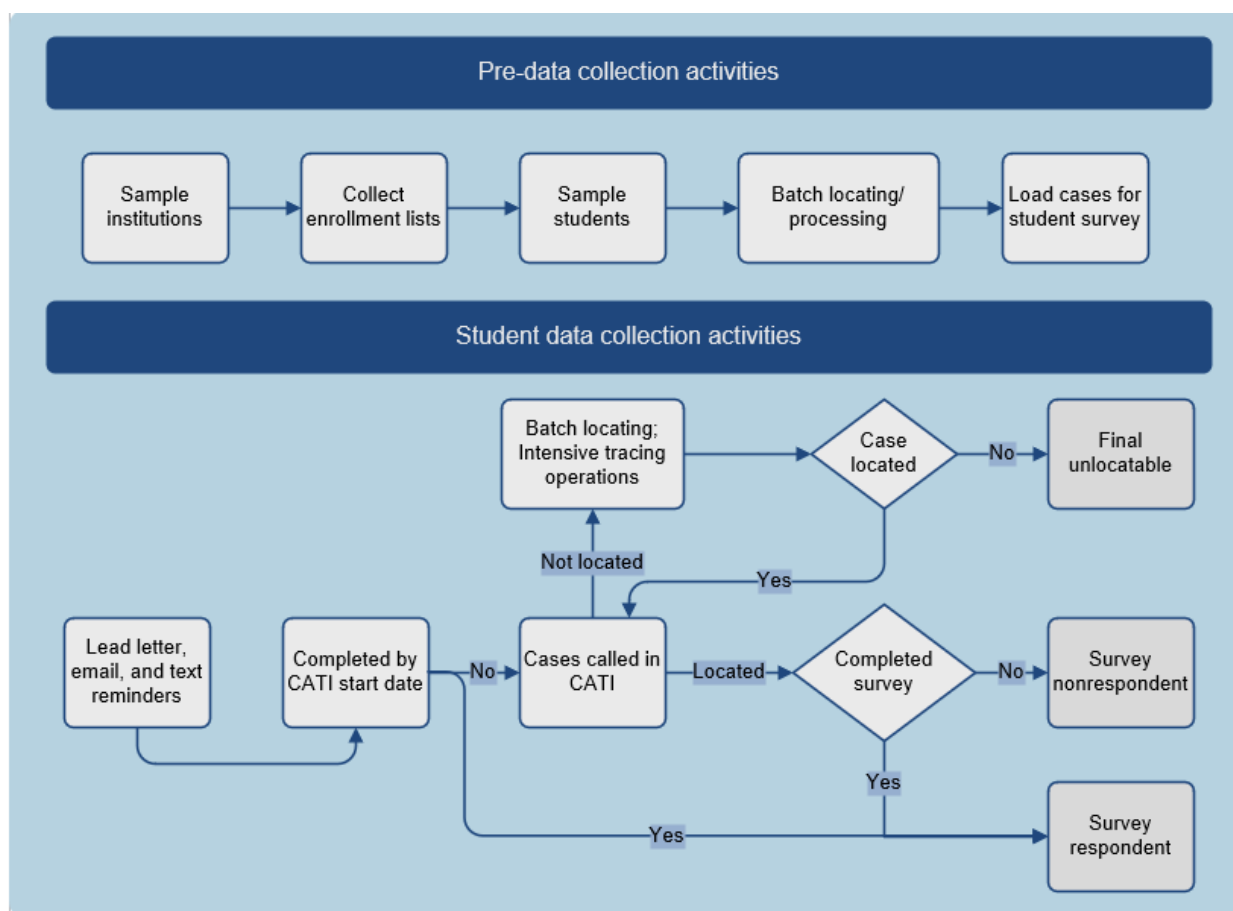
NPSAS:20 telephone interviewers also served as help-desk staff and responded to sample members' questions related to technical issues or completing the web survey. For each call received, staff confirmed contact information for the sample member for security purposes and recorded a description of the problem and resolution in a shared database.

Two common types of help-desk incidents were requests to retrieve log-in credentials and requests to complete the survey over the telephone. To minimize the need for telephone assistance, a "Forgot Password?" feature was included on the study website. After sample members entered a few pieces of identifying information, their log-in credentials were automatically provided to them via e-mail.

4.2.3 Locating and Contacting Sample Members

Before the start of data collection for each sample wave, batch locating information was used to confirm or update sample member address, telephone, and e-mail information received during collection of institution enrollment lists. At the start of each data collection wave, project staff sent a mailing, an e-mail, and a Short Message Service (SMS) text to sample members. Once outbound telephone efforts began, project staff conducted additional batch tracing, and specially trained tracers conducted intensive tracing as needed for sample members who could not be located by telephone. Once a telephone number was located for the sample member through batch or intensive tracing, interviewers contacted and invited them to complete the survey. See figure 4 for a diagram of locating activities.

Figure 4. NPSAS:20 sample member locating and contacting activities



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

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

Step 1: Project staff sent cases with a valid SSN to LexisNexis’s Single Best Address search. Single Best Address uses a combination of name and SSN to provide a new address or confirm an input address. Single Best Address uses progressive search logic to return the most current address available. Concurrently with Single Best Address, project staff sent cases with a valid SSN to CPS for record matching. CPS contains information on students who have applied for financial aid using the FAFSA. Project staff then compared records obtained from CPS to existing contact data, updating locating information as necessary.

Step 2: As a final step before the first mailing, project staff sent cases through the U.S. Postal Service (USPS) National Change of Address (NCOA) database for matching. Project staff updated records with new or updated address information based on the match. Additionally, NCOA provided a standardized version of each

input address, which was used to standardize and deduplicate the addresses in the database.

Along with the NCOA search, survey staff obtained phone information from LexisNexis Phone Append, which returns phone information based on a search by name and address.

Data collection mailings, e-mails, and SMS text messages. Using the best-known mailing address information, project staff sent mailings to the two best addresses. Mailing addresses were obtained or updated from enrollment lists, batch and intensive tracing, and other sources of locating information such as CPS or NSLDS, then prioritized based on the most recent source report date and an analysis of rates of undeliverable addresses for each source. Mailings proceeded on a flow basis, as institutions provided sample member information and as batch tracing procedures provided additional contact information. All mail correspondence was sent via USPS mail and contained a lead letter and a study brochure. The lead letter notified sample members of the start of data collection and the incentive they were eligible to receive for completing the survey. The letter also included unique log-in information for the web survey instrument and encouraged participation during the early-response period. The brochure provided information about the purpose of the study, confidentiality and security concerns, and study contact information. Project staff sent additional mailings such as postcards and letters periodically as reminders to complete the study.

Project staff sent e-mail communications to all sample member e-mail addresses collected from institutions and updated via batch tracing procedures. E-mails also went out on a flow basis and provided sample members with a link and unique log-in information to complete the survey. Additionally, SMS text messages were sent on a flow basis, providing sample members with a unique log-in link to access the survey. See appendix J for examples of the mailing and e-mail contact materials sent to sample members.

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

sample member. When the interviewer had exhausted all other sources of tracing data, production tracing was initiated.

Production tracing. Project staff referred cases that could not be located by other methods to further batch tracing and intensive tracing as needed. These tracing cases included those with no telephone number to load into CATI-CMS or for which all known numbers had failed. As a first step before intensive tracing, project staff sent cases through LexisNexis Single Best Phone and Premium Phone searches.

- **LexisNexis Single Best Phone.** Single Best Phone has access to more than 405 million phone numbers not typically published, including cell phone numbers. Single Best Phone uses a name and SSN (and can alternatively use name and address) to search multiple data sources using progressive search logic to return the most current phone number available.
- **LexisNexis Premium Phone.** Premium Phone is a residential telephone number lookup service that uses combinations of name and address information to match phone numbers to sample members. Project staff sent cases without new phone information from these searches to intensive tracing.

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

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

- Accurant database searches for sample members, parents, and other contacts;
- LexisNexis database searches including FastData reverse phone, SSN search, address search, and name search;

- Experian Address Update and Social Search; and
- matches with public records (e.g., driver’s license searches through state departments of motor vehicles).

After exhausting all possible phone leads for locating sample members, project staff set these cases to a status of “unlocatable.” If the sample member did not complete the survey by the end of data collection, that sample member was classified as a nonrespondent.

4.2.4 Training of Student Data Collection Staff

Members of the NPSAS:20 data collection team held one of four roles: telephone interviewer, quality expert (QE), quality control supervisor (QCS), or intensive tracing staff. All data collection staff, regardless of role, completed a comprehensive training program before beginning work on the study (see appendix K for training agendas). Additionally, all data collection staff completed a general training program on topics such as proper interviewing techniques, confidentiality procedures, and sample member rights.

In total, 117 telephone interviewers, QEs, and QCSs were trained over nine training sessions between January and November 2020. Due to the coronavirus pandemic, training sessions were modified to take place virtually on Zoom videoconferencing software, rather than in person as with trainings conducted for previous NPSAS surveys. Of the nine training sessions, eight were conducted in a virtual setting. Hard-copy training materials were distributed in the initial in-person training, and electronic versions of training materials were provided for virtual trainings, with hard-copy training manuals available upon request.

Training sessions, except for an abbreviated 1-night cross-training session with experienced B&B:16/20 telephone interviewers, were conducted over 3 nights and allocated equal time between direct instruction and hands-on practice.

Telephone interviewers. Telephone interviewers acted as the primary point of contact for sample members, conducting telephone interviews and employing strategies to avert or convert refusals. Telephone interviewers also served as help-desk agents to respond to sample member concerns, e-mail log-in instructions, and address incentive receipt inquiries and issues. Telephone interviewers familiarized themselves with the survey instrument and received training specific to each survey question. They developed proficiency with the survey through mock interviews, hands-on practice with case management systems, and instruction on conversational interviewing techniques. Training materials included

a telephone-interviewer manual and associated materials addressing survey administration and conversational interviewing. Project staff certified telephone interviewers after they conducted a mock interview and provided appropriate and accurate responses to NPSAS:20 frequently asked questions. Weekly Quality Circle meetings of QEs and telephone interviewers were held to review proper administration of the survey and ad hoc topics related to NPSAS:20 or general interview protocols. Project staff asked trainees for feedback to identify training needs or topics for future Quality Circle meetings.

QEs. QEs supervised telephone interviewers, performed general day-to-day monitoring responsibilities, and provided constructive feedback and coaching to interviewers after monitoring live or recorded NPSAS interviews. QEs attended interviewer training to learn survey basics and interviewing conventions.

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

Intensive tracing staff. Intensive tracing staff completed a 16-hour program on general tracing procedures with 2 additional hours of project-specific training, including the tracing techniques most appropriate for locating NPSAS:20 sample members.

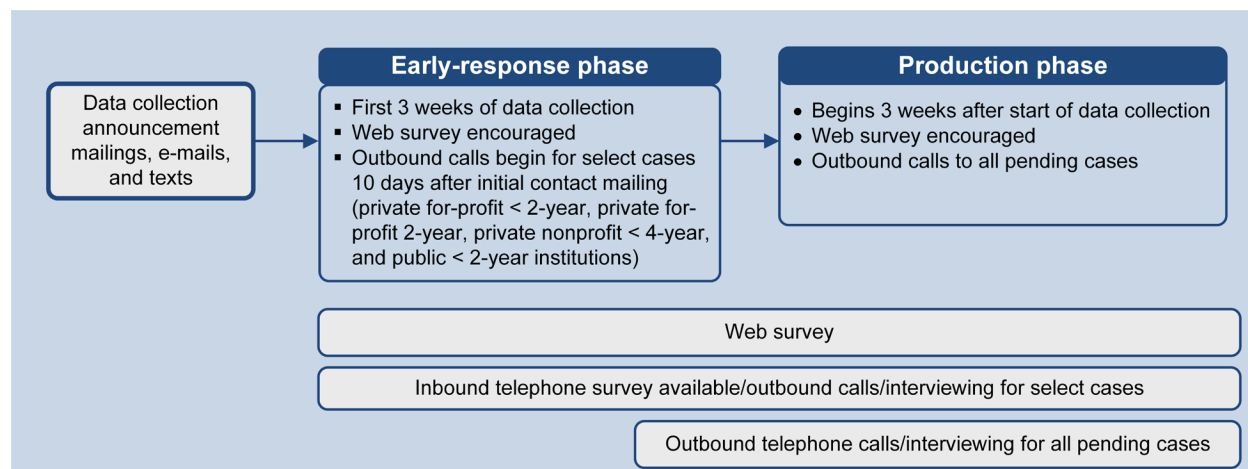
4.2.5 Interviewing

Data collection for the NPSAS:20 student survey consisted of early-response and production phases (figure 5). Regardless of when they chose to respond, sample members could complete the student survey by self-administration on the Web or by telephone.

Data collection was conducted in waves based on when sample member information became available and when staff completed batch tracing. For each wave, the initial 3-week early-response period included outbound calls to selected cases that have historically been low-response sample member types. Interviewers called sample members to encourage survey completion by web or telephone. The production phase began after week 3, at which time all remaining cases began to receive outbound telephone contacts. Project staff also sent multiple reminder

mailings, e-mails, and SMS text messages throughout the data collection period to encourage sample members to participate.

Figure 5. Data collection phases: 2019–20



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

Once outbound calling began, interviewers started locating sample members to gain their cooperation, answer questions about the study, and conduct interviews. Upon successfully reaching sample members, interviewers offered to complete the survey immediately via CATI. Alternatively, an interviewer could e-mail secure credentials for the web survey to sample members who preferred to complete the survey later. Interviewers followed up with sample members by telephone 10 days after contact if the sample member had selected the web option but had not completed the survey.

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

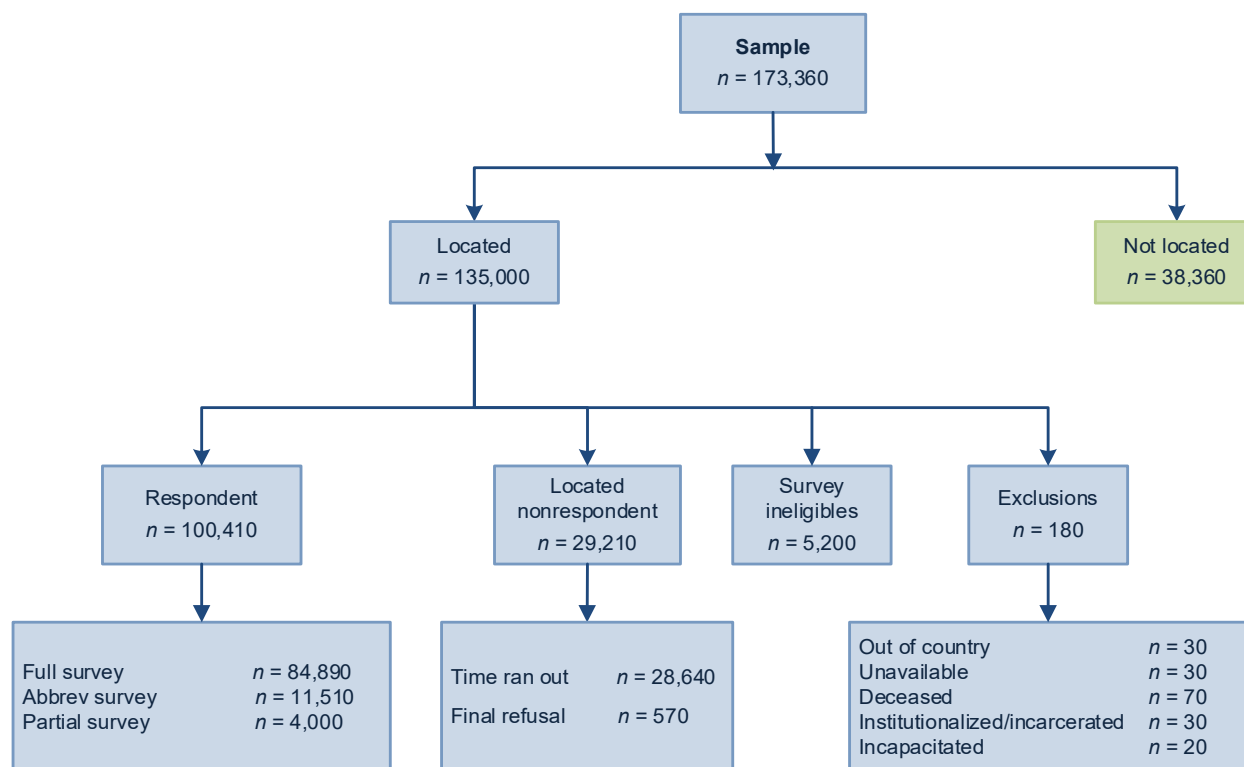
Late in data collection, an incentive boost and an abbreviated version of the survey were made available to selected sample members as part of responsive design

efforts. For further detail on responsive design, see section 4.3.6. The abbreviated survey included fewer questions and required less time, approximately 15 minutes, to complete. Interviewers used the incentive boost and abbreviated survey as refusal aversion tactics with sample members in an effort to gain their cooperation.

4.3 Student Survey Data Collection Outcomes

To assess student survey data collection outcomes, NPSAS staff determined the number of sample members located and surveyed, the time spent contacting sample members, and the rate of conversion for survey refusals. Overall, NPSAS staff located approximately 135,000 (78 percent) of the 173,360 total NPSAS:20 sample members. Of the sample members located, approximately 4 percent ($n = 5,200$) were determined to be ineligible by survey response, resulting in a total of 168,160 eligible sample members. Approximately 74 percent ($n = 100,410$) of located sample members responded, resulting in an overall response rate of 60 percent of eligible sample members ($n = 168,160$). See figure 6 for overall locating and surveying results. Section 4.3.2 provides more detail on student survey response rates.

Figure 6. NPSAS:20 student survey locating and response results: 2019–20



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

4.3.1 Student Locating Results

Locating rates by control and level of institution, shown in table 21, ranged from a high of approximately 84 percent for students enrolled at either a public 4-year, non-doctorate-granting or a private nonprofit 4-year, non-doctorate-granting institution to a low of about 69 percent for students enrolled at either a public less-than-2-year or a private for-profit 2-year institution. Data collection staff located potential undergraduate FTBs (74.5 percent) at a lower rate than other undergraduate students (77.2 percent; $\chi^2 = 133.84, p < .001$). Additionally, graduate students were located at a higher rate (86 percent) than undergraduate students (76.3 percent; $\chi^2 = 1253.83, p < .001$).

Table 21. Student locating results, by control and level of institution and student type: 2019–20

Control and level of institution and student type	Total Sample	Located ¹	
		Number	Percent of total sample
Total	173,360	135,000	77.9
Control of institution			
Public	112,040	86,410	77.1
Private nonprofit	33,250	27,310	82.1
Private for-profit	28,070	21,290	75.8
Level of institution			
Less-than-2-year	7,070	5,030	71.1
2-year	62,380	45,710	73.3
4-year, non-doctorate-granting	40,510	32,670	80.7
4-year, doctorate-granting	63,410	51,600	81.4
Control and level of institution			
Public less-than-2-year	1,580	1,090	69.2
Public 2-year	53,300	39,470	74.1
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	9,300	6,970	75.0
Public 4-year, non-doctorate-granting, primarily baccalaureate	10,530	8,820	83.8
Public 4-year, doctorate-granting	37,320	30,050	80.5
Private nonprofit less-than-4-year	1,380	950	69.1
Private nonprofit 4-year, non-doctorate-granting	11,990	10,050	83.8
Private nonprofit 4-year, doctorate-granting	19,880	16,300	82.0
Private for-profit less-than-2-year	5,440	3,900	71.7
Private for-profit 2-year	7,750	5,320	68.7
Private for-profit 4-year	14,890	12,070	81.1
Student type			
Total undergraduate	145,570	111,120	76.3
Potential FTB	48,790	36,360	74.5
Other undergraduate	96,780	74,760	77.2
Graduate	27,790	23,890	86.0

¹ Sample members are counted as located if they were ever located at some point during data collection.

NOTE: FTB = first-time beginning student. 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

Batch tracing. Matching the NPSAS sample with the CPS database resulted in updated or confirmed contact information for about 66 percent of the cases submitted for batch tracing (see table 22). Concurrently to the CPS matching, staff matched the NPSAS sample with the NSLDS database, which resulted in updated or confirmed contact information for about 70 percent of the cases submitted. NPSAS staff also sent cases to LexisNexis’s Single Best Address search. Of the 159,000 cases sent, Single Best Address returned 144,210 (91 percent) with new or confirmed addresses. NPSAS staff then submitted all existing and updated contact information received from CPS, NSLDS, and Single Best Address to the NCOA database. Of the 167,560 cases sent to NCOA, NCOA returned a new address for 26,830 (16 percent).

As part of the NCOA batch tracing step, NPSAS staff submitted sample member information to PhoneAppend for telephone number updates. Of the 167,560 cases submitted, PhoneAppend returned 87,060 (52 percent) with new or confirmed telephone numbers. Before intensive tracing, NPSAS staff submitted a small group of cases to Single Best Phone and Premium Phone after exhausting all other leads. Of the 13,380 cases submitted, Single Best Phone returned 7,320 (55 percent) with new or confirmed telephone numbers. Of the 5,510 cases submitted, Premium Phone returned 2,030 (37 percent) with new or confirmed telephone numbers.

Table 22. Batch processing match rates, by method of tracing: 2019–20

Method of tracing	Number of records sent	Number of records matched	Percent matched
Single Best Address	159,000	144,210	90.7
NSLDS	173,360	120,440	69.5
CPS	159,050	104,940	66.0
Single Best Phone	13,380	7,320	54.7
Phone Append	167,560	87,060	52.0
Premium Phone	5,510	2,030	36.9
NCOA	167,560	26,830	16.0

NOTE: CPS = Central Processing System; NCOA = National Change of Address; NSLDS = National Student Loan Data System. Percent is based on the number of records sent for batch tracing. Because records were sent to multiple tracing sources, multiple record matches were possible. Match rate includes instances when sample member contact information was confirmed and when new information was provided. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding. SOURCE: U.S. Department of Education, National Center for Education Statistics, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

Intensive tracing. Staff initiated intensive tracing throughout data collection for those sample members who were without a good telephone number once CATI efforts began. Overall, 11,230 cases, 7 percent of the total sample, required intensive tracing (see table 23). By type of institution, the percentage requiring intensive tracing ranged from a high of 10 percent of students at both public less-than-2-year and private nonprofit 4-year, doctorate-granting institutions to

3 percent of students at both private for-profit less-than-2-year and private for-profit 4-year institutions.

Phone and/or address information was successfully located for 70 percent of the cases that went through intensive tracing. Of those located, 1,890 sample members (24 percent) completed the survey (see table 24).

Table 23. Cases requiring intensive tracing, by control and level of institution and student type: 2019–20

Control and level of institution and student type	Total sample	Cases requiring intensive tracing	
		Number	Percent of total sample
Total	173,360	11,230	6.5
Control of institution			
Public	112,040	7,380	6.6
Private nonprofit	33,250	2,780	8.4
Private for-profit	28,070	1,070	3.8
Level of institution			
Less-than-2-year	7,070	340	4.8
2-year	62,380	3,970	6.4
4-year, non-doctorate-granting	40,510	2,110	5.2
4-year, doctorate-granting	63,410	4,810	7.6
Control and level of institution			
Public less-than-2-year	1,580	160	9.9
Public 2-year	53,300	3,420	6.4
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	9,300	620	6.7
Public 4-year, non-doctorate-granting, primarily baccalaureate	10,530	490	4.6
Public 4-year, doctorate-granting	37,320	2,700	7.2
Private nonprofit less-than-4-year	1,380	120	8.7
Private nonprofit 4-year, non-doctorate-granting	11,990	720	6.0
Private nonprofit 4-year, doctorate-granting	19,880	1,940	9.8
Private for-profit less-than-2-year	5,440	180	3.3
Private for-profit 2-year	7,750	440	5.7
Private for-profit 4-year	14,890	450	3.0
Student type			
Total undergraduate	145,570	9,230	6.3
Potential FTB	48,790	3,860	7.9
Other undergraduate	96,780	5,370	5.6
Graduate	27,790	2,000	7.2

NOTE: FTB = first-time beginning student. Count of cases requiring intensive tracing excludes cases initiated to intensive tracing that were not traced but includes cases for which intensive tracing work began but work was stopped. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.

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

Table 24. Located and survey response rates of cases requiring intensive tracing procedures: 2019–20

Intensive tracing round	Total cases	Located in tracing operations ¹		Responded to survey ²	
		Number	Percent of total cases	Number	Percent of located in tracing operations
Total	11,230	7,860	70.0	1,890	24.0
Intensive tracing—Round 1	11,230	7,630	67.9	1,860	24.4
Intensive tracing—Round 2	1,430	790	55.2	110	13.8

¹ Total cases count excludes cases initiated to intensive tracing that were not traced. Intensive tracing for NPSAS:20 entailed a trained tracer spending dedicated time on an individual case, conducting proprietary database searches to attempt to find up-to-date telephone information for the sample member. Intensive tracing—Round 2 cases are a subset of Intensive tracing—Round 1 cases that required additional intensive tracing efforts after Intensive tracing—Round 1; therefore, the total is not the sum of Intensive tracing—Round 1 and Intensive tracing—Round 2.

² Survey respondent count includes eligible students who met the criteria for qualification as a survey respondent, which required completing at least a partial survey.

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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

4.3.2 Student Survey Response Rates

Sixty percent (100,410 students) of the eligible sample of 168,160 completed the NPSAS:20 survey (see table 25). Partial survey completers were considered respondents if they completed at least the enrollment and FAFSA sections of the student survey. Across control and level of institution, response rates ranged from about 70 percent for students at private nonprofit 4-year, non-doctorate-granting institutions to 46 percent at both public less-than-2-year and private for-profit 2-year institutions. Potential undergraduate FTBs were significantly less likely to respond than other undergraduates (54 percent compared to 59 percent) ($\chi^2 = 285.46, p < .001$). Graduate students responded at a much higher rate (72 percent) than undergraduate students (57 percent; $\chi^2 = 2062.28, p < .001$).

Table 25. Student survey response rates, by control and level of institution and student type: 2019–20

Control and level of institution and student type	Eligible sample ²	Total respondents ¹	
		Number	Percent of eligible sample
Total	168,160	100,410	59.7
Control of institution			
Public	108,230	62,930	58.1
Private nonprofit	32,690	22,290	68.2
Private for-profit	27,230	15,190	55.8
Level of institution			
Less-than-2-year	6,730	3,180	47.2
2-year	59,520	30,420	51.1
4-year, non-doctorate-granting	39,610	25,190	63.6
4-year, doctorate-granting	62,300	41,620	66.8
Control and level of institution			
Public less-than-2-year	1,520	690	45.6
Public 2-year	50,810	26,400	51.9
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	8,920	4,840	54.3
Public 4-year, non-doctorate-granting, primarily baccalaureate	10,350	6,830	66.0
Public 4-year, doctorate-granting	36,640	24,160	66.0
Private nonprofit less-than-4-year	1,340	650	48.4
Private nonprofit 4-year, non-doctorate-granting	11,800	8,290	70.3
Private nonprofit 4-year, doctorate-granting	19,550	13,350	68.3
Private for-profit less-than-2-year	5,160	2,460	47.6
Private for-profit 2-year	7,420	3,400	45.9
Private for-profit 4-year	14,650	9,330	63.7
Student type			
Total undergraduate	140,890	80,760	57.3
Potential FTB	48,790	26,470	54.3
Other undergraduate	92,100	54,290	58.9
Graduate	27,270	19,650	72.1

¹ Total respondents count includes eligible students who met the criteria for qualification as a student survey respondent, which required completing at least a partial survey.

² Eligible sample count excludes 5,200 cases determined to be ineligible for the study using data obtained from one or more sources.

NOTE: FTB = first-time beginning student. 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

Completion by phase. As described in section 4.2.5, the NPSAS:20 student survey occurred in two main phases, an early-response phase and a production phase, and in two modes, web and telephone. Of the 100,410 cases that completed the survey, about 47 percent (46,680 cases) completed during the early-response phase, and about 54 percent (53,730 cases) completed in the production phase (see table 26).

Table 26. Student survey response rates, by data collection phase, control and level of institution, and student type: 2019–20

Control and level of institution and student type	Eligible sample ¹	Total respondents		Data collection phase			
		Number	Percent of eligible	Early-response		Production	
				Number	Percent of respondents	Number	Percent of respondents
Total	168,160	100,410	59.7	46,680	46.5	53,730	53.5
Control of institution							
Public	108,230	62,930	58.1	28,510	45.3	34,410	54.7
Private nonprofit	32,690	22,290	68.2	11,360	51.0	10,930	49.0
Private for-profit	27,230	15,190	55.8	6,800	44.8	8,390	55.2
Level of institution							
Less-than-2-year	6,730	3,180	47.2	1,330	41.8	1,850	58.2
2-year	59,520	30,420	51.1	12,890	42.4	17,530	57.6
4-year, non-doctorate-granting	39,610	25,190	63.6	11,860	47.1	13,340	52.9
4-year, doctorate-granting	62,300	41,620	66.8	20,600	49.5	21,010	50.5
Control and level of institution							
Public less-than-2-year	1,520	690	45.6	280	40.8	410	59.2
Public 2-year	50,810	26,400	51.9	11,150	42.2	15,250	57.8
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	8,920	4,840	54.3	2,080	43.0	2,760	57.0
Public 4-year, non-doctorate-granting, primarily baccalaureate	10,350	6,830	66.0	3,230	47.3	3,600	52.7
Public 4-year, doctorate-granting	36,640	24,160	66.0	11,770	48.7	12,400	51.3
Private nonprofit less-than-4-year	1,340	650	48.4	290	44.6	360	55.4
Private nonprofit 4-year, non-doctorate-granting	11,800	8,290	70.3	4,200	50.7	4,090	49.3
Private nonprofit 4-year, doctorate-granting	19,550	13,350	68.3	6,870	51.5	6,480	48.5
Private for-profit less-than-2-year	5,160	2,460	47.6	1,040	42.1	1,420	57.9
Private for-profit 2-year	7,420	3,400	45.9	1,460	42.9	1,940	57.1
Private for-profit 4-year	14,650	9,330	63.7	4,310	46.1	5,030	53.9
Student type							
Total undergraduate	140,890	80,760	57.3	36,530	45.2	44,230	54.8
Potential FTB	48,790	26,470	54.3	12,090	45.7	14,380	54.3
Other undergraduate	92,100	54,290	58.9	24,440	45.0	29,850	55.0
Graduate	27,270	19,650	72.1	10,150	51.6	9,500	48.4

¹ Respondent count includes eligible students who met the criteria for qualification as a student survey respondent, which required completing at least a partial survey. Excludes 5,200 cases determined to be ineligible for the study using data obtained from one or more sources.

NOTE: FTB = first-time beginning student. "Early-response" refers to the first 21 days of data collection; "Production" refers to the remaining data collection after the early-response phase. 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

Survey response results by survey type. All sample members were offered a full survey, and over the course of data collection, certain nonrespondents were offered an abbreviated version of the survey as a nonresponse conversion technique (see section 4.3.6).

Fifty-one percent (84,890 cases) of eligible respondents completed a full survey. Of the 79,680 nonrespondents who were offered an abbreviated survey, 14 percent (11,510 cases) completed the abbreviated survey. These completion percentages include telephone interviews and self-administered web surveys. See table 27 for response rates across control and level of institution

Table 27. Student survey response rates, by survey type, control and level of institution, and student type: 2019–20

Control and level of institution and student type	Completed full survey			Offered abbreviated survey		Completed abbreviated survey	
	Eligible sample ¹	Number	Percent of eligible sample	Number	Percent of eligible sample	Number	Percent of offered abbreviated survey
Total	168,160	84,890	50.5	79,680	47.4	11,510	14.4
Control of institution							
Public	108,230	52,900	48.9	53,180	49.1	7,610	14.3
Private nonprofit	32,690	19,330	59.1	12,730	38.9	2,270	17.8
Private for-profit	27,230	12,660	46.5	13,760	50.5	1,630	11.8
Level of institution							
Less-than-2-year	6,730	2,560	38.1	3,970	59.0	400	10.1
2-year	59,520	24,900	41.8	33,210	55.8	3,940	11.9
4-year, non-doctorate-granting	39,610	21,550	54.4	17,170	43.4	2,660	15.5
4-year, doctorate-granting	62,300	35,880	57.6	25,320	40.6	4,520	17.9
Control and level of institution							
Public less-than-2-year	1,520	560	36.9	930	61.4	110	11.5
Public 2-year	50,810	21,670	42.6	28,000	55.1	3,450	12.3
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	8,920	4,060	45.5	4,690	52.5	590	12.6
Public 4-year, non-doctorate-granting, primarily baccalaureate	10,350	5,860	56.6	4,270	41.3	730	17.1
Public 4-year, doctorate-granting	36,640	20,750	56.6	15,290	41.7	2,740	17.9
Private nonprofit less-than-4-year	1,340	530	39.5	770	57.5	80	10.0
Private nonprofit 4-year, non-doctorate-granting	11,800	7,190	60.9	4,370	37.0	840	19.2
Private nonprofit 4-year, doctorate-granting	19,550	11,620	59.4	7,590	38.8	1,360	17.9
Private for-profit less-than-2-year	5,160	1,980	38.4	3,010	58.4	290	9.5
Private for-profit 2-year	7,420	2,720	36.7	4,470	60.2	420	9.3
Private for-profit 4-year	14,650	7,960	54.3	6,280	42.9	920	14.7
Student type							
Total undergraduate	140,890	67,750	48.1	70,040	49.7	9,540	13.6
Potential FTB	48,790	22,140	45.4	25,340	51.9	2,890	11.4
Other undergraduate	92,100	45,610	49.5	44,700	48.5	6,650	14.9
Graduate	27,270	17,150	62.9	9,640	35.3	1,980	20.5

¹ Excludes 5,200 cases determined to be ineligible for the study using data obtained from one or more sources.

NOTE: FTB = first-time beginning student. 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

Completion by mode. While the web survey was available from the outset of data collection, telephone contacting efforts began 1 to 3 weeks after sample members were notified of their inclusion in the study, depending on the control and level of the institution where they were enrolled. Among respondents, 96 percent (92,190 cases) completed the survey by web and the remaining 4 percent (4,220 cases) completed by telephone (table 28).

Table 28. Student survey response rates, by mode of administration, control and level of institution, and student type: 2019–20

Control and level of institution and student type	Eligible sample ²	Total completes ³		Mode of administration ¹							
				Web total		Web nonmobile		Web mobile		Telephone	
		Number	Percent of eligible	Number	Percent of completes	Number	Percent of completes	Number	Percent of completes	Number	Percent of completes
Total	168,160	96,410	57.3	92,190	95.6	47,430	46.4	44,760	49.2	4,220	4.4
Control of institution											
Public	108,230	60,510	55.9	57,850	95.6	29,270	48.4	28,590	47.2	2,660	4.4
Private nonprofit	32,690	21,600	66.1	20,910	96.8	12,890	59.7	8,020	37.1	690	3.2
Private for-profit	27,230	14,290	52.5	13,430	93.9	5,270	36.9	8,150	57.1	870	6.1
Level of institution											
Less-than-2-year	6,730	2,960	44.0	2,740	92.3	700	23.7	2,030	68.6	230	7.7
2-year	59,520	28,840	48.5	27,180	94.2	11,250	39.0	15,930	55.2	1,660	5.8
4-year, non-doctorate-granting	39,610	24,200	61.1	23,100	95.5	11,460	47.3	11,650	48.1	1,100	4.5
4-year, doctorate-granting	62,300	40,400	64.9	39,170	97.0	24,010	59.4	15,160	37.5	1,230	3.0
Control and level of institution											
Public less-than-2-year	1,520	670	43.9	620	92.5	220	33.3	400	59.2	50	7.5
Public 2-year	50,810	25,120	49.4	23,770	94.6	10,340	41.2	13,430	53.5	1,350	5.4
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	8,920	4,650	52.1	4,400	94.6	1,900	40.8	2,500	53.8	250	5.4
Public 4-year, non-doctorate-granting, primarily baccalaureate	10,350	6,590	63.7	6,290	95.4	3,220	48.8	3,070	46.6	300	4.6
Public 4-year, doctorate-granting	36,640	23,490	64.1	22,780	97.0	13,590	57.9	9,190	39.1	720	3.0
Private nonprofit less-than-4-year	1,340	610	45.2	540	89.9	190	31.1	360	58.8	60	10.1
Private nonprofit 4-year, non-doctorate-granting	11,800	8,030	68.0	7,740	96.4	4,400	54.8	3,340	41.6	290	3.6
Private nonprofit 4-year, doctorate-granting	19,550	12,970	66.3	12,630	97.3	8,300	64.0	4,320	33.3	350	2.7
Private for-profit less-than-2-year	5,160	2,270	44.0	2,100	92.3	480	21.0	1,620	71.3	180	7.7
Private for-profit 2-year	7,420	3,140	42.3	2,890	91.9	730	23.3	2,160	68.6	250	8.1
Private for-profit 4-year	14,650	8,880	60.6	8,440	95.1	4,060	45.8	4,380	49.3	440	4.9
Student type											
Total undergraduate	140,890	77,280	54.9	73,590	95.2	34,530	44.7	39,060	50.5	3,690	4.8
Potential FTB	48,790	25,030	51.3	23,770	95.0	10,600	42.3	13,180	52.6	1,260	5.0
Other undergraduate	92,100	52,250	56.7	49,820	95.3	23,940	45.8	25,880	49.5	2,440	4.7
Graduate	27,270	19,120	70.1	18,600	97.3	12,890	67.4	5,700	29.8	530	2.7

¹ Mode of administration is the mode in which the respondent completed the student survey; this mode may be different from the starting mode for respondents who completed the survey in more than one session.

² Excludes 5,200 cases determined to be ineligible for the study using data obtained from one or more sources.

³ The number of total completes excludes the 4,000 partial surveys because mode of completion is not determined until the full survey is completed.

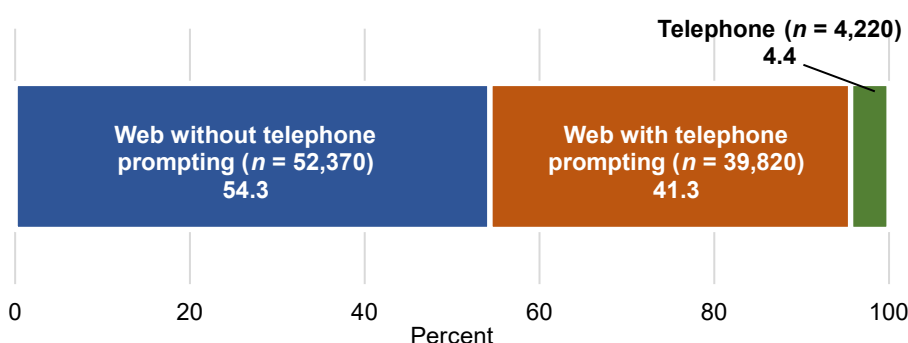
NOTE: FTB = first-time beginning student. Sample sizes rounded to the nearest 10. Percentages are based on unrounded numbers. Detail may not sum to totals because of rounding.

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

During the early-response phase, sample members received mail, e-mail, and text announcements. Select sample members (as described in section 4.2.5) were contacted by telephone. Although CATI was an option during the early-response phase, some sample members completed the survey before receiving an initial telephone contact. Additionally, some sample members did not receive any telephone contacts during data collection because no good telephone number was ever determined for the case.

In total, telephone interviews made up 4 percent of all completed surveys, while web surveys with telephone prompting made up 41 percent of the surveys completed. Web surveys without telephone prompting represented 54 percent of completed surveys (see figure 7).

Figure 7. Distribution of student survey respondents, by mode of administration: 2019–20



NOTE: Excludes 4,000 partial completes. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

4.3.3 Survey Timing Burden

This section assesses the burden associated with completing the NPSAS:20 student survey by analyzing the time required for each respondent to complete the full-scale survey. Analyses include the average timing burden for respondents to complete the survey based on mode of administration (web nonmobile, web mobile, and telephone), survey type (full or abbreviated), and student type (FTB, other undergraduate, and graduate student).

To calculate the time required to complete the survey, the survey instrument recorded the elapsed time respondents took to complete each form, or web screen. The summed form-level timing values resulted in section times and total survey times. Respondents were able to complete the survey in multiple sessions. When beginning a new session to continue the survey, respondents began on the form they last saw in their prior session. Among cases that completed in multiple sessions, the timing for the last form the respondent saw before the breakoff could

not be measured; therefore, it was imputed to the median time other respondents spent on the same form.

Of the 100,410 complete and partial full and abbreviated full-scale surveys, 70,390 (70 percent) are included in the timing report (table 29), and 30,020 cases (30 percent) were excluded from analyses. Excluded cases included partial cases, cases who completed in multiple sessions (i.e., generally cases with more than two forms requiring mean time imputation³³), and total time outliers.³⁴

Table 29. Number and percentage of NPSAS surveys included and excluded from the timing report: 2019–20

Inclusion in timing analysis and survey type	Number of cases	Percent
Total surveys (including partials)	100,410	100.0
Total completed surveys included in timing report	70,390	70.1
Full	59,290	59.0
Abbreviated	11,100	11.1
Surveys excluded from timing report	30,020	29.9
Completed in multiple sessions ¹	25,180	25.1
Total time outliers ²	840	0.8
Partial surveys	4,000	4.0

¹ Interviews completed in more than one session were included in the timing analysis when possible. Form timing is missing when the respondent is disconnected from a survey session, although some logouts do not result in disconnection from the database. Generally, if a respondent completed in more than two sessions, more than two forms would be missing timing information and thus excluded from analysis.

² Outliers were calculated by normalizing the data and excluding extreme values identified by Tukey's formula (1977), which is not sensitive to distributional assumptions. Timing cases outside of the interquartile range (IQR = third quartile - first quartile) multiplied by 1.5 were identified as total time outliers.

NOTE: Analysis includes only completed cases that have two or fewer forms of imputed timing data. Respondents who completed the survey in multiple sessions, total time outliers, and respondents who did not complete the entire survey (partial surveys) were excluded. 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

The NPSAS:20 survey contained eight key content areas: (1) enrollment; (2) FAFSA; (3) education experiences; (4) financial aid; (5) employment; (6) income and expenses; and (7) background, which included (8) a new COVID module. The COVID module contained an additional set of questions in the background section to collect information pertaining to widespread postsecondary education changes in response to the coronavirus pandemic. Because these questions were not in the initial NPSAS:20 survey design, the additional burden

³³ To avoid introducing excessive imputation and uncertainty into the timing estimates, cases that required more than two form imputations were excluded.

³⁴ To detect total time outliers, the distribution of all survey times (highly right-skewed) was first normalized using a Box-Cox power transformation (Box and Cox, 1964). Cases were then removed from the overall total using an interquartile range formula adopted from Tukey (1977) with a multiplier of 1.5. Cases were excluded as outliers if total time was greater than the 75th percentile + (1.5 * interquartile range) or if total time was less than the 25th percentile – (1.5 * interquartile range).

associated with completing the COVID module is separate from the burden estimated for the rest of the background section in this analysis.

Survey respondents had the opportunity to complete by telephone interview and self-administered web survey. In addition to accommodating different modes of administration, the survey was translated into Spanish. On average, the English full survey took 34.5 minutes to complete, and the Spanish full survey took 53 minutes to complete. Across both English and Spanish language surveys, the full survey took an average of 34.7 minutes to complete. Given that Spanish language surveys represent only 1 percent of all survey completions, the remainder of this timing analysis reports burden of English and Spanish combined. All burden comparisons reported below are significant at the $p < .0001$ level unless otherwise noted.³⁵

Table 30 provides the average timing burden of full and abbreviated surveys, by mode of administration and student type. Web nonmobile full surveys took an average of 34.4 minutes, and web mobile full surveys took substantially less time, with an average of 32.6 minutes. Full telephone surveys took an average of 60.7 minutes to complete, the greatest timing burden of all modes of administration, compared with both web nonmobile and web mobile. Because data collection interviewers read survey questions aloud to respondents, telephone interviews often result in longer completion times than self-administered web surveys.

The NPSAS:20 full-scale survey contained three distinct respondent subgroups based on student type: FTB respondents eligible for the BPS follow-up, other undergraduates, and graduate students. Potential FTB respondents, those who first began postsecondary enrollment in the 2019–20 academic year, received more survey questions than other respondents. Graduate student respondents received fewer questions than both the potential FTB respondents and the other undergraduate student groups.

The abbreviated survey included a subset of questions from the full survey, collecting high-level enrollment, FAFSA, employment, and background information, including the COVID module. The abbreviated survey took an average of 21 minutes to complete. Web nonmobile abbreviated surveys took an average of 20.6 minutes to complete, and web mobile abbreviated surveys took an average of 19.7 minutes to complete. Web nonmobile and web mobile abbreviated surveys took substantially less time than the average 35.1 minutes it took to complete abbreviated telephone surveys.

³⁵ Results use Satterthwaite (1946) approximation in difference-of-means tests with unequal variances.

Table 30. Average time, in minutes, to complete the full and abbreviated NPSAS surveys, by mode of administration and student type: 2019–20

Student type	All respondents			Mode of administration ¹								
				Web nonmobile			Web mobile			Telephone		
	Number of cases	Mean time	Median time	Number of cases	Mean time	Median time	Number of cases	Mean time	Median time	Number of cases	Mean time	Median time
Full survey	59,290	34.73	30.88	30,650	34.35	30.67	26,100	32.64	29.63	2,540	60.69	58.63
Potential FTB	14,280	44.66	40.65	6,420	45.02	41.58	7,120	41.29	38.22	740	74.13	72.57
Other undergraduate	32,310	32.76	29.15	15,480	33.23	29.67	15,340	30.05	27.38	1,480	55.99	53.91
Graduate student	12,700	28.54	25.50	8,750	28.50	25.73	3,630	26.60	24.16	320	51.58	49.75
Abbreviated survey	11,100	21.00	18.45	4,300	20.63	18.25	6,110	19.66	17.53	690	35.13	34.27
Potential FTB	2,810	24.79	22.20	890	25.09	22.68	1,730	22.99	21.08	190	39.79	39.67
Other undergraduate	6,370	20.55	18.00	2,360	20.74	18.30	3,610	18.93	16.92	400	33.96	33.23
Graduate student	1,920	16.95	14.28	1,050	16.64	14.31	770	15.55	13.68	100	30.99	29.59

¹ Mode of administration is the mode in which the respondent completed the student survey; this mode may be different from the starting mode for respondents who completed the survey in more than one session.

NOTE: FTB = first-time beginning student. Analysis includes only completed cases that have two or fewer forms of imputed timing data. Form timing is missing when the respondent is disconnected from a survey session, though some logouts do not result in disconnection from the database. Respondents who completed the survey in multiple sessions, total time outliers, and respondents who did not complete the entire survey (partial surveys) were excluded. 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

As anticipated, potential FTB respondents took longer than other student types to complete the full survey across modes, averaging 44.7 minutes. This was significantly longer than other undergraduate students, who took 32.8 minutes on average. Graduate students had the shortest completion time for the full survey, with an average of 28.5 minutes, significantly less time than both potential FTB respondents and other undergraduate respondents. This pattern holds for the abbreviated survey as well, with potential FTB respondents taking 24.8 minutes on average to complete, compared to 20.6 minutes for other undergraduates and 17 minutes for graduate students.

FTB respondents completing by telephone represented the longest average timing burden to complete the full survey. These respondents took an average of 74.1 minutes to complete the full survey, significantly longer than both other undergraduates (56 minutes) and graduate students (51.6 minutes). This was anticipated given that potential FTB respondents received the most questions in the full survey and questions were read aloud to respondents by telephone interviewers.

The remainder of this analysis addresses timing by survey section and timing of individual forms.

4.3.3.1 *Timing by survey section*

The survey design for NPSAS:20 separated questions into sections. Each section contained questions related to a general content area such as education experiences or employment. Table 31 provides the average timing for the full survey, by survey section and student type. Overall, the enrollment section took respondents the longest time to complete: 6.1 minutes on average. Potential FTB respondents spent an average of 12.4 minutes for the education experiences questions in the full survey, representing 28 percent of their entire full survey burden, significantly longer than the average 4.7 minutes for other undergraduates and the average 2.1 minutes for graduate students. Potential FTB respondents received a greater number of education experiences questions because the BPS follow-up will examine their experiences during and transitions through postsecondary education and into the labor force.

Across all student types, respondents spent an average of 2.6 minutes completing the COVID module. Graduate students took the least amount of time to answer the COVID module at an average of 2.5 minutes, significantly shorter than other undergraduates (2.6 minutes) and potential FTB respondents (2.8 minutes). This module was added after the NPSAS:20 survey was initially designed as an estimated 30-minute full survey. Without including the time required to complete the COVID module, the total timing burden to complete the full survey was 32 minutes, on average.

Table 31. Average time, in minutes, to complete the full NPSAS survey, by student type and survey section: 2019–20

Survey section	All respondents			Student type								
	Number of cases	Average time	Median time	Potential undergraduate FTB			Other undergraduate			Graduate student		
				Number of cases	Average time	Median time	Number of cases	Average time	Median time	Number of cases	Average time	Median time
Total	59,290	34.73	30.88	14,280	44.66	40.65	32,310	32.76	29.15	12,700	28.54	25.50
Enrollment	59,290	6.13	5.25	14,280	6.35	5.53	32,310	6.24	5.32	12,700	5.60	4.73
FAFSA	59,290	4.31	3.48	14,280	4.66	3.83	32,310	4.42	3.58	12,700	3.61	3.00
Education Experiences	59,290	5.98	4.02	14,280	12.39	10.90	32,310	4.68	3.72	12,700	2.09	1.63
Financial Aid	59,290	4.26	3.52	14,280	3.87	3.22	32,310	4.17	3.42	12,700	4.95	4.12
Employment	59,290	2.17	1.80	14,280	2.58	2.22	32,310	2.05	1.72	12,700	2.00	1.67
Income and Expenses	59,290	3.51	2.98	14,280	3.51	3.00	32,310	3.58	3.03	12,700	3.34	2.83
Background	59,290	2.79	2.32	14,280	3.32	2.80	32,310	2.71	2.23	12,700	2.38	1.95
COVID	59,290	2.60	2.33	14,280	2.81	2.53	32,310	2.56	2.32	12,700	2.48	2.20

NOTE: FAFSA = Free Application for Federal Student Aid; FTB = first-time beginning student. Analysis includes only completed cases that have two or fewer forms of imputed timing data. Form timing is missing when the respondent is disconnected from a survey session, though some logouts do not result in disconnection from the database. Respondents who completed the survey in multiple sessions, total time outliers, and respondents who did not complete the entire survey (partial surveys) were excluded. 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

4.3.3.2 *Timing of individual forms*

Finally, this section of the analysis details the average time to administer each individual substantive form within the survey—excluding the front-end, locating, and incentive sections. Coder forms are associated with an underlying database that provided standardized responses to an open-ended response from the respondent. Coders required respondents or the telephone interviewer to interact with the predictive text search functionality of the form, which made it a more involved process than simply typing in a response or selecting a response option on noncoder forms (i.e., forms not linked to an underlying database of standardized responses). The form-level timing analysis separated coders from other form types. Table 32 provides the average form-level timing for all coders and the top 10 longest noncoder forms. As with the other timing analyses reported, individual form results exclude cases in which more than two forms had missing timing data. Form timing is missing when the respondent is disconnected from a survey session, though some logouts do not result in disconnection from the database.

Four of the nine coders took an average of 40 seconds or more. The *Parent 1 occupation coder* (N20HPAROCC1) had the longest average form time at 59.9 seconds (table 32). This form asked respondents to code a parent or guardian's occupation using O*NET-SOC version 24.0; it may be cognitively burdensome for some respondents to correctly identify parent occupation. The next two highest burden coders in the survey also used the O*NET-SOC database. The *Expected occupation coder* (N20BEXOCC) took an average of 48.1 seconds for respondents to code their own expected occupation. *Parent 2 occupation coder* (N20HPAROCC2) took respondents an average of 46.4 seconds to code the second parent's occupation. The shorter time to report the occupation of the second parent or guardian is likely due to the respondent's increased familiarity with the occupation coder form after reporting on the occupation of the first parent or guardian, and the shorter time to code a respondent's own expected occupation may point to lower cognitive burden for the requested information compared to parents' occupation. Finally, the last coder that took more than 40 seconds on average was the *Major/field of study 1 at NPSAS institution* (N20BMAJ1), with an average administration time of 42.4 seconds. This form asked respondents to provide and code their major or field of study at the NPSAS institution using an underlying database of the 2020 CIP codes.

Table 32. Forms with the highest average administration time, in seconds, by form and item type: 2019–20

Form name ¹	Form description	Type	Number	Average (seconds)	Median (seconds)
Coders					
N20HPAROCC1	Parent 1 occupation coder	Coder	39,460	59.94	35.00
N20BEXOCC	Expected occupation coder	Coder	18,670	48.10	32.00
N20HPAROCC2	Parent 2 occupation coder	Coder	37,440	46.43	26.00
N20BMAJ1	Major/field of study 1 at NPSAS institution	Coder	54,340	42.36	26.00
N20BHSCDR	Last high school attended	Coder	44,920	32.81	16.00
N20ASCH01 ²	School 1: other postsecondary institution attended in NPSAS academic year	Coder	6,770	29.83	18.00
N20BMAJ2	Major/field of study 2 at NPSAS institution	Coder	2,060	28.00	16.00
N20BOMJ1A	Original declared major at NPSAS institution	Coder	2,850	24.91	16.00
N20BSABCOUNT	Last country in which studied abroad	Coder	5,310	12.24	9.00
Noncoder forms					
N20CCSTBKs	Cost of required textbooks and other required instructional materials	Text box	59,210	57.23	43.00
N20FCOVCOMM	Communication received from NPSAS institution due to COVID-19	Grid	55,600	46.52	37.00
N20CAIDGATE2	Financial aid used in 2019–20 academic year: graduate students	Grid	14,600	42.77	35.00
N20FCOVEXP	General experiences or disruptions due to COVID-19	Check box	68,010	39.94	31.00
N20BSRVUSE	Types of school services used in 2019–20 academic year	Grid	18,680	39.64	31.00
N20ANENRLFT	Months of full-time attendance at NPSAS institution	Months form	47,620	38.79	30.00
N20CAIDGATE1	Financial aid used in 2019–20 academic year: undergraduate students	Grid	55,720	36.66	28.00
N20BACDPART	Participation in academic activities at NPSAS	Grid	18,690	36.39	28.00
N20EFIN1YEAR	Financial literacy: amount in savings account after 1 year	Radio buttons	59,270	35.83	26.00
N20FCOVRFND	Refunds received from NPSAS institution due to COVID-19	Grid	55,630	35.77	28.00

¹ Results are based on respondent behavior on the survey forms and may not align with variables found in source files on the restricted-use file.

² This form may be administered multiple times to each respondent, so each respondent's time value is the average of all administrations. That is, it is the number of seconds to complete all iterations of the form divided by the number of iterations.

NOTE: Analysis includes only completed cases that have two or fewer forms of imputed timing data. Form timing is missing when the respondent is disconnected from a survey session, though some logouts do not result in disconnection from the database. Respondents who completed the survey in multiple sessions, total time outliers, and respondents who did not complete the entire survey (partial surveys) were excluded. For individual form time calculations, forms from the front-end, locating, and incentive sections were excluded from the analysis. Sample sizes rounded to the nearest 10.

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

The noncoder forms that took the longest to answer represent some of the most cognitively burdensome questions in the survey. These forms either relied on the respondents' recall of detailed information or took longer to read due to the form type. Six of these longest noncoder forms were in a grid format, which displays several items and scaled response options in a grid. The noncoder form with the highest timing burden, *Cost of required textbooks and other required instructional materials* (N20CCSTBKS) took respondents an average of 57.2 seconds to enter numeric values in text boxes associated with the cost of five different instructional materials. The high burden may reflect that this is one of the only forms in the NPSAS:20 survey that contains multiple numeric text boxes on a single screen, in addition to the cognitive burden required of respondents to estimate these educational costs. *Communication received from NPSAS institution due to COVID-19* (N20FCOVCOMM) took an average of 46.5 seconds to respond to a grid of five communication topics received from the NPSAS institution during students' enrollment from January 1, 2020, through June 30, 2020, which is different from the time frame referenced most in the survey (i.e., entire 2019–20 academic year). This subset time period likely introduced additional burden on the respondents, reflected in the fact that three of the 10 longest noncoder form times were from the COVID module that collected information from this subset time frame. The next highest noncoder form time was *Financial aid used in 2019–20 academic year: graduate students* (N20CAIDGATE2), which took an average of 42.8 seconds for graduate student respondents to select all forms of financial aid received in the 2019–20 academic year from a grid list of 14 aid types.

4.3.4 Contacts to Sample Members

NPSAS:20 sample members received contacts of various forms, including an initial contact mailing, e-mail, and SMS text message notifying them of their inclusion in the survey. Any remaining sample members who had not responded to these initial contacts began receiving telephone call reminders 1 to 3 weeks after initial mailing contacts, in addition to periodic reminder e-mails, mailings, and SMS text messages. On average, sample members received 20 contacts during the data collection period. Sample members received, on average, just over three calls during data collection, a reflection of declining CATI contact rates and increased use of SMS text messages (sample members received around four text messages on average). Table 33 shows the average number of contacts by control and level of institution and student type.

Table 33. Average number of contacts to sample members, by contact mode, control and level of institution, and student type: 2019–20

Control and level of institution and student type	Eligible sample ¹	Total contacts ²		Contact mode							
		Number	Average per case	Total telephone calls		Total mailings		Total e-mails		Total text messages	
				Number	Average per case	Number	Average per case	Number	Average per case	Number	Average per case
Total	168,160	3,416,360	20.3	577,240	3.4	633,550	3.8	1,552,510	9.2	653,050	3.9
Control of institution											
Public	108,230	2,259,240	20.9	381,760	3.5	418,410	3.9	1,029,540	9.5	429,530	4.0
Private nonprofit	32,690	579,370	17.7	94,260	2.9	113,060	3.5	265,250	8.1	106,800	3.3
Private for-profit	27,230	577,750	21.2	101,220	3.7	102,080	3.7	257,720	9.5	116,720	4.3
Level of institution											
Less-than-2-year	6,730	154,000	22.9	27,710	4.1	26,400	3.9	67,820	10.1	32,070	4.8
2-year	59,520	1,313,370	22.1	226,740	3.8	240,770	4.0	595,420	10.0	250,440	4.2
4-year, non-doctorate-granting	39,610	799,360	20.2	135,460	3.4	146,670	3.7	360,940	9.1	156,290	3.9
4-year, doctorate-granting	62,300	1,149,630	18.5	187,320	3.0	219,710	3.5	528,340	8.5	214,250	3.4
Control and level of institution											
Public less-than-2-year	1,520	30,590	20.1	5,820	3.8	5,600	3.7	13,540	8.9	5,630	3.7
Public 2-year	50,810	1,127,340	22.2	193,250	3.8	207,700	4.1	512,460	10.1	213,930	4.2
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	8,920	202,740	22.7	33,970	3.8	36,300	4.1	91,220	10.2	41,250	4.6
Public 4-year, non-doctorate-granting, primarily baccalaureate	10,350	210,070	20.3	36,590	3.5	38,600	3.7	95,010	9.2	39,890	3.9
Public 4-year, doctorate-granting	36,640	688,490	18.8	112,130	3.1	130,220	3.6	317,310	8.7	128,830	3.5
Private nonprofit less-than-4-year	1,340	24,450	18.3	4,680	3.5	4,450	3.3	10,680	8.0	4,650	3.5
Private nonprofit 4-year, non-doctorate-granting	11,800	213,410	18.1	35,210	3.0	41,130	3.5	97,230	8.2	39,840	3.4
Private nonprofit 4-year, doctorate-granting	19,550	341,510	17.5	54,370	2.8	67,490	3.5	157,340	8.0	62,320	3.2
Private for-profit less-than-2-year	5,160	122,280	23.7	21,710	4.2	20,610	4.0	53,770	10.4	26,200	5.1
Private for-profit 2-year	7,420	162,710	21.9	29,000	3.9	28,820	3.9	72,780	9.8	32,110	4.3
Private for-profit 4-year	14,650	292,760	20.0	50,520	3.4	52,650	3.6	131,170	9.0	58,420	4.0
Student type											
Total undergraduate	140,890	2,960,840	21.0	502,700	3.6	543,540	3.9	1,345,620	9.6	568,980	4.0
Potential FTB	48,790	1,107,500	22.7	186,910	3.8	199,580	4.1	512,650	10.5	208,360	4.3
Other undergraduate	92,100	1,853,340	20.1	315,780	3.4	343,970	3.7	832,970	9.0	360,620	3.9
Graduate	27,270	455,520	16.7	74,540	2.7	90,010	3.3	206,890	7.6	84,070	3.1

¹ Eligible sample excludes 5,200 cases determined to be ineligible for the study using data obtained from one or more sources.

² Total contacts includes all telephone calls, mailings, e-mails, and text messages sent (at a case level).

NOTE: FTB = first-time beginning student. 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

The average contacts for completed cases varied by mode of administration. Respondents who completed a telephone survey required more total contacts (5.4 more contacts on average) than those who completed the web survey. Many web survey respondents did not receive any telephone calls because they had completed the survey before telephone efforts began; those who completed the survey via web received less than two calls on average, compared to those who completed via telephone (just under four calls) (table 34).

Table 34. Average number of contacts to sample members, by contact mode, response status, and mode of administration: 2019–20

Response status and mode	Eligible sample ¹	Total contacts ²		Contact mode							
				Total telephone calls		Total mailings		Total e-mails		Total text messages	
		Number	Average per case	Number	Average per case	Number	Average per case	Number	Average per case	Number	Average per case
Response status											
Completed full survey	84,890	702,680	8.3	92,660	1.1	191,180	2.3	332,130	3.9	86,720	1.0
Completed abbreviated survey	11,510	332,930	28.9	67,590	5.9	65,270	5.7	141,330	12.3	58,750	5.1
Partial survey respondent	4,000	131,220	32.8	27,620	6.9	20,760	5.2	57,080	14.3	25,760	6.4
Nonrespondent or exclusion ³	67,750	2,249,520	33.2	389,370	5.7	356,350	5.3	1,021,980	15.1	481,820	7.1
Total	168,160	3,416,360	20.3	577,240	3.4	633,550	3.8	1,552,510	9.2	653,050	3.9
Mode of administration ⁴											
Web survey	92,190	968,560	10.5	144,770	1.6	241,800	2.6	446,250	4.8	135,750	1.5
Telephone survey	4,220	67,050	15.9	15,470	3.7	14,650	3.5	27,210	6.4	9,720	2.3
Total	96,410	1,035,620	10.7	160,240	1.7	256,450	2.7	473,460	4.9	145,470	1.5

¹ Excludes 5,200 cases determined to be ineligible for the study using data obtained from one or more sources.

² Total contacts includes all telephone calls, mailings, e-mails, and text messages sent (at a case level).

³ Exclusion cases are sample members who are out of the country, unavailable for the duration of study, deceased, institutionalized, incarcerated, or incapacitated.

⁴ Mode of administration is the mode in which the respondent completed the student survey; this mode may be different than the starting mode for respondents who completed the survey in more than one session. The “web survey” mode of administration category excludes the 4,000 partial survey respondents.

NOTE: Sample sizes rounded to the nearest 10. Detail may not sum to totals because of rounding.

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

4.3.5 Refusal Conversion

NPSAS staff integrated refusal conversion techniques into telephone-interviewer training, revisiting them throughout data collection Quality Circle meetings. Project staff sorted sample members who had ever refused to be interviewed, or had a gatekeeper refuse on their behalf, into a separate queue managed by a subset of interviewers who had received specialized refusal training. Overall, 4 percent of eligible sample members ever refused or had someone refuse on their behalf; of those, 16 percent subsequently completed the survey (table 35).

Sample members and gatekeepers provided a variety of reasons for refusing to complete the survey. The most common reasons provided were a general “not interested” response (with no mention of NPSAS:20), specifically mentioning that they did not want to participate in NPSAS:20, being too busy, or asking to be removed from the calling list. Other less common reasons for refusing were not wanting the government to have their information, survey legitimacy/confidentiality concerns, or concerns about the overall length of the survey.

Table 35. Refusal and refusal conversion rates, by control and level of institution and student type: 2019–20

Control and level of institution and student type	Eligible sample ¹	Ever any refusal		Survey respondent, given refusal ²		
		Number	Percent of eligible	Number	Percent of refused	Percent of eligible
Total	168,160	6,590	3.9	1,060	16.1	0.6
Control of institution						
Public	108,230	4,500	4.2	720	15.9	0.7
Private nonprofit	32,690	1,160	3.5	200	17.2	0.6
Private for-profit	27,230	930	3.4	140	15.3	0.5
Level of institution						
Less-than-2-year	6,730	240	3.6	30	13.3	0.5
2-year	59,520	2,480	4.2	340	13.5	0.6
4-year, non-doctorate-granting	39,610	1,510	3.8	260	17.1	0.7
4-year, doctorate-granting	62,300	2,360	3.8	430	18.3	0.7
Control and level of institution						
Public less-than-2-year	1,520	70	4.8	10	17.8	0.9
Public 2-year	50,810	2,200	4.3	300	13.6	0.6
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	8,920	400	4.5	60	15.2	0.7
Public 4-year, non-doctorate-granting, primarily baccalaureate	10,350	460	4.5	80	17.7	0.8
Public 4-year, doctorate-granting	36,640	1,370	3.7	260	19.2	0.7
Private nonprofit less-than-4-year	1,340	50	4.0	10	11.3	0.4
Private nonprofit 4-year, non-doctorate-granting	11,800	380	3.2	70	19.4	0.6
Private nonprofit 4-year, doctorate-granting	19,550	730	3.7	120	16.5	0.6
Private for-profit less-than-2-year	5,160	170	3.2	20	11.5	0.4
Private for-profit 2-year	7,420	230	3.1	30	13.9	0.4
Private for-profit 4-year	14,650	540	3.7	90	17.2	0.6
Student type						
Total undergraduate	140,890	5,440	3.9	850	15.6	0.6
Potential FTB	48,790	1,930	4.0	300	15.5	0.6
Other undergraduate	92,100	3,510	3.8	550	15.7	0.6
Graduate	27,270	1,150	4.2	210	18.1	0.8

¹ Excludes 5,200 cases determined to be ineligible for the study using data obtained from one or more sources.

² Sample members who provided a refusal response at some point during data collection but eventually completed the survey.

NOTE: FTB = first-time beginning student. 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

4.3.6 Responsive Design

The multiphase design allowed design interventions to be implemented when there was stagnation in metrics (e.g., an increase in refusals and calls per complete) and an increase in refusals and CATI hours per complete. The calibration sample experimental design for incentive and nonresponse follow-up strategies demonstrated that the protocol for Group 3 (i.e., control condition) did not yield significantly different results from the other two protocols, which involved change in the incentive offer between phases 1 and 2. As a result, Group 3 protocol was implemented during main data collection (see section 4.2.1). Thus, there were no design differences between phases 1 and 2 (both offering a \$30 promised incentive), and phase 3 introduced to respondents an abbreviated survey (average 20-minute completion time) for the same \$30 promised incentive. Because there was no difference in the data collection protocol between phases 1 and 2, before the launch of phase 3, a CATI-focused week was implemented for each wave. During that week, respondents were prompted by telephone to complete the survey online or were offered to complete via CATI. The duration of phases 1 and 2 varied by wave and was determined by monitoring metrics on the NPSAS:20 dashboard indicative of reaching phase capacity (e.g., static response rates, increased refusals and noncontacts, increased CATI hours per complete, etc.). Student survey response rates are covered in more detail in section 4.3.2.

To monitor data collection progress and inform when to implement design changes, time series regression models were run to estimate the daily cumulative response rate for each wave. The response rate was modeled as a function of various design features, including data collection interventions (e.g., reminders, prioritizing cases for telephone interviews and prompting, offering the abbreviated interview) and sample composition for each wave (e.g., institution control), allowing for autocorrelations among interventions. The overall response rate projection was the weighted average of the projections for each wave at the planned end date of data collection (January 31, 2021). Response rates were one of the most visible effects of the coronavirus pandemic for NPSAS:20.

Because response rates for student data collection were lagging model projections, with only 2 months of data collection left, a \$10 incentive boost was implemented for three key analysis groups for which the abbreviated offer was not attracting at desired rates: (1) potential FTBs; (2) undergraduate students from private for-profit institutions; and (3) undergraduate students who did not file a FAFSA. This change in the data collection protocol was needed to ensure an adequate number of participants from each group to minimize the potential for nonresponse bias and improve accuracy of estimates.

To evaluate the effectiveness of the incentive boost, the response rates for the three groups were compared to their projected response rates (based on the time series model described above) without the incentive boost. Across the waves, the incentive boost resulted in a 17.5 percent increase above the projected response rate for all three key analysis groups, and the offer of an abbreviated survey in phase 3 added an additional 5.1 percent to the overall effect. The incentive boost and the abbreviated survey resulted in increases in observed response rates compared to the projected outcomes for all data collection waves in which the boost was introduced.

4.3.7 Identification of Potential BPS-Eligible Sample Members

In past NPSAS administrations, institutions have not always been able to identify FTBs accurately. Specifically, some institutions had difficulty differentiating students who were simply new to the institution from “true” FTBs, that is, students enrolling in postsecondary education for the first time after completing high school. As described in section 2.4, although presampling matching helped to identify true FTBs, survey staff determined in the survey that some students listed and sampled as FTBs were not FTBs (false positives). Likewise, the survey identified as true FTBs some students originally listed and sampled as “not FTBs” (false negatives).

Of the 28,490 survey respondents sampled as potential FTBs, NPSAS staff confirmed that 23,700 were FTBs, based on the survey and NSLDS data, for an unweighted false-positive rate of 17 percent. Conversely, of the 71,920 survey respondents who staff sampled as other undergraduate or graduate students, about 2,770 were FTBs, for an unweighted false-negative rate of 4 percent. Table 36 shows the percentage of surveyed students who were confirmed by both the survey and NSLDS to be FTBs, based on their student type used for sampling.

Presampling matching likely reduced the false-positive rate from what it would have been without the matching. NPSAS:04 did not use presampling matching to help identify FTBs, and the false-positive rate was 54 percent (Cominole et al. 2006). Presampling matching to help identify FTBs was introduced in NPSAS:12 to reduce the false-positive rate, and the rate was 22 percent (Wine, Bryan, and Siegel 2014).

When an institution did not identify FTBs, a proxy was used to identify potential FTBs, as described in section 2.4. Sixty-five percent of the surveyed students who were sampled as FTBs using the proxy definition were successfully identified as FTBs, as shown in table 36.

Table 36. First-time beginning student status determination, by student type: 2019–20

Student type ¹	Students surveyed ²	Confirmed FTB eligibility ³	
		Number	Unweighted percent
Total sample	100,410	26,470	26.4
Total undergraduate	81,290	26,470	32.6
Potential FTB	28,490	23,700	83.2
Identified on student lists	26,800	22,610	84.4
Identified by proxy definition	1,690	1,090	64.7
Other undergraduate	52,800	2,760	5.2
Graduate	19,120	10	#

Rounds to zero.

¹ Some institution classifications of student type on the enrollment lists (e.g., undergraduate or 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.

² Students surveyed includes eligible students who met the criteria for qualification as a student survey respondent, which required completing at least a partial survey. Excludes 3,250 cases determined to be ineligible for the study using data obtained from one or more sources.

³ Confirmed FTB eligibility is defined by FTB status confirmed by both the National Student Loan Data System and the survey.

NOTE: FTB = first-time beginning student. Graduate students were not sampled as potential FTBs but can be confirmed FTBs. 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

4.4 Data Collection Quality Control

This section describes the quality control procedures employed throughout data collection. These procedures consisted of monitoring interviews, holding quality review meetings, and conducting debriefings with project staff during and after data collection.

4.4.1 Interview Monitoring

Telephone interviewers were regularly monitored during NPSAS:20 data collection to meet the following data quality objectives:

- identify difficult items in the survey;
- reduce the number of interviewer errors;
- improve interviewer performance through reinforcement of effective strategies; and
- assess the quality of the data collected.

As quality experts monitored interviewer interactions with sample members, they recorded feedback on standardized forms that covered such topics as interviewer professionalism, question administration, conversational interviewing, and familiarity with the survey instrument. Quality review meetings frequently incorporated issues identified during monitoring to improve the overall quality of telephone interviews. Segments of interviews recorded and stored in the CATI-CMS were used as training aids during project trainings and quality meetings.

4.4.2 Quality Review Meetings

Because of the coronavirus pandemic, all quality review meetings were conducted in a virtual setting utilizing Zoom videoconferencing software, rather than in-person quality review meetings that were conducted during previous NPSAS surveys. Other than being handled in a virtual setting, the quality review meetings were conducted no differently than on previous NPSAS surveys. Supervisors reinforced concepts from interview training sessions in these biweekly quality review meetings, reminding interviewers of proper administration of the survey and other topics as needed. Supervisors encouraged trainees to ask questions, which helped identify training topics for subsequent quality meetings. During NPSAS:20, some of the topics covered during quality meetings follow:

- use of help text within the survey;
- clarification of questions and item responses in the survey;
- proper administration of specific survey questions;
- successful refusal conversion techniques;
- guidelines for providing detailed sample member-level comments in the CATI-CMS;
- strategies for gaining cooperation from sample members and other contacts;
- data security protocols; and
- study progress and outline of activity schedules.

After each quality review meeting, notes were disseminated to data collection staff via an online portal. The notes provided guidance on the topics discussed at each meeting and were posted in a cumulative format, so that staff had an updated and searchable document containing all quality meeting notes compiled over the course of the project.

4.4.3 Debriefings

At both the midpoint and the end of data collection, supervisors debriefed interviewers regarding their experiences during the study. Data collection staff offered feedback to project leaders through an anonymous online survey and in-person meetings. Topics of the survey and debriefing discussions included interviewer training, interviewer support and monitoring, challenges with remote dialing, procedures for gaining sample member cooperation, and NPSAS:20 survey design and administration. Feedback from interviewers and supervisory staff will be used to inform the planning and implementation of future NPSAS surveys.

For example, in response to feedback from prior data collections, NPSAS:20 training included more active experiences with the systems and the survey instrument. NPSAS:20 interviewers reported that they appreciated these hands-on activities conducted during training but would like more time to learn how to navigate the CATI-CMS. Interviewers also noted that while the mock interview practice provided in training was beneficial, allowing interviewers to administer the interview to another interviewer while being closely supervised by monitors, they would like more time focusing on how to administer survey questions rather than just reading the questions (e.g., conversion text, interviewer instructions, etc.).

Specific to the challenges with remote dialing, interviewers reported they rarely had issues accessing projects systems and materials or experienced technical issues (bad connections, etc.). Overall, interviewers reported that they enjoyed the quality review meetings, but also suggested future meetings could include listening to clips of interviews, doing role-playing activities with difficult sample member scenarios, and reviewing proper coding procedures in the CATI-CMS.

4.5 Evaluation of Student Survey Items

This section provides an evaluation of the student survey items including analyses of the data collected in the instrument coders and a review of help text access rates, success rates for conversion text, and item nonresponse.

4.5.1 *Coder System Usability*

An analysis of each coding system's coding success rate was completed to determine the performance and usability of coding systems by respondents and telephone interviewers. For details about the reliability of coded responses, and the process of assigning codes to text strings not coded in the survey instrument, see section 6.2.2.2. Success rates for each coding system were calculated by dividing the number of responses where a complete code was selected by either the respondent or a telephone interviewer in the survey, by the total number of times the coder was administered. See section 4.1.1.2 for more information on coding systems. Analysis of successful coding rates in the survey is limited to cases with a final complete or final partial complete status, including both full and abbreviated surveys ($n = 100,410$). Overall, respondents and telephone interviewers successfully coded responses 89 percent of the time when administered a coder (table 37).

Overall coding success rates by mode of administration ranged from 92 percent for web nonmobile, 87 percent for web mobile, and 82 percent for telephone. By

coding system, coding success rates ranged from 98 percent overall success with the postsecondary institution coder to 83 percent overall success with the parent occupation coder. All coders in the survey, except for the parent occupation and expected occupation coders, had a success rate of 86 percent or above.

The postsecondary institution coder had the highest success rate of any coding system across all three modes, with 95 percent or greater rates of success in each mode. The parent occupation coder had the lowest success rate of any coding system in web nonmobile mode. The study abroad country coder had the lowest success rate of any coding system in web mobile mode. The expected occupation coder had the lowest success rate of any coding system in telephone mode. Significance tests were conducted for each of the six coders to determine significant differences in coding success rates between modes of administration. Results are significant at the $p < .0001$ level unless otherwise noted.

Rates of successful coding were significantly higher in web nonmobile mode than in web mobile mode for every coder system: *postsecondary institution*, *major*, *expected occupation*, *parent occupation*, *high school*, and *study abroad country*. These differences could be attributed to the coder forms being more difficult to navigate on the smaller mobile screens. Additionally, coding success rates were significantly higher in web nonmobile mode than in telephone mode for all coding systems. Coding success rates were significantly higher in web mobile mode than in telephone mode for five out of the six coding systems: *postsecondary institution* ($p < .05$), *major*, *expected occupation*, *parent occupation*, and *high school*. While telephone interviewers receive extensive training on using the various coders within the survey, respondents may provide text strings that do not produce desired results, making it difficult to choose a final code.

Table 37 shows a summary of coding success rate, by mode of administration and coding system.

Table 37. Percentage of responses coded in the survey, by mode of administration and coding system: 2019–20

Coding system	Mode of administration ¹			
	Overall	Web nonmobile	Web mobile	Telephone
Total	88.8	91.7	86.7	82.0
Postsecondary institution ²	97.5	98.1	96.9	95.0
Major ²	96.2	97.5	95.1	91.9
High school	86.3	88.2	85.2	80.4
Study abroad country	86.0	91.7	75.9	76.7
Occupation: expected occupation	84.7	89.4	82.2	73.6
Occupation: parent occupation ²	82.5	86.6	79.4	74.2

¹ Mode of administration is the mode in which the respondent completed the student survey; this mode may be different from the starting mode for respondents who completed the survey in more than one session.

² When a form is administered multiple times to a respondent (e.g., the postsecondary institution coder is administered for every additional institution attended in the 2019–20 academic year), the coder system success rate is calculated using the total number of administrations across respondents (i.e., the number of successful coding attempts of the form divided by the number of administrations).

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

4.5.2 Help Text

Respondents and telephone interviewers were able to click a help button provided on each form to obtain form-specific help text. Additionally, some questions provided embedded hyperlinks within question wording or response option text to prompt respondents to access the help text on the form. Whether accessed through the help button or through an embedded hyperlink, the help text provided definitions of key terms and phrases used in question wording and response options, and any explanations thought to help clarify and standardize interpretation for respondents.

Overall, respondents and interviewers accessed help text less than 1 percent of the time. Overall help text access rates were calculated by dividing the number of times that help text was accessed by the number of times that forms were administered. Form-level help text access rates were calculated by dividing the number of times help text was accessed on each form by the number of times that form was administered. The rate of help text access was also analyzed at the form level, overall and by mode of administration, to identify specific forms that may have been difficult for users. Table 38 shows the 10 forms administered to at least 10 percent of respondents or which help text was accessed at a rate of 1 percent or greater (to ensure sufficient information for analysis), by mode of administration.

Aware of emergency aid program at NPSAS institution (N20CEAIDAWARE) had the highest rate of help text access at 8 percent. This was a new construct to NPSAS:20, added to capture awareness of institution emergency aid programs. The phrase “emergency aid program” is a broad term that likely contributed to the higher help text access rate. Telephone interviewers accessed help text on this form at a significantly higher rate (14 percent; $p < .0001$) than both web nonmobile respondents (10 percent) and web mobile respondents (5 percent).³⁶

³⁶ Results use a chi-square test of association ($\alpha = 0.05$).

Table 38. Forms with at least 1 percent help text access rate, by mode of administration: 2019–20

Form ¹	Form description	Overall		Web nonmobile		Web mobile		Telephone	
		Analysis cases	Percent of help text access	Analysis cases	Percent of help text access	Analysis cases	Percent of help text access	Analysis cases	Percent of help text access
N20CEAIDAWARE	Aware of emergency aid program at NPSAS institution	83,760	7.9	42,600	10.3	37,760	4.5	3,400	14.2
N20ECARRYBAL	Credit card amount carried over each month	52,890	4.4	28,170	5.7	22,800	2.9	1,920	4.4
N20BHSMATH	Math courses taken in high school	40,710	3.2	17,330	2.5	21,520	3.6	1,860	5.5
N20BIBEXP	Took IB courses in high school	55,630	3.2	25,610	3.5	27,610	2.5	2,400	7.2
N20CPRVLN	Took out private loans in 2019–20 academic year	40,280	2.8	18,860	3.7	19,830	2.1	1,590	1.6
N20CUGPRVT	Total amount borrowed in private loans for undergraduate education	30,900	2.0	15,440	2.5	14,360	1.3	1,100	3.3
N20BEDBENEFTS	Importance of benefits	22,120	2.0	9,650	2.1	11,400	1.9	1,060	1.3
N20BAPEXP	Took AP courses in high school	55,630	1.8	25,610	2.0	27,610	1.4	2,400	4.5
N20HPAROCC1	Parent 1: occupation	54,440	1.7	25,500	2.5	26,590	1.1	2,360	0.8
N20ASTST	Enrollment intensity at NPSAS institution in 2019–20 academic year	83,760	1.1	42,600	1.3	37,760	0.7	3,400	2.2

¹ Results are based on respondent behavior on the survey forms and may not align with variables found in source files on the restricted-use file.

NOTE: AP = Advanced Placement; IB = International Baccalaureate. Mode of administration is the mode in which the respondent completed the student survey; this mode may be different from the starting mode for respondents who completed the survey in more than one session. This table only includes those items that were administered to at least 10 percent of respondents. Because of exclusions resulting from data processing, the analysis cases may differ from the total cases administered the form. 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

4.5.3 Conversion Text

To minimize item-level nonresponse (see section 4.5.4 for details on item-level nonresponse) in the full-scale survey, “conversion text” was used to encourage reluctant respondents to provide a response to forms collecting high-priority data points. On these critical forms, if a respondent moved forward in the survey without providing a response, additional text (i.e., conversion text) displayed on the same screen in web modes and was read aloud to the respondent during telephone interviews. This additional text emphasized the importance of the data being collected, why it was being requested, and encouraged respondents to respond. For some forms, a previously hidden “Don’t know” response option appeared in addition to conversion text. However, if the respondent chose not to provide an answer after conversion text was displayed, there was no additional prompting and the survey advanced to the next form. Conversion text was included on a subset of 28 critical forms. Of the 28 critical forms, this section provides results for the eight forms with sufficient information to analyze, that is, forms administered to at least 10 percent of respondents and a conversion text triggered rate of 1 percent or higher.

Overall, conversion text was triggered in 2 percent of the instances in which a critical form was administered, calculated by dividing the total number of times conversion text was triggered on a form by the number of times the form was administered. The rate of successful conversion was calculated as the total number of valid responses provided on a form after conversion text was triggered divided by the total number of instances in which conversion text was triggered. Overall, the conversion text trigger converted a missing response to a valid response 85 percent of the time. Web nonmobile surveys accounted for 35 percent of the total instances of triggering conversion text and 35 percent of the total converted instances. Web mobile surveys accounted for 59 percent of the total instances of triggering conversion text and 61 percent of total converted instances. Telephone interviews made up the remaining 6 percent of total instances of triggering conversion text and 4 percent of total converted instances.

Table 39 shows rates at which conversion text was triggered and rates at which nonresponse was converted to a response, overall and by mode of administration, for the six of the eight critical forms in the survey without a hidden “Don’t know” response option. Of these six forms, successful conversion rates ranged from 49 percent to 99 percent, with four forms resulting in conversion rates higher than 80 percent. The two forms with conversion rates lower than 80 percent were both custom forms requesting detailed information of the respondents. The first form with a conversion rate lower than 80 percent asked respondents to provide their

salary with a numeric text-box entry, along with a radio-button selection of salary time frame: *Employer 1: amount earned and earning time frame* (N20DEARN01). The second form with a conversion rate lower than 80 percent is a calendar form in which all months of the academic year were available for individual selection by the respondent: *Months worked at selected job in 2019–20 academic year* (N20DWRKMON01).

Significance tests were conducted to determine differences in conversion rates between mode of administration for each of the critical forms (results are significant at the $p < .0001$ level unless otherwise noted). These significance tests can indicate the efficacy of conversion text by mode of administration, that is, regardless of initial nonresponse rates, does conversion text perform better in particular modes of administration? Five of the six critical forms had significantly higher conversion rates in both web nonmobile mode and web mobile mode compared to telephone mode: *Number of people financially supported by parents or guardians in 2019–20 academic year* (N20HPRHSD), *Parent 1: identification and highest level of education* (N20HPARED1), *Employer 1: amount earned and earning time frame* (N20DEARN01), *Months worked at selected job in 2019–20 academic year* (N20DWRKMON01), and *Financially supported others in 2019–20 academic year* (N20HOTDEPS).

Number of people financially supported by parents or guardians in 2019–20 academic year (N20HPRHSD) had a significantly higher conversion rate in web nonmobile than in web mobile mode, while *Parent 1: identification and highest level of education* (N20HPARED1) and *Employer 1: amount earned and earning time frame* (N20DEARN01) both had significantly higher conversion rates in web mobile than in web nonmobile mode ($p < .01$ and $p < .05$, respectively). No forms had significantly higher conversion rates in telephone mode compared to other modes.

Table 39. Conversion rates for critical forms, by mode of administration: 2019–20

Form ¹	Form description	Total			Web nonmobile			Web mobile			Telephone		
		Analysis cases	Percent triggered ²	Percent converted ³	Analysis cases	Percent triggered ²	Percent converted ³	Analysis cases	Percent triggered ²	Percent converted ³	Analysis cases	Percent triggered ²	Percent converted ³
N20HPRHSD	Number of people financially supported by parents or guardians in 2019–20 academic year	2,1710	14.0	81.6	10,160	11.3	86.0	10,540	17.4	79.6	1,010	5.1	55.8
N20HPARED1	Parent 1: identification and highest level of education	100,410	3.4	94.6	48,590	2.2	95.2	47,200	4.4	97.1	4,620	4.7	66.7
N20HPARED2	Parent 2: identification and highest level of education	99,930	1.9	98.5	48,450	1.5	98.9	46,970	2.3	98.3	4,520	3.1	98.6
N20DEARN01	Employer 1: amount earned and earning time frame	70,730	1.5	48.7	34,440	1.3	47.4	33,080	1.6	54.5	3,200	2.8	22.0
N20DWRKMON01	Months worked at selected job in 2019–20 academic year	70,730	1.3	71.6	34,440	1.1	73.4	33,080	1.5	73.1	3,200	1.5	41.7
N20HOTDEPS	Financially supported others in 2019–20 academic year	69,390	1.2	97.1	34,610	0.9	96.7	31,600	1.5	98.1	3,170	0.6	77.8

¹ Results are based on respondent behavior on the survey forms and may not align with variables found in source files on the restricted-use file.

² Percent triggered is the number of instances that conversion text was triggered divided by the number of instances in which the form was administered.

³ Percent converted is the number of instances a response was provided after triggering conversion text divided by the number of instances in which conversion text was triggered.

NOTE: Mode of administration is the mode in which the respondent completed the student survey; this mode may be different from the starting mode for respondents who completed the survey in more than one session. This table only includes items that were administered to at least 10 percent of respondents, with a conversion text trigger rate of 1 percent or greater. Because of exclusions resulting from data processing, the analysis cases may differ from the total cases administered the form. 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

Table 40 shows conversion rates for the two of the eight critical forms that displayed a previously hidden “Don’t know” option once conversion text was triggered. A “Don’t know” response is considered a valid conversion for the purposes of the overall conversion rate. Both forms had a conversion rate of 42 percent. *Spouse’s income from all sources in 2019* (N20HINCSP) and *Income from all sources in 2019* (N20HINCOM) had a significantly higher conversion rate in web nonmobile mode than in telephone mode ($p < .001$ and $p < .05$, respectively). Additionally, *Income from all sources in 2019* (N20HINCOM) had a significantly higher conversion rate in web mobile mode than in telephone mode ($p < .01$). There were no other significant differences across modes in the critical items with a hidden “Don’t know” option.

Table 40. Conversion rates for critical forms with a hidden “Don’t know” response option, by mode of administration: 2019–20

Form ¹	Form description	Total				Web nonmobile				Web mobile				Telephone			
		Analysis cases	Percent triggered ²	Percent converted ³	Percent converted to a “Don’t know” ⁴	Analysis cases	Percent triggered ²	Percent converted ³	Percent converted to a “Don’t know” ⁴	Analysis cases	Percent triggered ²	Percent converted ³	Percent converted to a “Don’t know” ⁴	Analysis cases	Percent triggered ²	Percent converted ³	Percent converted to a “Don’t know” ⁴
N20HINCSP	Spouse’s income from all sources in 2019	18,020	1.4	42.2	33.1	9,610	1.3	54.1	26.2	7,690	1.0	38.7	34.7	720	7.5	20.4	46.3
N20HINCOM	Income from all sources in 2019	69,390	1.0	42.0	32.6	34,610	0.9	43.2	32.0	31,600	0.8	45.6	33.3	3,170	3.2	30.1	33.0

¹ Results are based on respondent behavior on the survey forms and may not align with variables found in source files on the restricted-use file.

² Percent triggered is the number of instances that conversion text was triggered divided by the number of instances in which the form was administered.

³ Percent converted is the number of instances a response was provided after triggering conversion text divided by the number of instances in which conversion text was triggered.

⁴ Percent converted to a “Don’t know” is the number of instances a “Don’t know” response was provided after triggering conversion text divided by the number of instances that conversion text was triggered. Percent converted to a “Don’t know” is included within the percent converted.

NOTE: Mode of administration is the mode in which the respondent completed the student survey; this mode may be different from the starting mode for respondents who completed the survey in more than one session. This table only includes items that were administered to at least 10 percent of respondents, with a conversion text trigger rate of 1 percent or greater. Because of exclusions resulting from data processing, the analysis cases may differ from the total cases administered the form. 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

Overall, conversion text is generally more effective in web nonmobile and web mobile modes than in telephone mode. Although conversion text provides context for why the requested information is critical, this may be less effective for apprehensive respondents, particularly when the requested information would need to be shared over the telephone with an interviewer.

4.5.4 Item-Level Nonresponse

The rate of nonresponse for individual items in the survey is used to identify potentially burdensome survey questions and to better understand the experiences of respondents completing the survey. An item in the survey is a single-response option (e.g., check box) or set of response options (e.g., radio-button list) that will make a variable in data processing. A form in the survey can include more than one item. Item-level nonresponse rates in the NPSAS:20 full-scale student survey were calculated for all items administered to at least 100 respondents, and all items with a nonresponse rate of 10 percent or more are reported here. Of the 520 items administered to more than 100 respondents, less than 3 percent (15 items) had a nonresponse rate of 10 percent or more. Table 41 shows item nonresponse rates for these 15 items with 10 percent or more of data missing, overall and by mode of administration.

Table 41. Item nonresponse for items with 10 percent or more of data missing, by mode of administration: 2019–20

Item ¹	Item description	Overall		Web nonmobile		Web mobile		Telephone	
		Analysis cases	Percent missing	Analysis cases	Percent missing	Analysis cases	Percent missing	Analysis cases	Percent missing
N20FLGBTQ_DK	Sexual orientation: don't know: please describe	2,130	36.2	990	35.7	1,050	38.1	90	19.6
N20HP2OCC6	Parent 2: 6-digit occupation code	51,380	30.2	24,380	26.1	24,880	35.4	2,120	15.3
N20HP2OCC3	Parent 2: 3-digit occupation code	51,380	29.1	24,380	25.2	24,880	34.2	2,120	14.4
N20BHSGPES	Estimated high school GPA	450	29.1	110	33.0	260	33.0	80	10.4
N20HP2OCC2	Parent 2: 2-digit occupation code	51,380	28.5	24,380	24.8	24,880	33.4	2,120	13.3
N20FIMGEST	Estimated age when arrived in U.S.	800	26.6	330	16.5	460	34.0	10	18.2
N20HP1OCC6	Parent 1: 6-digit occupation code	54,440	25.7	25,500	21.9	26,590	30.6	2,360	11.9
N20HP1OCC3	Parent 1: 3-digit occupation code	54,440	24.5	25,500	20.8	26,590	29.2	2,360	11.3
N20HP1OCC2	Parent 1: 2-digit occupation code	54,440	23.9	25,500	20.5	26,590	28.4	2,360	10.5
N20DSCHCURWK	Work study: currently employed	780	23.0	450	19.4	320	27.3	10	42.9
N20FGENDERQR_OTHER	Gender identity: genderqueer or gender nonconforming: please describe	1,260	15.4	710	13.8	500	17.6	60	16.1
N20CGLNEST	Estimated total amount borrowed for graduate education	330	14.7	200	18.5	130	8.3	10	28.6
N20HLT30	Respondent age range	170	13.8	80	12.7	70	19.1	30	3.7
N20CULNEST	Estimated total amount borrowed for undergraduate education	3,900	11.8	1,360	13.1	2,430	11.3	120	6.8
N20BEXOCC6	Expected occupation: 6-digit occupation code	25,000	10.5	10,590	8.1	13,170	12.5	1,250	9.5

¹ Results are based on respondent behavior on the survey forms and may not align with variables found in source files on the restricted-use file.

NOTE: GPA = grade point average. Mode of administration is the mode in which the respondent completed the student survey; this mode may be different from the starting mode for respondents who completed the survey in more than one session. This table only includes those items that were administered to at least 100 respondents. Because of exclusions resulting from data processing, the analysis cases may differ from the total cases administered the form. 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

The item with the highest overall nonresponse rate was *Sexual orientation: don't know: please describe* (N20FLGBTQ_DK). This item was administered to 2,130 respondents and had an overall rate of nonresponse of 36 percent. This item was administered to respondents who selected “Unsure” when asked about their sexual orientation. The text-box item appeared underneath the selected option to allow respondents to provide further explanation of their response. The high rate of nonresponse is likely due to the sensitive nature of the question.

The next highest overall nonresponse rates were associated with the form *Parent 2: occupation* (N20HPAROCC2), administered to 51,380 respondents. In the FAFSA section, respondents were asked to report the occupation for each of their parents using a coder form with occupations from O*NET-SOC version 24.0 as an underlying database that populated results based on an entered text string. This database used three levels to classify occupations ranging from a general area (2-digit) to a detailed classification (6-digit), all of which were distinct items in the survey. *Parent 2: 6-digit occupation code* (N20HP2OCC6) had an overall nonresponse rate of 30 percent. *Parent 2: 3-digit occupation code* (N20HP2OCC3) and *Parent 2: 2-digit occupation code* (N20HP2OCC2) had nonresponse rates of 29 percent, meaning that 1 percent of the respondents who did not provide a full 6-digit occupation code were able to partially code the occupation by providing a general 3-digit or 2-digit occupation area for their second parent. The difficulty associated with finding their parent's occupation within the populated list or their potential lack of knowledge of their parent's occupation could have contributed to the nonresponse rate for this item.

Estimated total amount borrowed for undergraduate education (N20CULNEST) was administered to 3,900 respondents and had an overall nonresponse rate of 12 percent, the second lowest of the 15 items with 10 percent or more of data missing. In the financial aid section, respondents were asked to provide a point estimate of the total amount borrowed for their undergraduate education. If respondents left the point estimate missing, they were then administered N20CULNEST, which presented a radio-button list (i.e., list of options in which only one option can be selected) of amount borrowed ranges. For the 3,900 respondents administered N20CULNEST as a nonresponse follow-up, the question effectively prompted a response for 88 percent of respondents who had left the original point estimate missing.

After reporting enrollment information, respondents were asked what their expected occupation would be after they completed all of their education. Similar to the parent occupation form, this was a coder form with an underlying database that populated possible matches based on an entered text string. Among the

15 items with 10 percent or more of data missing, *Expected occupation: 6-digit occupation code* (N20BEXOCC6) had the lowest overall nonresponse rate of 11 percent. The nonresponse rate of this question could be due to respondents having difficulty finding their specific expected job title within the populated list of results.

Item-level nonresponse rates were also examined by mode of administration for the 15 items with 10 percent or more missing data. Results are significant at the $p < .0001$ level unless otherwise noted.³⁷ Four items had significantly higher nonresponse rates in web mobile mode than in web nonmobile mode: *Work study: currently employed* (N20DSCHCURWK, $p < .05$), *Estimated age when arrived in U.S.* (N20FIMGEST), *Gender identity: genderqueer or gender nonconforming: please describe* (N20FGENDERQR_OTHER, $p < .05$), and *Expected occupation: 6-digit occupation code* (N20BEXOCC6). The higher levels of nonresponse in web mobile mode could be a result of difficulty navigating these forms, which required a text-box entry on the smaller screen. *Estimated high school GPA* (N20BHSGPES) and *Estimated total amount borrowed for undergraduate education* (N20CULNEST) had significantly higher nonresponse rates in web nonmobile mode than in telephone mode ($p < .05$). Finally, three items had significantly higher nonresponse rates in web mobile mode than in telephone mode: *Estimated high school GPA* (N20BHSGPES), *Respondent age range* (N20HLT30, $p < .05$) and *Expected occupation: 6-digit occupation code* (N20BEXOCC6). Telephone interviewers are trained to use survey tools, such as help text and conversion text, and receive extensive training on the various coders in the survey, which likely led to the lower nonresponse rates for these items in telephone mode.

³⁷ Results use Satterthwaite (1946) approximation in difference-of-means tests with unequal variances.

Chapter 5. Administrative Records

Matching Overview and Outcomes

In addition to student data obtained from the student survey and from institutions, student data for NPSAS:20 also came from administrative databases, including two from FSA: CPS and NSLDS. Additional data sources included NSC, ACT, College Board, and VBA. These additional data sources were useful in obtaining information that could not be collected from institutions or students, and when assessing the accuracy of similar information from other sources. Furthermore, some sample members for whom data were not obtained from the student survey or from student records from the institutions may only have data from administrative sources. This chapter provides detail on administrative data matching processes and outcomes.

5.1 Administrative Records Matching and Outcomes

5.1.1 *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. 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:20 against CPS data for the 2019–20 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 without sample members’ SSNs were not submitted for matching to CPS. Data were transmitted to FSA using their 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. Therefore, after September 2020, the 2019–20 FAFSA data were no longer accessible through CPS. Because NPSAS:20 sample selection was not completed by that time, only a portion of the sample members were matched to CPS for 2019–20. For the sample members selected after September 2020, FAFSA data were obtained from tables in NSLDS instead. Data from both sources were then merged.

Table 42 summarizes the results of matching student data to CPS overall and by institution and student characteristics. The overall matching rate for the 2019–20 academic year was 61 percent. Match rates varied by control and level of institution, ranging from a low of 40 percent for public less-than-2-year institutions to a high of 70 percent for private nonprofit 4-year, non-doctorate-granting institutions. These match rates met expectations based on trend estimates of Pell Grant and federal student loan funding recipients over time.³⁸

³⁸ For details on administrative matching outcomes from prior collections, see NPSAS:18-AC at <https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2022477> and NPSAS:16 at <https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2018482>.

Table 42. Results of Central Processing System (CPS) matching for 2019–20, by control and level of institution and student type: 2019–20

Control and level of institution and student type	Sample	Sent to CPS 2019–20 ¹		Matched to CPS	
		Number	Percent	Number	Percent ²
Total	380,100	351,620	92.5	213,140	60.6
Control of institution					
Public	277,470	257,340	92.7	154,420	60.0
Private not-for-profit	69,330	62,380	90.0	38,330	61.4
Private for-profit	33,300	31,900	95.8	20,400	63.9
Level of Institution					
Less-than-2-year	9,360	8,900	95.1	5,450	61.2
2-year	146,530	135,480	92.5	77,550	57.2
4-year, non-doctorate-granting	78,010	73,510	94.2	48,560	66.1
4-year, doctorate-granting	146,200	133,730	91.5	81,590	61.0
Control and level of institution					
Public less-than-2-year	2,010	1,760	87.1	700	39.9
Public 2-year	136,110	125,750	92.4	71,600	56.9
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	18,220	17,060	93.6	10,010	58.7
Public 4-year, non-doctorate-granting, primarily baccalaureate	23,250	22,150	95.3	14,630	66.1
Public 4-year, doctorate-granting	97,880	90,630	92.6	57,480	63.4
Private nonprofit less-than-4-year	1,740	1,390	79.8	840	60.8
Private nonprofit 4-year, non-doctorate-granting	27,320	25,240	92.4	17,540	69.5
Private nonprofit 4-year, doctorate-granting	40,280	35,760	88.8	19,940	55.8
Private for-profit less-than-2-year	7,280	7,080	97.3	4,700	66.3
Private for-profit 2-year	8,750	8,400	96.1	5,160	61.4
Private for-profit 4-year	17,270	16,410	95.1	10,540	64.2
Student type					
Total undergraduate	351,150	325,460	92.7	202,400	62.2
Potential FTB student	62,230	56,180	90.3	41,750	74.3
Other undergraduate	288,920	269,290	93.2	160,640	59.7
Graduate	28,950	26,160	90.4	10,750	41.1

¹ “Sent to CPS” includes cases sent for matching to CPS 2019–20 as well as cases selected after September 2020 whose Free Application for Federal Student Aid data were taken from tables in National Student Loan Data System after CPS 2019–20 was no longer available for matching.

² Percentage of cases sent to CPS.

NOTE: FTB = first-time beginning student. 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

Of all undergraduate students, 62 percent matched to CPS for the 2019–20 academic year, while only 41 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 specifically applying for federal student loans. Graduate students often apply for financial aid directly through their institution or department. Fellowship and assistantship

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

5.1.2 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 the U.S. Department of Education's Federal Student Aid office, a records match was initiated between NPSAS:20 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 a 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.

As stated previously, NSLDS matching only returned records of sample members who, at some point in time, had received Pell Grant or federal student loan funding. The NSLDS database is historical and includes information not only for the 2019–20 academic year but also for prior years. Table 43 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 2019–20. Based on existing estimates from FSA and NCES, approximately 42 percent of students attending a postsecondary institution received federal grants or loans during 2019–20.³⁹ The decline in NSLDS match rates is consistent with the literature; the College Board reports that total federal grant aid decreased by 11 percent between 2009–10 and 2019–20, and Pell Grants also declined.⁴⁰

³⁹ According to the NCES Trend Generator (<https://nces.ed.gov/ipeds/TrendGenerator>), there were about 26,093,000 students enrolled in postsecondary institutions in the 2019–20 academic year; and according to FSA's Aid Recipient Summary estimates (<https://studentaid.gov/data-center/student/title-iv>), there were approximately 10,947,400 aid recipients in the 2019–20 academic year.

⁴⁰ See more details on aid trends at <https://research.collegeboard.org/pdf/trends-college-pricing-student-aid-2020.pdf>.

Table 43. Results of National Student Loan Data System (NSLDS) loan and Pell Grant matching, by control and level of institution and student type: 2019–20

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	380,100	351,360	92.4	198,320	56.4	199,220	56.7
Control of institution							
Public	277,470	257,150	92.7	130,740	50.8	144,720	56.3
Private not-for-profit	69,330	62,340	89.9	41,380	66.4	30,210	48.5
Private for-profit	33,300	31,880	95.7	26,200	82.2	24,290	76.2
Level of Institution							
Less-than-2-year	9,360	8,890	95.0	6,550	73.6	6,740	75.7
2-year	146,530	135,380	92.4	61,770	45.6	84,980	62.8
4-year, non-doctorate-granting	78,010	73,440	94.1	45,800	62.4	43,190	58.8
4-year, doctorate-granting	146,200	133,650	91.4	84,210	63.0	64,320	48.1
Control and level of institution							
Public less-than-2-year	2,010	1,750	86.6	400	22.8	780	44.5
Public 2-year	136,110	125,660	92.3	54,180	43.1	77,270	61.5
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	18,220	17,040	93.6	7,220	42.3	10,330	60.6
Public 4-year, non-doctorate-granting, primarily baccalaureate	23,250	22,120	95.2	13,590	61.4	12,470	56.4
Public 4-year, doctorate-granting	97,880	90,570	92.5	55,370	61.1	43,880	48.4
Private nonprofit less-than-4-year	1,740	1,390	79.8	760	54.7	970	69.9
Private nonprofit 4-year, non-doctorate-granting	27,320	25,210	92.3	17,650	70.0	13,440	53.3
Private nonprofit 4-year, doctorate-granting	40,280	35,740	88.7	22,970	64.3	15,810	44.2
Private for-profit less-than-2-year	7,280	7,080	97.2	6,110	86.2	5,910	83.5
Private for-profit 2-year	8,750	8,400	96.0	6,880	81.9	6,790	80.9
Private for-profit 4-year	17,270	16,400	95.0	13,220	80.6	11,580	70.6
Student type							
Total undergraduate	351,150	325,240	92.6	179,470	55.2	187,910	57.8
Potential FTB	62,230	56,140	90.2	24,030	42.8	28,960	51.6
Other undergraduate	288,920	269,100	93.1	155,440	57.8	158,950	59.1
Graduate	28,950	26,120	90.2	18,850	72.2	11,310	43.3

¹ Percentage of cases sent to NSLDS.

NOTE: FTB = first-time beginning student. Both institution and student classifications were verified to correct classification errors on the sampling frame. Matching was completed on historical files that include awards made in 2019–20 and prior years. Percentage is of the number sent to NSLDS. 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

Match results are presented among over 350,000 sample members with a recorded SSN. In total, approximately 198,320 sample members (56 percent) were matched to loan data. NSLDS loan match rates by institution control ranged from 51 percent of public institutions to 82 percent of private for-profit institutions. The match rates by institution level ranged from 46 percent for 2-year institutions to 74 percent for less-than-2-year institutions. Match rates for control and level of institution varied from a low of 23 percent for public less-than-2-year institutions to a high of 86 percent for private for-profit less-than-2-year institutions. Of undergraduate students, 55 percent matched to the loan database, and 72 percent of graduate students matched.

NSLDS match produced Pell Grant matches for 199,220 sample members (57 percent). Match rates by control and level of institution ranged from 44 percent for private nonprofit 4-year, doctorate-granting institutions to 84 percent for private for-profit less-than-2-year institutions. Of undergraduate students, 58 percent matched to the Pell Grant database, and 43 percent of graduate students had a match.

5.1.3 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. NSC collects enrollment records from institutions representing over 97 percent of students enrolled in U.S. institutions.⁴¹ StudentTracker data were requested toward 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.

The NSC match provided information on enrollment and degree records for the 2019–20 academic year. Of the total sample members, 305,500 (80 percent) matched to NSC for their NPSAS-sampled institution. By control and level of institution, the match rate ranged from 3 percent for private for-profit less-than-2-year institutions to 92 percent for public 4-year, doctorate-granting institutions. Match rates by level of institution ranged from 10 percent for less-than-2-year

⁴¹ For more information on NSC participation, visit <https://www.studentclearinghouse.org/colleges/studenttracker>.

institutions to 89 percent for 4-year, doctorate-granting institutions. Match rates by control of institution ranged from 33 percent for private for-profit institutions to 86 percent for public institutions. Matches to institutions other than the sample members' NPSAS institutions yielded results for 48,180 sample members (13 percent). Because sample members could match to multiple institutions or to a single institution other than that reported in NPSAS, these subsets are not mutually exclusive. Table 44 shows NSC match rates by control and level of institution and student type.

Table 44. Results of National Student Clearinghouse matching, by control and level of institution and student type: 2019–20

Control and level of institution and student type	Sample	Matched for the NPSAS institution ¹		Matched for another institution ¹	
		Number	Percent	Number	Percent
Total	380,100	305,500	80.4	48,180	12.7
Control of institution					
Public	277,470	234,610	84.6	39,320	14.2
Private not-for-profit	69,330	59,840	86.3	4,980	7.2
Private for-profit	33,300	11,050	33.2	3,880	11.6
Level of Institution					
Less-than-2-year	9,360	890	9.5	660	7.1
2-year	146,530	111,940	76.4	23,390	16.0
4-year, non-doctorate-granting	78,010	62,870	80.6	10,800	13.8
4-year, doctorate-granting	146,200	129,800	88.8	13,330	9.1
Control and level of institution					
Public less-than-2-year	2,010	660	32.8	190	9.6
Public 2-year	136,110	109,290	80.3	22,200	16.3
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	18,220	15,180	83.3	2,900	15.9
Public 4-year, non-doctorate-granting, primarily baccalaureate	23,250	19,780	85.1	3,900	16.8
Public 4-year, doctorate-granting	97,880	89,700	91.6	10,140	10.4
Private nonprofit less-than-4-year	1,740	900	51.9	140	7.9
Private nonprofit 4-year, non-doctorate-granting	27,320	23,940	87.6	2,180	8.0
Private nonprofit 4-year, doctorate-granting	40,280	35,000	86.9	2,660	6.6
Private for-profit less-than-2-year	7,280	230	3.2	470	6.4
Private for-profit 2-year	8,750	1,750	19.9	1,060	12.1
Private for-profit 4-year	17,270	9,070	52.5	2,350	13.6
Student type					
Total undergraduate	351,150	282,560	80.5	46,270	13.2
Potential FTB	62,230	49,880	80.2	4,000	6.4
Other undergraduate	288,920	232,680	80.5	42,270	14.6
Graduate	28,950	22,930	79.2	1,910	6.6

¹ Sample members matched to only the NPSAS-year enrollment period (July 1, 2019–June 30, 2020).

NOTE: FTB = first-time beginning student. 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

5.1.4 ACT and SAT

NPSAS sample member data files were merged with ACT data files to obtain admissions test data. The ACT files contained survey data and a record of the most recent test score registered by each student between the 2013–14 and 2018–19 academic years. This record matching was performed toward the end of data collection to use the most updated personally identifying data (first and last name, middle initial, DOB, last four digits of the SSN, and permanent address) 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.

In total, 95,350 sample members (25 percent) matched to ACT (table 45). The match rate by control and level of institution ranged from 6 percent for students sampled from private for-profit 4-year institutions to 36 percent for students sampled from public 4-year, doctorate-granting institutions. Match rates also varied by student type, with 27 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.

To obtain SAT test scores and questionnaire data, sample member data files were merged with College Board records spanning high school graduation years from 2014 to 2019. If the file merge yielded multiple test records per student, it returned only the most recent record. Files were merged after the end of data collection using name, DOB, 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.

As shown in table 45, staff matched SAT data records for 67,870 sample members (18 percent). Match rates by institution control and level of institution ranged from 6 percent of students from private for-profit 4-year institutions to 24 percent of students from public 4-year, non-doctorate-granting, primarily baccalaureate institutions.

Table 45. Results of ACT and SAT matching, by control and level of institution and student type: 2019–20

Control and level of institution and student type	Sample	Matched to ACT ¹		Matched to SAT ²	
		Number	Percent	Number	Percent
Total	380,100	95,350	25.1	67,870	17.9
Control of institution					
Public	277,470	71,340	25.7	49,890	18.0
Private not-for-profit	69,330	21,540	31.1	15,620	22.5
Private for-profit	33,300	2,460	7.4	2,360	7.1
Level of Institution					
Less-than-2-year	9,360	910	9.7	860	9.2
2-year	146,530	26,770	18.3	18,270	12.5
4-year, non-doctorate-granting	78,010	19,610	25.1	15,800	20.3
4-year, doctorate-granting	146,200	48,070	32.9	32,940	22.5
Control and level of institution					
Public less-than-2-year	2,010	260	13.1	190	9.6
Public 2-year	136,110	25,770	18.9	17,460	12.8
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	18,220	3,520	19.3	3,420	18.8
Public 4-year, non-doctorate-granting, primarily baccalaureate	23,250	6,170	26.5	5,570	24.0
Public 4-year, doctorate-granting	97,880	35,630	36.4	23,250	23.8
Private nonprofit less-than-4-year	1,740	190	11.2	160	9.2
Private nonprofit 4-year, non-doctorate-granting	27,320	9,230	33.8	5,990	21.9
Private nonprofit 4-year, doctorate-granting	40,280	12,120	30.1	9,460	23.5
Private for-profit less-than-2-year	7,280	640	8.8	660	9.1
Private for-profit 2-year	8,750	820	9.3	650	7.4
Private for-profit 4-year	17,270	1,010	5.8	1,050	6.1
Student type					
Total undergraduate	351,150	93,710	26.7	66,770	19.0
Potential FTB	62,230	17,840	28.7	17,650	28.4
Other undergraduate	288,920	75,880	26.3	49,130	17.0
Graduate	28,950	1,640	5.7	1,090	3.8

¹ Sample members were matched to the 2013–14 through 2018–19 academic years for ACT scores.

² Sample members were matched to high school graduation years 2014–2019 for SAT scores.

NOTE: FTB = first-time beginning student. 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

5.1.5 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.

Veterans education benefits information was obtained for 23,030 of the sample members with a recorded SSN (7 percent), as shown in table 46. Match rates across control and level of institution ranged from 4 percent of cases sampled

from private for-profit less-than-2-year institutions to 15 percent at private for-profit 4-year institutions. Undergraduate students matched to VBA data at a rate of 6 percent, and graduate students had a match rate of 13 percent.

Table 46. Results of Veterans Benefits Administration (VBA) matching, by control and level of institution and student type: 2019–20

Control and level of institution and student type	Sample	Sent to VBA		Matched to VBA	
		Number	Percent	Number	Percent ¹
Total	380,100	351,780	92.5	23,030	6.5
Control of institution					
Public	277,470	257,490	92.8	15,340	6.0
Private not-for-profit	69,330	62,450	90.1	4,500	7.2
Private for-profit	33,300	31,840	95.6	3,190	10.0
Level of Institution					
Less-than-2-year	9,360	8,910	95.1	370	4.2
2-year	146,530	135,590	92.5	8,160	6.0
4-year, non-doctorate-granting	78,010	73,480	94.2	4,990	6.8
4-year, doctorate-granting	146,200	133,800	91.5	9,510	7.1
Control and level of institution					
Public less-than-2-year	2,010	1,760	87.4	100	5.9
Public 2-year	136,110	125,870	92.5	7,520	6.0
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	18,220	17,030	93.5	980	5.7
Public 4-year, non-doctorate-granting, primarily baccalaureate	23,250	22,100	95.0	1,270	5.7
Public 4-year, doctorate-granting	97,880	90,730	92.7	5,470	6.0
Private nonprofit less-than-4-year	1,740	1,380	79.3	110	7.6
Private nonprofit 4-year, non-doctorate-granting	27,320	25,290	92.6	1,790	7.1
Private nonprofit 4-year, doctorate-granting	40,280	35,780	88.8	2,610	7.3
Private for-profit less-than-2-year	7,280	7,080	97.2	270	3.8
Private for-profit 2-year	8,750	8,400	96.1	540	6.4
Private for-profit 4-year	17,270	16,360	94.7	2,380	14.6
Student type					
Total undergraduate	351,150	325,600	92.7	19,580	6.0
Potential FTB	62,230	56,220	90.3	1,790	3.2
Other undergraduate	288,920	269,380	93.2	17,780	6.6
Graduate	28,950	26,180	90.4	3,450	13.2

¹ Percentage is based on number of cases sent to VBA.

NOTE: FTB = first-time beginning student. 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

5.2 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, the file received from NSLDS had to have a header and a footer, otherwise the file and data were incomplete. If a source provided PII from their

database, names and DOBs were compared to what was in the study database to make sure the data were for the correct person. If the names and DOBs 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.

Chapter 6. Data File Processing and Preparation

NPSAS:20 student- and institution-level data were compiled from institution student records, the student survey, and matches to governmental and administrative databases. These files are fully documented and available to researchers as a set of restricted-use, microlevel 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 the NPSAS:20 study files, including processing, editing, and documentation.

6.1 Overview of the NPSAS:20 Study Files

The primary analysis files (or derived files) for NPSAS:20 contain data for 296,000 study respondents⁴² and include more than 700 derived variables created by combining data from multiple sources. Some sources have source data files available, while some do not. Complete data for NPSAS:20 is available to researchers in restricted-use files (RUFs) and documented in detail in the associated codebooks. RUFs are only released to those who have applied for and received authorization from NCES to access these data. Researchers may obtain authorization by contacting the Institute of Education Sciences (IES) Data Security Office.⁴³ The NPSAS:20 RUFs are listed in table 47. SAT, ACT, NSC, and VBA data were also used to create derived variables, in combination with data from the other primary sources (e.g., student survey, student records data

⁴² Study respondents are either (a) survey respondents: any undergraduate or graduate sample member who is determined to be eligible for the study and completed, at a minimum, the enrollment and FAFSA sections of the student survey, or (b) administrative student respondents: any undergraduate sample member who is determined to be eligible for the study and enrolled for at least 1 month (based on student records or administrative sources) and has valid data from student records. See section 7.1.1.2 for additional information about study respondents.

⁴³ More information on obtaining RUFs can be found at <https://nces.ed.gov/statprog/rudman/>.

from institutions, CPS, NSLDS). The SAT, ACT, NSC, and VBA data files are not available as source files.⁴⁴

Preliminary findings on the effects of the coronavirus pandemic on undergraduate students were released in June 2021 as a special report entitled *2019–20 National Postsecondary Student Aid Study (NPSAS:20): First Look at the Impact of the Coronavirus (COVID-19) Pandemic on Undergraduate Student Enrollment, Housing, and Finances (Preliminary Data)*. The data used in this report came primarily from new questions added to the NPSAS:20 student survey in April 2020. Data were preliminary given the report’s accelerated publication due to the urgency of the issues surrounding the pandemic. Full and summary versions of the report can be found at <https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2021456>. Final versions of the variables used in the report and additional variables documenting financial aid related to the pandemic are included in the NPSAS:20 primary analysis files, and the data files for student records and student survey.

⁴⁴ The NPSAS:20 design is unique among the NPSAS series. It contains both survey respondents (similar to NPSAS:16 and earlier rounds) and some students for whom only administrative data were available (similar to the NPSAS:18 Administrative Collection). All administrative-only respondents were undergraduates, some of whom were survey nonrespondents and others who were part of an administrative oversample that was not invited to participate in the survey. Collectively, undergraduates who were either survey respondents or administrative-only respondents are referred to as study respondents in NPSAS:20. This approach led to two statistical analysis weights: a survey weight and a study weight. The weight for the survey respondents is the survey weight (WTA000), which was designed for analysis of undergraduate or graduate student data at the national level and within each category of control and level of institution. The weight for the study respondents is the study weight (WTB000), which was designed for analysis of undergraduate data at the state level and within the public 2-year and public 4-year institution sectors within states for which the sample is sufficient for representation.

Table 47. Restricted-use NPSAS:20 files, by file name, description, and file path: 2019–20

Restricted-use file	Description	File path
NPSAS undergraduate analysis (derived) file	Contains analytic variables derived from all NPSAS:20 data sources for the 276,350 undergraduate study respondents. This file contains two weights and can be compiled to produce estimates for all study respondents or survey respondents.	/DATA/DERIVED/UNDERGRADUATE/np20derivedug_datafile.csv
NPSAS graduate analysis (derived) file	Contains analytic variables derived from all NPSAS:20 data sources for the 19,650 graduate survey respondents.	/DATA/DERIVED/GRADUATE/np20derivedgr_datafile.csv
Student survey data file	Contains data collected from the student surveys for the 100,410 survey respondents.	/DATA/SOURCE/NP20SURVEY/np20survey_datafile.csv
Student records data file	Contains data collected from the student records of the 282,410 study respondents whose institution provided data.	/DATA/SOURCE/NP20STUDRECS/np20studrecs_datafile.csv
Imputation flag	Contains imputation flags for any NPSAS:20 derived variable imputed for 296,000 NPSAS:20 study respondents.	/DATA/SOURCE/NP20FLAG/np20flag_datafile.csv
Institution data	Contains institution-level data for the NPSAS:20 sample members' NPSAS institution; 2,190 institutions are represented. Data can be linked to the student survey and student records data files by the IPEDS UNITID number.	/DATA/SOURCE/NP20INSTITUTION/np20institution_datafile.csv
CPS 2019–20 data	Contains data received from CPS for the 174,630 study respondents who matched to the 2019–20 financial aid application files.	/DATA/SOURCE/NP20CPS20/np20cps20_datafile.csv
CPS 2020–21 data	Contains data received from CPS for the 149,470 study respondents who matched to the 2020–21 financial aid application files.	/DATA/SOURCE/NP20CPS21/np20cps21_datafile.csv
CPS 2021–22 data	Contains data received from CPS for the 51,220 study respondents who matched to the 2021–22 financial aid application files.	/DATA/SOURCE/NP20CPS22/np20cps22_datafile.csv
GIS census tract data	Contains census tract-level characteristics based on geocoded location information of each student's residence in 2019–20 matched to the 2015–19 (5-year) estimates from the America Community Survey from the U.S. Census Bureau; 279,940 study respondents with valid permanent addresses are represented.	/DATA/SOURCE/NP20GIS/np20gis_datafile.csv
NSLDS loan	Contains loan-level data received from NSLDS for 155,430 matched study respondents who had received federal loans as of December 2020. Includes one record for each federal loan received and provides the most recent information for that loan.	/DATA/SOURCE/NP20NSLDSLOAN/np20nslsloan_datafile.csv
NSLDS loan disbursement	Contains loan-disbursement-level data from NSLDS for 155,040 matched study respondents who borrowed federal loans as of December 2020. Includes one record for each disbursement made on a loan.	/DATA/SOURCE/NP20NSLDSLOANDIS/np20nslsloanandis_datafile.csv
NSLDS award origin	Contains student-award-year-level data on federal Direct Loans awarded to 143,500 study respondents as of December 2020. Includes one record for each student and year during which the student was awarded a federal Direct Loan between 2012 and 2021. 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/NP20NSLDSAWARD/np20nslsaward_datafile.csv

See notes at end of table.

Table 47. Restricted-use NPSAS:20 files, by file name, description, and file path: 2019–20—Continued

Restricted-use file	Description	File path
NSLDS program enrollment status	Contains student-school-program-level enrollment information from NSLDS for 184,220 study respondents as of December 2020. Includes one record for each program and enrollment status change for a student respondent as reported to NSLDS.	/DATA/SOURCE/NP20NSLDSENROLLPROG/np20nslds enrollprog_datafile.csv
NSLDS Pell Grant data	Contains Pell Grant data received from NSLDS for 153,630 study respondents who received a federal grant as of December 2020. Includes one record for each grant received and provides distribution information for that grant.	/DATA/SOURCE/NP20NSLDSPELL/np20nsldspell_datafile.csv
NSLDS certification data	Contains student-school-level information from NSLDS on the certification of the accuracy of enrollment data for 198,220 study respondents, as reported through the Student Status Confirmation Report process as of December 2020. Contains one record for each institution a student attended.	/DATA/SOURCE/NP20NSLDS CERT/np20nsldscert_datafile.csv
NSLDS enrollment data	Contains student-school-level enrollment information from NSLDS for 209,040 study respondents as of December 2020. Includes one record for each institution a student attended.	/DATA/SOURCE/NP20NSLDENROLL/np20nsldsenroll_datafile.csv
NSLDS financial profile data	Contains records of the financial profile of 208,040 study respondents and their relatives as of December 2020. Relatives include parents, father, mother, and spouse. Includes one record for each student.	/DATA/SOURCE/NP20NSLDSFINANC PROF/np20nsldsfinancprof_datafile.csv
NSLDS gainful employment data	Contains student-school-program-level information for 21,080 study respondents on the Gainful Employment data elements of records submitted to NSLDS as of December 2020.	/DATA/SOURCE/NP20NSLDSGE/np20nsldsge_datafile.csv
NSLDS student demographic data	Contains award-year-level information on the student demographics for 210,860 study respondents as of December 2020. 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/NP20NSLDSSTUDEM/np20nsldsstudem_datafile.csv
NSLDS student demographic supplement data	Contains award-year-level information on supplemental student demographics for 135,390 study respondents as of December 2020.	/DATA/SOURCE/NP20NSLDSSTUDEMSUPP/np20nsldsstudemsupp_datafile.csv
NSLDS preferred school data	Contains information on 210,760 study respondents' plans for attending schools and corresponding housing plans as of December 2020. A FAFSA can have up to 10 preferred school(s) selected for attendance.	/DATA/SOURCE/NP20NSLDS PREFSCH/np20prefsch_datafile.csv
NPSAS:20 undergraduate weights	Contains the final NPSAS:20 undergraduate weight and variance estimation variables as a separate record for each undergraduate student respondent.	/DATA/SOURCE/NP20WEIGHTSUG/np20weightsug_datafile.csv
NPSAS:20 undergraduate weights history	Contains the intermediate undergraduate weight adjustment factors as well as the final undergraduate student weights and the variance estimation variables as a separate record for each undergraduate student respondent.	/DATA/SOURCE/NP20WEIGHTHUG/np20weighthug_datafile.csv
NPSAS:20 graduate weights	Contains the final NPSAS:20 graduate weight and variance estimation variables as a separate record for each graduate student respondent.	/DATA/SOURCE/NP20WEIGHTSGR/np20weightsgr_datafile.csv
NPSAS:20 graduate weights history	Contains the intermediate graduate weight adjustment factors as well as the final graduate student weights and the variance estimation variables as a separate record for each graduate student respondent.	/DATA/SOURCE/NP20WEIGHTHGR/np20weighthgr_datafile.csv

NOTE: COD = Common Origination and Disbursement; CPS = Central Processing System; FAFSA = Free Application for Federal Student Aid; IPEDS = Integrated Postsecondary Education Data System; NPSAS = National Postsecondary Student Aid Study; NSLDS = National Student Loan Data System.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

6.2 Post–Data-Collection Editing

This section details NPSAS:20 data editing activities for student records and the student survey.

6.2.1 Student Records

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. The student records data file, which serves as a source file for derived variables, includes student-level data reported by NPSAS institution staff. Quality control 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, the data were extracted and placed into a SAS dataset for further processing, which included various activities to ensure completeness and consistency across student records data.

The data cleaning and editing process for the NPSAS:20 student records data file 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 but 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 and NPSAS:18-AC, 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. After first using machine learning to predict probable CIP codes based on major fields coded in prior studies, these fields were reviewed by at least two expert coders in order to provide a valid CIP code where possible. This process supported the encoding of valid, consistent, and accurate majors.

Assigning missing data codes. All student records data were assigned missing data codes to indicate why data were missing. Some values were missing because an item did not apply to the student (e.g., if a student was enrolled in a bachelor’s degree program, then staff entered a value of –3 for the doctoral degree type variable). Sometimes an item was missing when, due to prior missing or unknown data, it was not possible to determine whether the item applied to the student; these items were assigned a value of –4.

Some items received a missing data code when they were left missing, but the value could be inferred from other values. If a response was provided for either other students at the institution or other items in the group (i.e., when all provided responses are positive), then a zero or no are implied for other items in the group left blank. For example, if the institution provided a positive dollar value for Pell Grant amount for a student, but left all other federal aid amounts missing, it was assumed that the missing items indicated zero. A value of –5 was assigned to indicate an “implied 0.”

A missing data code was assigned for values that were outside of a known set of values. For example, an SAT mathematics score over 800 would be assigned a code of –6. Any final missing data codes were determined to be missing because the institution did not provide an answer and were assigned a code of –9. Student records missing data codes and descriptions are provided in table 48.

Table 48. Description of missing data codes for student records: 2019–20

Student Records missing data code	Data label	Description
–1	Don't know	Institution did not have this information for the student; explicit response provided by institution.
–3	Not applicable	Item does not apply to the student.
–4	Missing – unable to determine applicability	Preceding gate item was left blank or had a “Don't know” response and it cannot be determined whether nested items apply.
–5	Missing – implied no/zero	Item 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 positive, 0/“no” are implied for other items in the group that were left blank.
–6	Out of range	Institution provided a response that was outside of the determined range for this item. For example, an ACT English score over 36 would be edited to a -6.
–9	Missing – response not provided	Institution did not provide an answer.

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

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.
- Minimum, maximum, mean, and median values of continuous variables were examined to assess reasonableness of responses. Anomalous distributions and values were investigated. If a value was impossible, such as an SAT mathematics or reading/writing score outside of the 200–800 range, it was set to –6 as discussed above. However, most often this investigation revealed places where the institution had provided data in the wrong field and those issues were corrected.
- Similar and related items were cross tabulated to verify that 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.

While data were processed and edited, question wording, response options, logical recoding, and the “applies to” text (e.g., description of students to whom the item was administered) were documented for each delivered variable from the student records data collection (appendix E).

6.2.2 Student Survey

During data collection, quality control checks were performed on all survey items to ensure the quality and accuracy of data. The survey data file, which serves as a source file for derived variables, includes data as reported by respondents. Self-reported data inconsistencies are not edited nor are they documented in source files. During derived variable construction, inconsistencies that are identified may be remedied at that point. Outlier data in continuous variables are not bottom-coded, top-coded, or blanked out, but remain in the source file for reference during derived variable construction. Documentation for these variables includes question wording, response options, logical imputations, and “applies to” text (e.g., description of respondents to whom the question was administered) (see the full-scale survey instrument in appendix C). Preparing this survey item data file was a multifaceted process described in the steps below.

Assigning missing data codes. All missing data from the survey were assigned missing data codes to indicate why data were missing. Project staff defined gate-nest question relationships, in which “gate” questions must first be answered before dependent “nest” questions. Some values were missing due to appropriate question routing (e.g., a respondent with no dependents would not be administered a question about a dependent’s age). These values were assigned a missing data code of –3, to indicate item “not applicable.” If a value was missing because the respondent completed the abbreviated or mini survey and the item was excluded from those surveys, the value was assigned a –7, “not included in abbreviated survey.” If enough information was available to determine that a variable not included in the abbreviated or mini survey would still not apply had the respondent been administered the full survey, then a –3 was applied in place of the –7. Therefore, all variables not included in the abbreviated or mini survey will not have the same number of –7s. Sometimes an item was not administered when, due to prior missing data, it was not possible to determine whether the item applied to the respondent; these items were assigned a value of –4. Because of the coronavirus pandemic, items were added to the survey in April 2020 that measured the impact on student experiences. Respondents completing the survey prior to April 2020 were assigned a value of –10 to indicate they were not administered those items and did not have an opportunity to provide an answer.

Some items received a missing data code when they were administered, but the response could be inferred from other responses. For example, if a form displayed multiple items, each with yes/no checkboxes, and the respondent checked “yes” for at least one item but left all other items missing, it was assumed that the

respondent intended for the missing items to be “no.” A value of –5 to indicate an “implied no” was assigned.

Assigning these codes during data collection served as a quality control check for the instrument operation, final data file quality, and documentation accuracy. For example, if an investigation revealed survey routing was not operating properly, an update was deployed to the survey and the item was assigned a value of –8, signifying that the item was missing due to an instrument error.

Any final missing data codes were determined to be missing because the respondent did not provide an answer and were assigned a code of –9. See survey missing data codes and descriptions in table 49.

Table 49. Description of missing data codes for student survey: 2019–20

Survey missing data code	Data label	Description
–1	Don't know	Respondent did not know the correct response; explicit response provided by respondent.
–3	Skipped – not applicable	Item does not apply to the respondent.
–4	Missing – unable to determine applicability	Preceding gate was left blank and cannot determine whether dependent nested items apply.
–5	Missing – implied no	Item left blank by respondents, but a positive response was provided for other items in the group. (When some grouped items with a response are positive, a “0” or “no” value is implied for other items in the group left blank.)
–7	Missing – not included in abbreviated survey	Item not included in the abbreviated or mini survey, so respondent did not have an opportunity to provide an answer.
–8	Missing – instrument error	Item is missing data due to an instrument error.
–9	Missing – response not provided	Respondent saw item but did not provide an answer.
–10	Missing – added to the survey after data collection began	Item was not administered at beginning of data collection. Respondents completing the survey prior to April 2020 did not have an opportunity to provide an answer to items that measure the impact of the coronavirus pandemic (COVID-19) on student experiences.

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

Applying logical recodes. Logical recodes of values were performed when the value of missing items could be definitively determined (as opposed to implied) from answers to previous survey questions. For example, if a respondent answered, “studied abroad for the entire year” to *did you attend [NPSAS institution] in 2019–20 academic year* (N20AELIG), then the item, *ever study abroad during undergraduate education* (N20BSABEVR) was skipped and logically recoded to “yes.”

Sanitizing. All open-ended responses collected in the survey were systematically reviewed. Any PII was sanitized or removed from the text string. Any other text was left unchanged. All open-ended text strings released on the restricted-use datasets were sanitized.

Coding. Predictive coding systems, or coder forms, were used to help respondents assign a code to standardized data elements such as postsecondary institutions, major, expected occupation, parent occupation, high school, and study abroad country. For each coder form, respondents entered their answer as a text string. As respondents completing in English typed, a keyword search of an underlying database returned a list of possible matches that were displayed in a drop-down menu for respondents to select. The predictive search function was disabled for respondents completing the survey in Spanish as all underlying databases are only available in English. See section 4.1.1.2 for detailed coder form descriptions and examples and see section 4.5.1 for respondent coding rates.

When an item on a coder form was not coded in the survey but an open-ended response was provided, the responses were reviewed to assign a valid code. First, the open-ended responses were automatically processed to match them to a database code, based on an exact match or similar match to database code labels. The remaining uncoded responses were loaded into an application where staff searched the coder database and assigned a code when possible. For example, if the respondent typed “Education-math” into the open-ended form for major but did not select a CIP code, the text string would be compared to all CIP code labels. Though similar, the string is not an exact match to CIP code 13.1311, “Mathematics Teacher Education,” so the text would be loaded into an application for staff review. Then, upon review, staff could assign “Education-math” to CIP code 13.1311 based on the text’s similarity to “Mathematics Teacher Education.” Results of post-data-collection edits to coder responses (i.e., recoding and upcoding) are provided below.

6.2.2.1 *Recoding*

Recoding is a process in which expert coding staff review the codes chosen in the survey on the coder forms. Using the original text strings entered on a coder form, expert coding staff selected the code that most accurately described the text string provided, then compared the original selection to the staff selection. The recoding review only applied to text strings from respondents who completed the survey in telephone mode and who therefore had their full responses entered for them by an interviewer. Due to the predictive text search functionality of the coder forms (see section 4.1.1.2 for more information on coding systems), web nonmobile and web mobile respondents often provided only partial strings before seeing and selecting an option, making it difficult for the expert coding staff to accurately review the text strings and determine a code selection.

The recoding process results in one of three recoding scenarios: (1) assigning the same code as the original selected in the survey (agreement rate), (2) recoding to a different code than selected in the survey (recode rate), or (3) determining that the original text string provided by the interviewer on behalf of the respondent was too vague to code. Because of the large variability in names given to programs of study across institutions and in job titles given to the same or similar jobs across employers, the recoding review was conducted for both majors and occupations. Overall, expert coding staff agreed with respondents' selections at least 92 percent of the time across all three coding systems. None of the recoded values were determined to be "too vague to code" for any coding system. Table 50 shows the rate of recodes for the major and occupation coding systems administered in telephone mode.

Table 50. Percentage of recoded values, by coding system: 2019–20

Coding system	Percent of recoded values		
	Recoded same as original	Recoded to a different value	Text string too vague to code
Occupation: parent occupation	94.0	6.0	0
Major	96.9	3.1	0
Occupation: expected occupation	91.6	8.4	0

NOTE: The recoding analysis only includes survey data obtained by telephone interviewers. Recode rates are calculated by dividing the total of recoded values as a result of expert staff coding by the total number of nonmissing coder values in the survey data. For coders administered multiple times to a respondent, this rate is calculated using all nonmissing coder values across all survey questions with the listed coding system, administered by telephone. Coded values resulting from the upcoding process are excluded from this table.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

6.2.2.2 Upcoding

Upcoding is a process by which expert coding staff attempt to assign a code to any text string that a respondent or interviewer did not select a code for in the survey. Text strings from each coding system went through the upcoding process. Upcode rates were calculated by dividing the number of nonmissing coder values because of expert staff upcoding by the total number of nonmissing coder values in the survey data. Overall, the *study abroad country* coder had an upcode rate of 9 percent and the *parent occupation* coder had an upcode rate of 5 percent. The remaining four coders had upcode rates lower than 5 percent. Table 51 shows the upcode rates for *study abroad country*, *parent occupation*, *expected occupation*, *postsecondary institution*, *major*, and *high school* coding systems, overall and by mode of administration.

Table 51. Summary of upcoding rates, by mode of administration and coding system: 2019–20

Coding system	Percent of text strings			
	Overall	Web nonmobile	Web mobile	Telephone
Study abroad country	9.2	5.8	15.6	15.9
Occupation: parent occupation	4.9	3.6	4.3	21.5
Occupation: expected occupation	2.6	1.4	2.5	13.6
Postsecondary institution	2.5	1.9	3.1	5.0
Major	1.8	1.1	2.3	5.8
High school	1.7	1.7	1.7	2.0

NOTE: Ucode rates are calculated by dividing the number of nonmissing coder values as a result of expert staff coding by the total number of nonmissing coder values in the survey data. For coders administered multiple times to a respondent, this rate is calculated using all nonmissing coder values across all survey questions with the listed coding system.

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

6.3 Derived Variable Construction

Project staff created the derived variables by examining student-level data available from the various data sources, prioritizing data sources specific to each item, and reconciling discrepancies both within and between sources. In some cases, staff created derived variables by assigning the value from the available source with the highest priority. In other cases, they recoded or combined source data to create a derived variable (for a listing of the analysis variables derived for NPSAS:20, see appendix L). If there was no available data from any source, the derived variable was imputed (see section 7.4 for more information about imputation procedures). Users interested in the source and imputation patterns can consult the variables that describe the sources used in derivation. These are known as “z-variables” and begin with the letter “z” and are only available on the RUF. Taken as a whole, the undergraduate and graduate derived files provide the most accurate blended data for the convenience of end users and when combined, is considered the main analysis file. As NPSAS is a fully integrated data product, incorporating multiple sources is a critical step in increasing variable accuracy and reducing missingness. Further detail on variable derivation is available in PowerStats on the “Get more info” tab for each variable, and in the RUF codebooks.

For derived variables, a value of –3 was applied when the variable does not apply to the study respondent (e.g., a respondent with no dependents would not have a valid value derived for number of dependents). Further, certain derived variables are not applicable to administrative-only student respondents (i.e., those with only administrative sources available without student survey data). For example, reliable information about private student loan borrowing is unavailable from administrative sources, so self-reported private loan information from the student survey is utilized as the sole source of this information. Therefore, for private student loan derived

variables, administrative-only student respondents have a reserve code of –13 applied. See derived variable missing data codes and descriptions in table 52.

Table 52. Description of missing data codes for derived variables: 2019–20

Derived variable missing data code	Data label	Description
–3	Skipped	Item does not apply to the respondent.
–13	Skipped – administrative-only respondent	Item does not apply to the administrative-only student respondents.

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

6.3.1 Comparison to Previous NPSAS Administrations

A subset of the variables created in NPSAS:20 will be comparable to NPSAS:18-AC but will not be comparable to administrations of NPSAS prior to 2018. These variables are different primarily due to the lack of student survey data for the students with only administrative data available. These variables include versions that are applicable only to survey respondents (the standard NPSAS variables) and versions that are derived from administrative sources only (the administrative collection variables). Another set of variables—“hybrid variables”—are variables that are derived slightly differently for survey respondents and non-survey respondents but are still considered the same variable and are comparable to administrations of NPSAS prior to 2018 (see further description of “hybrid variables” in section 6.3.1.19).

6.3.1.1 Puerto Rico

All administrations except NPSAS:87 and NPSAS:12 sampled institutions in Puerto Rico. There are approximately 90 institutions in Puerto Rico, enrolling about 1 percent each of undergraduate and graduate students nationally. In NPSAS:20, institutions in Puerto Rico accounted for 5 percent of Hispanic undergraduate students and 9 percent of Hispanic graduate students nationally. Analysts who wish to compare other NPSAS administrations to NPSAS:87 or NPSAS:12 may filter on COMPTO87 to exclude Puerto Rico.

6.3.1.2 Title IV eligibility

Starting with NPSAS:2000, samples were limited to institutions participating in federal Title IV student aid programs. In the earlier surveys (NPSAS:87, NPSAS:90, NPSAS:93, and NPSAS:96), about 1 percent of undergraduate students, mainly concentrated in private for-profit less-than-2-year institutions, attended institutions that were not eligible for Title IV aid. Analysts wishing to

exclude students from institutions that were not eligible for Title IV aid may filter on T4ELIG in these earlier administrations.

6.3.1.3 Community colleges

Over the past two decades, community colleges in many states began conferring bachelor's degrees in selected fields. In the IPEDS data used to determine control and level of institution in NPSAS, the community colleges that award bachelor's degrees are categorized as public 4-year, non-doctorate-granting institutions. Beginning with NPSAS:16, this group of institutions has been subdivided into two categories: (1) those that conferred mainly subbaccalaureate awards (certificates and associate's degrees) and (2) those that conferred mainly bachelor's degrees. This differs from NPSAS administrations prior to NPSAS:16 in which these institutions were classified together to reflect the highest level of award offered. The variable SECTOR11 makes this distinction among public 4-year institutions.

6.3.1.4 Estimates from imputed versus unimputed data

Starting with NPSAS:04, missing values were imputed for almost all, rather than a selected subset of, variables. Analysts should use caution when comparing estimates based on imputed data with estimates based on unimputed data. Distributions of imputed and unimputed variables are not directly comparable because imputed variables have no missing values, and imputation may appreciably change the distribution of valid values for variables with a substantial proportion of missing data.

6.3.1.5 Enrollment period

Starting with NPSAS:90, all NPSAS samples have been based on 12-month enrollment, which is higher than fall enrollment commonly reported by IPEDS because those who enroll only in spring, summer, or winter terms are included in the 12-month total. This is particularly true at private for-profit less-than-4-year institutions, where a substantial proportion of students may enroll throughout the year and not necessarily during the fall. The NPSAS:87 sample was based on fall 1986 enrollment. Analysts wishing to compare NPSAS:87 results with those from subsequent administrations can filter on COMPTO87 in the later NPSAS administrations.

6.3.1.6 Federal loans older than 10 years

In NPSAS administrations before 2016, federal student loans older than 10 years as of the beginning of the study were excluded from cumulative borrowing and outstanding loan amount variables. Starting in NPSAS:16, this was changed so that loans older than 10 years were included in these variables (e.g., FEDCUM1, STFCUM1). As a result, cumulative borrowing estimates in NPSAS:20,

especially for older student subpopulations, may differ from estimates for NPSAS administrations before 2016, with prior studies underestimating these amounts.

6.3.1.7 Federal veterans education benefits

Beginning with NPSAS:16, an administrative data match to VBA databases was conducted to obtain information on sampled students' receipt of federal veterans education benefits (VETBEN) and their military service. The VBA data was the sole source for federal veterans education benefits amounts, and the data include payments for tuition and fees, books and supplies, work-study, housing, and other education expenses. Estimates of federal veterans education benefits in NPSAS cycles before 2016 were derived from self-reported amounts, amounts reported by the recipient's NPSAS institution, and stochastic imputation and were significantly lower on average than amounts in NPSAS:16 and later. These earlier values may not include all the benefits included in the VBA data, particularly housing benefits, which were not explicitly requested from students or their institutions.

6.3.1.8 Financial aid variables including private loans

In NPSAS:20, a subsample of undergraduates did not complete a student survey and are considered administrative-only respondents. Without student-reported information on private loans, the percentage of missing data on private loan amounts in NPSAS:20 is too high to accurately estimate and impute private loans for students without a completed student survey. The absence of data on private loans (PRIVLOAN) has cascading effects on other financial aid variables. Variables affected by the absence of private loan data are identified with a suffix “_AC” (for “administrative collection”). These variables are not comparable to administrations of NPSAS prior to NPSAS:18-AC; however, they are comparable to those from NPSAS:18-AC. Analysts wishing to compare NPSAS:20 results with administrations from NPSAS:16 and earlier should choose financial aid variables that do not end in “_AC.” Students who completed a student survey have data on private loans (PRIVLOAN); thus, non-AC versions of these financial aid variables are available for survey respondents only.

6.3.1.9 Cumulative financial aid variables that include data on private loans

Cumulative financial aid variables that include data on private loans (e.g., BORAMT1) are not computed for the subsample of undergraduates who did not complete a student survey. The variables are available for survey respondents only.

6.3.1.10 Degree completion and parents' education

The absence of a student interview for non-survey respondents in NPSAS:20 also has implications for measuring degree completion in the NPSAS year

(PROGSTAT) and parents' education (PARED1/2, PAREduc). Administrative data alone are insufficient to create versions of these variables for students without a complete student survey that are comparable to administrations of NPSAS before NPSAS:18-AC. Variables describing degree completion and parents' education that use only administrative data have all been named with a suffix “_AC” and are comparable to NPSAS:18-AC. Non-AC versions of these variables are available for survey respondents only.

6.3.1.11 Survey-only variables

Some variables in NPSAS:20 come directly from the student survey and are not replicable from administrative sources. The subsample of undergraduate students who did not complete a student survey do not have estimates for these variables. These variables are only available for survey respondents.

6.3.1.12 Survey respondent definition

In NPSAS:20, a survey respondent was defined as any undergraduate or graduate sample member who was determined to be eligible for the study and completed at least the enrollment and FAFSA sections of the student survey. In NPSAS:16 and previous NPSAS administrations that included survey data, respondents were defined as any sampled student who was determined to be study eligible and had valid data from any source for a predetermined list of key variables. These studies had no distinction between survey and study respondents because all respondents were administered the survey.

6.3.1.13 Gender

The student survey in NPSAS:20 allowed for self-identification of gender, including: male; female; transgender, male-to-female; transgender, female-to-male; genderqueer or gender nonconforming; and a different gender identity. The variable GENDER3 includes these categories and applies only to students who completed a student survey, while the variable GENDER2 is comparable to prior administrations of NPSAS and applies to nearly all administrative student respondents, including those who did not complete a student survey. Students who reported in the survey that they were genderqueer, gender nonconforming, or had a different gender identity are skipped (-3) on GENDER2.

6.3.1.14 Graduate assistantships

In NPSAS administrations before NPSAS:20, the student survey collected graduate assistantship amounts through three separate items each addressing a different type of assistantship (teaching, research, or other). To avoid double-counting or an inaccurate division of a total assistantship amount across these

three items, the student survey in NPSAS:20 was modified to collect a single total amount of all graduate assistantships.

6.3.1.15 *Income, federal benefits, and tax variables*

Without a student survey for non-survey respondents, data on income for FAFSA nonfilers were missing not at random. Because FAFSA nonfilers have, on average, higher incomes than FAFSA filers, missing income information was imputed for the administrative-only student respondents using donors from the student survey. Due to the high degree of missingness in federal benefits (e.g., FEDBEN) and tax-related variables (e.g., PFEDTAX) for the non-survey respondents, these variables are only available for survey respondents.

6.3.1.16 *Perkins Loans*

The Perkins Loan program was discontinued in September 2017, and disbursements ceased after June 30, 2018. The variable for academic year Perkins Loans (PERKAMT) is not included in NPSAS:20 and is not incorporated in other financial aid variables that included Perkins Loans in previous administrations of NPSAS (e.g., FEDNEED). Due to these changes, the variables SUBLOAN and T4LNAME1 are not included in NPSAS:20 because, without PERKAMT, they are identical to STAFSUB and STAFFAMT, respectively. Other variables that previously included PERKAMT are still comparable to previous administrations of NPSAS because they represent the same construct (e.g., federal need-based aid), despite excluding PERKAMT.

6.3.1.17 *State-level analyses*

Although NPSAS:20 was designed to provide state-representative estimates for undergraduates overall, in the public 2-year sector and the public 4-year sector, not all states and sectors had sufficient response to be state representative. Analysts wishing to make state-representative estimates for all undergraduates within a state should utilize the state representation indicator (STATEREPI) in conjunction with the state identifier (INSTSTAT) to determine whether a state is representative of all undergraduates. Analysts wishing to make state-representative estimates of the public 2-year, or public 4-year sector should use the sector representation indicator (SECTORREPI) in conjunction with the state identifier (INSTSTAT) and SECTOR11 to identify state-sector pairs that are suitable for state-level analyses. No states have representative samples for individual sectors other than public 2-year and public 4-year institutions. A list of state and within-state representativeness by institution sector is available in section 1.3.1.

6.3.1.18 *FAFSA concordance*

Certain student characteristics that are relevant for federal aid can change between the date a student files their FAFSA and the end of the 2019–20 academic year. When values differ across data sources for these student characteristics, many NPSAS variables prioritize responses on the FAFSA in order to be consistent with students' dependency status (DEPEND), which is used in the determination of federal financial aid awards. For instance, MILTYPE2 can be used to analyze military type in conjunction with dependency status and federal financial aid. (In contrast, the alternative version MILTYPE gives priority to data from VBA.)

6.3.1.19 *Hybrid variables*

Hybrid variables have just one version—the standard version—and are derived for all respondents. These variables (e.g., ENR01) are derived using both administrative data sources and the student survey. The derivation of these variables for survey respondents thus differs slightly from the derivation for non-survey respondents, as it includes an additional source of data. However, the resulting variables are comparable to one another and to prior administrations of NPSAS and thus do not have an administrative collection version for non-survey respondents. These variables are considered “hybrid” because they combine the standard approach to derivation with the administrative collection approach to derivation, resulting in a single derived variable that applies to all respondents.

6.3.2 *COVID-19 Aid Variables*

This section briefly covers how COVID-19 aid variables differ across sources at the student level. The CARES Act was passed by Congress on March 27, 2020. This bill allotted \$2.2 trillion to provide direct economic aid to the American people negatively impacted by the COVID-19 pandemic. Of that money, approximately \$14 billion was given to the Office of Postsecondary Education as the Higher Education Emergency Relief Fund (HEERF).⁴⁵ Of those funds, more than \$6 billion was allotted to Title IV eligible institutions to be disbursed to students as emergency financial aid grants. Students who were eligible to receive Title IV financial aid were also eligible to receive emergency financial aid grants.

Although the CARES Act was signed into law on March 27, 2020, institutions began distributing aid to students at different times after this date and may not have begun disbursing emergency funds to students until well after that date. Institutions continued to disburse HEERF funds to students well into the

⁴⁵More information about HEERF can be found at <https://www2.ed.gov/about/offices/list/ope/caresact.html>.

following academic year. As the NPSAS:20 academic year begins July 1, 2019, and ends June 30, 2020, the emergency aid captured by both student records and student survey data collection represents only a subset of the total HEERF funds received by students.

NPSAS:20 collected data on emergency financial aid grants from two sources: the student survey and student records provided by institutions. These two sources show approximately the same rate of emergency financial aid received, but they have a low level of concordance for which students received emergency financial aid grants. As a result, both sources are reported separately on the derived data files for maximum utility for analysts.

Although there is no certainty which—if any—of these factors are driving the discrepancy between sources, the following are possible explanations for why the sources differ at the student level:

Timing of response to the student survey or institution records. The student or the institution submitted data before the student was awarded the emergency aid. While institutions have complex systems for tracking and reporting aid, it is unlikely that most students have this same capability. Particularly for students who responded to the survey several months or more after the receipt of aid, this can introduce errors in reporting aid that was awarded outside the NPSAS year or other dispensations that were not financial aid awards. Additionally, emergency financial aid that was awarded near the end of the NPSAS year might have been disbursed after the NPSAS year had ended. While the institution would have reported such aid as it was awarded within the NPSAS year, the student may have indicated no aid, since to them, the award was disbursed outside the NPSAS year.

Differences in measurement. The student records variable includes aid from any source, including federal, state, institution, and private. The student survey asks students to report only emergency aid received from their institution. Given the volume of students reporting that they received this aid in the survey and given the nature of how HEERF funds were disbursed to students (i.e., by the institution), it is likely that many students were reporting CARES Act funds in the survey, but it is also possible that some students maintained a strict interpretation of the question and did not report federal, state, or private awards in this question.

Chapter 7. Weighting, Variance Estimation, and Imputation

This chapter provides an overview of the statistical analysis weights for the NPSAS:20 sample, including institution and student weights. Unit-level and item-level nonresponse bias analysis, variance estimation, imputation for missing data, and disclosure procedures to protect student confidentiality are also covered.

7.1 Analysis Weights

Two statistical analysis weights were computed—the first for survey respondents (undergraduates and graduates) and the second for undergraduate study respondents. Both statistical analysis weights compensate for the unequal probability of selection of institutions and students into the NPSAS:20 sample.

As described in chapter 2, NPSAS:20 was designed to sample a total of 400,000 students (375,000 undergraduate students and 25,000 graduate students); student records and administrative data were to be collected for all these 400,000 sampled students. The planned undergraduate student sample of 375,000 (the undergraduate student study sample) was designed to be both nationally representative and state representative for public 2-year and public 4-year institutions, as well as overall for undergraduates.

All graduate students in the graduate sample and a planned subsample of 125,000 undergraduate students from the undergraduate study sample were to be asked to complete a survey. This planned subsample of 125,000 undergraduate students formed the undergraduate survey sample. The graduate sample and the undergraduate survey sample were designed to be nationally representative for graduates and undergraduates, respectively.

The first statistical analysis weight was the student survey analysis weight. The student survey analysis weight was computed separately for undergraduate and graduate students who responded to the survey so that the survey respondents represent the target population as defined in chapter 2. This weight was designed for analysis of undergraduate and graduate student survey data at the national level and within each category of control and level of institution.

The second statistical analysis weight was the student study weight, computed for use with study data so that study members represent the target population of undergraduates. This weight was designed for analysis of undergraduate student administrative (non-survey) data at the state level and within the public 2-year and public 4-year institution sectors within states for which the sample is sufficient for representation. There was no graduate student study weight because the graduate sample of 25,000 students was designed to be nationally representative. For analysis of graduate student administrative data, the survey weight described above may be used.

Since separate weights were developed for undergraduate survey and study respondents,⁴⁶ the terms “survey” and “study” are used in this chapter to distinguish the elements (i.e., institutions, students, adjustments) described in the construction of the undergraduate survey and study analysis weights. Furthermore, since there is only one weight for graduate students who responded to the survey, references to the elements of the graduate weight will not include the term “survey.” Table 53 provides a summary of the two statistical analysis weights.

Table 53. NPSAS:20 analysis weights: 2019–20

Weight	Target population	Variable name	Respondent name	Respondent size	National or state analysis	When to use
Survey ¹	All study-eligible students	WTA000	Survey respondent	100,410	National	Analysis of only student survey data or survey data in conjunction with administrative data; for undergraduate or graduate students
Study	All study-eligible undergraduate students	WTB000	Study respondent	276,350	National or state	Analysis of only administrative (non-survey) data; for undergraduate students

¹ The survey weight was computed separately for undergraduate and graduate students, but the same steps were followed. The variable name (WTA000) is the same for both. However, undergraduate and graduate students are on separate files in the restricted-use file and PowerStats.

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

As discussed in section 1.3.1, some states and public 2-year and public 4-year institution sectors within states did not have enough institution and/or 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

⁴⁶ See section 7.1.1.2 for more information about respondent definitions.

about states. These criteria were a combination of objective and subjective measures that vary by state. Generally, the following informed the determination:

- the institution response rate needed to be at least 50 percent;
- the student response counts needed to meet thresholds based on power calculations;
- the percentage of student-level 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; and
- most of the essential institutions (i.e., institutions with greater student enrollment and institutions within large systems) and unique institutions (based on multiple characteristics) needed to have participated in the NPSAS:20 institution data collection.

The steps in developing the weight components are described in sections 7.1.2 and 7.1.3 and include the following:

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

Table 54 lists the components used to construct the student analysis weights. The institution weights were computed first and then used to construct the student weights. Each weight component represents a probability of selection, a weight adjustment, or an interim weight. Although the weights for undergraduates and graduates were computed separately, the steps described in this section were followed for both undergraduates and graduates.

Table 54. Components of final analysis weights: 2019–20

Undergraduate and graduate survey analysis weight		Undergraduate study analysis weight	
Component	Description	Component	Description
WTA_WT1	Institution sampling weight	WTB_WT1	Institution sampling weight
WTA_adj1	Institution nonresponse adjustment	WTB_adj1	Institution nonresponse adjustment
WTA_adj2	Institution poststratification adjustment	WTB_adj2	Institution poststratification adjustment
WTA_INSTWT	Institution weight; $WTA_INSTWT = WTA_WT1 * WTA_adj1 * WTA_adj2$	WTB_INSTWT	Institution weight; $WTB_INSTWT = WTB_WT1 * WTB_adj1 * WTB_adj2$
WTA_WT4	Student within-institution sampling weight	WTB_WT4	Student within-institution sampling weight
WTA_WT5	Student sampling weight; $WTA_WT5 = WTA_INSTWT * WTA_WT4$	WTB_WT5	Student sampling weight; $WTB_WT5 = WTB_INSTWT * WTB_WT4$
WTA_adj3	Student multiplicity adjustment	WTB_adj3	Student multiplicity adjustment
WTA_adj4	Student unknown eligibility status adjustment	WTB_adj4	Student unknown eligibility status adjustment
WTA_WT7	Student base weight; $WTA_WT7 = WTA_WT5 * WTA_adj3 * WTA_adj4$	WTB_WT7	Student base weight; $WTB_WT7 = WTB_WT5 * WTB_adj3 * WTB_adj4$
WTA_adj5a	Student-not-located nonresponse adjustment	†	†
WTA_adj5b	Student refusal nonresponse adjustment	†	†
WTA_adj5c	Student other nonresponse adjustment	†	†
WTA_adj5	Student composite nonresponse adjustment; $WTA_adj5 = WTA_adj5a * WTA_adj5b * WTA_adj5c$	WTB_adj5	Student nonresponse adjustment
WTA_adj6	Student poststratification adjustment	WTB_adj6	Student poststratification adjustment
WTA000	Final survey analysis weight; $WTA000 = WTA_WT7 * WTA_adj5 * WTA_adj6$	WTB000	Final undergraduate study analysis weight; $WTB000 = WTB_WT7 * WTB_adj5 * WTB_adj6$

† Not applicable.

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

All nonresponse and poststratification adjustments were computed using the WTADJUST procedure in SUDAAN. The WTADJUST procedure used a constrained logistic model to predict response. A key feature of the WTADJUST procedure is that the weight adjustments and weight trimming and smoothing are all accomplished in one step. Bounds can be specified on the weight going into the WTADJUST procedure, and bounds can be specified on the weight adjustment factors that are output by the WTADJUST procedure. The exact formula for the weight adjustment factors calculated by the SUDAAN WTADJUST procedure is in the *SUDAAN Language Manual, Release 11.0* (RTI International 2012).

During model refinement, which involved collapsing categories of candidate predictor variables and/or excluding candidate predictor variables so that models would converge, no upper or lower bounds were specified. Once model

convergence was achieved, the weight adjustment bounds were tightened to reduce the occurrence of extreme weights and the variance inflation caused by unequal weighting (referred to as the unequal weighting effect, or UWE).

The UWE is a measure of unequal weighting on the variance. When there is no variation in weights (as with a simple random sample, or SRS, design), the UWE equals one. 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 adjusted weight.

UWEs at each weight adjustment stage were examined, overall and by stratum, before and after the adjustment to evaluate the effect of each weighting adjustment on the precision of survey estimates. If an adjustment resulted in a much larger UWE compared to the UWE before the adjustment, weight adjustment bounds were tightened to reduce the loss of precision.

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, the UWEs, and the minimum and maximum adjustment factors.

7.1.1 Unit Response Rates

This section provides institution and student respondent definitions and summarizes response rates overall and by control and level of institution.

7.1.1.1 Institution response rates

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

Of the 3,110 sampled institutions, 3,070 were determined to be eligible (see section 2.1.1 for more information about eligibility). Of the 3,010 eligible sample institutions that enrolled undergraduate students, 2,150 provided enrollment lists (71.5 unweighted percent, 63.9 weighted percent) from which a student was selected for the study, and 2,120 provided enrollment lists (70.7 unweighted percent, 63.0 weighted percent) from which a survey student was selected. Of the

1,410 eligible sample institutions that enrolled graduate students, 1,090 provided enrollment lists (77.2 unweighted percent, 75.4 weighted percent).

Institution response rates, by control and level of institution, for undergraduate- and graduate-enrolling institutions are shown in table 55. Institution response rates for undergraduate-enrolling study institutions, by state and institution stratum, are shown in table 56.

Table 55. Institution response rates for undergraduate- and graduate-enrolling institutions, by collection type and by control and level of institution: 2019–20

Control and level of institution	All institutions		Undergraduate-enrolling study institutions			Undergraduate-enrolling survey institutions			Graduate-enrolling institutions			
	Number of eligible sampled institutions	Number of respondents ¹	Response rates			Response rates			Number of eligible sampled institutions	Number of respondents ¹	Response rates	
			Number of respondents ¹	Un-weighted	Weighted	Number of respondents ¹	Un-weighted	Weighted			Un-weighted	Weighted
Total	3,070	3,010	2,150	71.5	63.9	2,120	70.7	63.0	1,410	1,090	77.2	75.4
Public less-than-2-year	40	40	20	53.8	68.5	20	51.3	68.2	†	†	†	†
Public 2-year	960	940	710	75.1	75.2	690	73.7	73.8	†	†	†	†
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	150	170	130	75.4	75.4	130	74.9	74.9	70	60	89.2	89.2
Public 4-year, non-doctorate-granting, primarily baccalaureate	230	220	180	81.2	81.2	180	81.2	81.2	180	150	82.3	82.3
Public 4-year, doctorate-granting	380	370	330	87.4	87.4	320	86.6	86.6	390	340	87.1	87.1
Private nonprofit 2-year or less	30	30	10	43.8	58.8	10	43.8	58.8	†	†	†	†
Private nonprofit 4-year, non-doctorate-granting	390	360	260	71.7	71.2	260	70.6	67.8	270	180	68.2	66.1
Private nonprofit 4-year, doctorate-granting	390	380	290	76.6	78.0	290	75.8	77.3	410	300	74.7	73.3
Private for-profit less-than-2-year	220	240	100	42.2	38.9	100	41.8	38.6	†	†	†	†
Private for-profit 2-year	160	140	70	49.7	49.7	70	49.7	49.7	†	†	†	†
Private for-profit 4-year	120	120	60	51.3	62.3	60	51.3	62.3	100	60	55.6	74.3

† 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

Table 56. Institution response rates for undergraduate-enrolling study institutions, by state and institution stratum: 2019–20

State and institution stratum	Undergraduate-enrolling study institutions		
	Number of eligible sampled institutions	Response rates ¹	
		Unweighted	Weighted
Alabama			
Public 2-year	20	91.7	91.7
Public 4-year	10	78.6	78.6
Other	30	55.2	37.2
Alaska			
Public 2-year	†	†	†
Public 4-year	#	75.0	75.0
Other	10	80.0	80.0
Arizona			
Public 2-year	20	80.0	80.0
Public 4-year	10	100.0	100.0
Other	30	51.6	39.8
Arkansas			
Public 2-year	20	81.8	81.8
Public 4-year	10	81.8	81.8
Other	30	42.9	36.3
California			
Public 2-year	110	69.5	69.5
Public 4-year	50	72.3	72.3
Other	30	63.3	27.3
Colorado			
Public 2-year	10	100.0	100.0
Public 4-year	20	95.2	95.2
Other	30	48.1	60.6
Connecticut			
Public 2-year	10	100.0	100.0
Public 4-year	10	50.0	50.0
Other	30	65.5	62.8
Delaware			
Public 2-year	†	†	†
Public 4-year	#	100.0	100.0
Other	10	42.9	42.9
District of Columbia			
Public 2-year	†	†	†
Public 4-year	#	100.0	100.0
Other	20	66.7	66.7
Florida			
Public 2-year	30	64.5	66.9
Public 4-year	40	83.3	83.3
Other	30	72.4	86.8
Georgia			
Public 2-year	20	73.9	73.9
Public 4-year	30	100.0	100.0
Other	30	71.4	54.9
Hawaii			
Public 2-year	10	100.0	100.0
Public 4-year	#	100.0	100.0
Other	10	54.5	54.5

See notes at end of table.

Table 56. Institution response rates for undergraduate-enrolling study institutions, by state and institution stratum: 2019–20—Continued

State and institution stratum	Undergraduate-enrolling study institutions		
	Number of eligible sampled institutions	Response rates ¹	
		Unweighted	Weighted
Idaho			
Public 2-year	#	33.3	33.3
Public 4-year	10	80.0	80.0
Other	30	65.5	65.5
Illinois			
Public 2-year	50	81.3	81.3
Public 4-year	10	91.7	91.7
Other	30	82.8	94.4
Indiana			
Public 2-year	#	100.0	100.0
Public 4-year	20	53.3	53.3
Other	30	79.3	74.3
Iowa			
Public 2-year	20	75.0	75.0
Public 4-year	#	100.0	100.0
Other	30	60.0	42.0
Kansas			
Public 2-year	30	40.0	40.0
Public 4-year	10	75.0	75.0
Other	30	55.2	38.6
Kentucky			
Public 2-year	20	68.8	68.8
Public 4-year	10	87.5	87.5
Other	30	78.6	67.2
Louisiana			
Public 2-year	10	71.4	71.4
Public 4-year	20	75.0	75.0
Other	30	57.1	45.6
Maine			
Public 2-year	10	57.1	57.1
Public 4-year	10	100.0	100.0
Other	20	78.9	78.9
Maryland			
Public 2-year	20	68.8	68.8
Public 4-year	10	100.0	100.0
Other	30	48.3	42.6
Massachusetts			
Public 2-year	20	93.8	93.8
Public 4-year	10	100.0	100.0
Other	30	69.0	89.1
Michigan			
Public 2-year	20	75.0	75.0
Public 4-year	20	86.4	86.4
Other	30	57.7	32.8
Minnesota			
Public 2-year	30	93.8	93.8
Public 4-year	10	100.0	100.0
Other	30	69.0	65.3

See notes at end of table.

Table 56. Institution response rates for undergraduate-enrolling study institutions, by state and institution stratum: 2019–20—Continued

State and institution stratum	Undergraduate-enrolling study institutions		
	Number of eligible sampled institutions	Response rates ¹	
		Unweighted	Weighted
Mississippi			
Public 2-year	20	86.7	86.7
Public 4-year	10	62.5	62.5
Other	30	28.6	28.6
Missouri			
Public 2-year	20	47.1	47.1
Public 4-year	10	69.2	69.2
Other	30	64.3	52.6
Montana			
Public 2-year	10	70.0	70.0
Public 4-year	10	87.5	87.5
Other	10	50.0	50.0
Nebraska			
Public 2-year	10	88.9	88.9
Public 4-year	10	42.9	42.9
Other	30	77.8	77.8
Nevada			
Public 2-year	†	†	†
Public 4-year	10	100.0	100.0
Other	30	48.3	48.3
New Hampshire			
Public 2-year	10	42.9	42.9
Public 4-year	10	100.0	100.0
Other	20	43.5	43.5
New Jersey			
Public 2-year	20	84.2	84.2
Public 4-year	10	100.0	100.0
Other	30	79.3	81.0
New Mexico			
Public 2-year	20	84.2	84.2
Public 4-year	10	66.7	66.7
Other	20	56.3	56.3
New York			
Public 2-year	40	75.7	75.7
Public 4-year	40	90.7	90.7
Other	30	75.0	80.2
North Carolina			
Public 2-year	60	82.8	82.8
Public 4-year	20	88.2	88.2
Other	30	80.6	85.8
North Dakota			
Public 2-year	10	20.0	20.0
Public 4-year	10	44.4	44.4
Other	10	35.7	35.7
Ohio			
Public 2-year	20	91.3	91.3
Public 4-year	40	89.7	89.7
Other	30	66.7	60.8

See notes at end of table.

Table 56. Institution response rates for undergraduate-enrolling study institutions, by state and institution stratum: 2019–20—Continued

State and institution stratum	Undergraduate-enrolling study institutions		
	Number of eligible sampled institutions	Response rates ¹	
		Unweighted	Weighted
Oklahoma			
Public 2-year	20	72.7	72.7
Public 4-year	20	81.3	81.3
Other	30	53.3	45.5
Oregon			
Public 2-year	20	58.8	58.8
Public 4-year	10	88.9	88.9
Other	30	51.7	53.6
Pennsylvania			
Public 2-year	20	72.2	72.2
Public 4-year	40	90.5	90.5
Other	30	76.7	77.4
Puerto Rico			
Public 2-year	10	100.0	100.0
Public 4-year	10	85.7	85.7
Other	30	80.0	44.9
Rhode Island			
Public 2-year	#	100.0	100.0
Public 4-year	#	0.0	0.0
Other	20	47.1	47.1
South Carolina			
Public 2-year	20	47.4	47.4
Public 4-year	10	78.6	78.6
Other	30	85.7	91.4
South Dakota			
Public 2-year	10	60.0	60.0
Public 4-year	10	75.0	75.0
Other	10	58.3	58.3
Tennessee			
Public 2-year	40	82.1	82.1
Public 4-year	10	90.0	90.0
Other	30	70.0	66.7
Texas			
Public 2-year	60	78.6	78.6
Public 4-year	50	86.0	86.0
Other	30	64.3	31.7
Utah			
Public 2-year	#	0.0	0.0
Public 4-year	10	85.7	85.7
Other	30	63.3	45.1
Vermont			
Public 2-year	#	100.0	100.0
Public 4-year	#	50.0	50.0
Other	10	53.8	53.8
Virginia			
Public 2-year	20	87.5	87.5
Public 4-year	20	93.3	93.3
Other	30	51.7	31.8

See notes at end of table.

Table 56. Institution response rates for undergraduate-enrolling study institutions, by state and institution stratum: 2019–20—Continued

State and institution stratum	Undergraduate-enrolling study institutions		
	Number of eligible sampled institutions	Response rates ¹	
		Unweighted	Weighted
Washington			
Public 2-year	10	83.3	83.3
Public 4-year	40	62.2	62.2
Other	30	78.6	68.4
West Virginia			
Public 2-year	10	41.7	41.7
Public 4-year	10	91.7	91.7
Other	30	40.0	29.8
Wisconsin			
Public 2-year	20	62.5	62.5
Public 4-year	20	66.7	66.7
Other	30	48.3	37.9
Wyoming			
Public 2-year	10	85.7	85.7
Public 4-year	#	100.0	100.0
Other	#	50.0	50.0

† 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.

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

7.1.1.2 Student response rates

The rules for defining respondents depended on the data elements that were collected from the surveys; student records; and administrative, federal, and private databases including the CPS, NSLDS, and NSC. Separate rules were established for undergraduate and graduate students.

Respondents were defined as follows:

- **Survey respondents.** A survey respondent was defined as any undergraduate or graduate sample member who was determined to be eligible for the study and completed at least the enrollment and FAFSA sections of the student survey. The weight for the survey respondents is the survey weight (WTA000).
- **Administrative student respondents.** An administrative student respondent was defined as any undergraduate sample member who was determined to be eligible for the study, was enrolled for at least 1 month (based on student records or administrative sources), and at a minimum, had valid data from student records for the following items:
 - federal work-study amount awarded;

- state aid recipient indicator;
- state program 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.

There is not a separate weight for administrative student respondents. The definition of administrative student respondents is used as a component in the study respondent definition.

- **Study respondents.** A study respondent was defined as any undergraduate sample member who was an administrative student respondent and/or survey respondent.⁴⁷ Study respondents are all undergraduates; there are no graduate study respondents. The weight for the study respondents is the study weight (WTB000).

Of the 140,080 eligible undergraduate survey students, 80,760 were respondents (58 unweighted percent, 62 weighted percent). Of the 27,260 eligible graduate students, 19,650 were respondents (72 unweighted percent, 71 weighted percent). Of the 345,640 eligible undergraduate study students, 276,350 were respondents (80 unweighted percent, 84 weighted percent). Student response rates, by control and level of institution, for undergraduate and graduate students are shown in table 57. Student response rates for undergraduate study students, by state and institution stratum are shown in table 58.

⁴⁷ Administrative data were imputed for study respondents who were survey respondents but not administrative student respondents (see section 7.4).

Table 57. Student response rates for undergraduate and graduate students, by collection type and by control and level of institution: 2019–20

Control and level of institution	Undergraduate survey students				Graduate students				Undergraduate study students			
	Number of eligible sampled students	Number of respondents	Response rates		Number of eligible sampled students	Number of respondents	Response rates		Number of eligible sampled students	Number of respondents	Response rates	
			Un-weighted	Weighted ¹			Un-weighted	Weighted ¹			Un-weighted	Weighted ¹
Total	140,080	80,760	57.7	61.5	27,260	19,650	72.1	71.3	345,640	276,350	80	83.7
Public less-than-2-year	1,390	660	47.9	55.4	†	†	†	†	2,450	2,240	91.6	96
Public 2-year	49,670	25,940	52.2	53.9	†	†	†	†	130,460	97,650	74.8	75.4
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	9,140	5,020	54.9	57.1	130	110	79.9	88.1	19,030	15,720	82.6	87.9
Public 4-year, non-doctorate-granting, primarily baccalaureate	8,620	5,650	65.5	67.3	1,820	1,260	69.6	71.7	21,220	16,390	77.2	81.2
Public 4-year, doctorate-granting	27,070	17,240	63.7	66.5	9,630	7,050	73.2	71.9	87,300	75,930	87	89.5
Private nonprofit 2-year or less	1,340	650	48.5	54.2	†	†	†	†	1,700	1,380	80.8	83.7
Private nonprofit 4-year, non-doctorate-granting	8,990	6,230	69.3	71.3	1,800	1,360	75.6	76.8	23,340	18,990	81.4	86.4
Private nonprofit 4-year, doctorate-granting	11,510	7,620	66.2	67.5	9,050	6,420	70.9	69.2	32,360	27,010	83.5	88.9
Private for-profit less-than-2-year	5,110	2,470	48.3	47.3	†	†	†	†	7,120	4,880	68.6	70.3
Private for-profit 2-year	7,450	3,390	45.5	47.8	†	†	†	†	8,460	5,680	67.1	72.3
Private for-profit 4-year	9,790	5,890	60.2	61.3	4,820	3,450	71.5	73.1	12,200	10,500	86	87.5

† Not applicable.

¹ The weighted response rates were calculated using the student base weight (the student sampling weight adjusted for multiplicity and unknown eligibility).

NOTE: Number of eligible sampled students and number of respondents rounded to the nearest 10. Response rates 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

Table 58. Student response rates for undergraduate study students, by state and institution stratum: 2019–20

State and institution stratum	Number of eligible sampled students	Number of respondents	Response rates	
			Unweighted	Weighted ¹
Alabama				
Overall	6,970	5,490	78.7	86.7
Public 2-year	3,420	2,180	63.9	67.6
Public 4-year	2,310	2,180	94.6	96
Alaska ²				
Overall	1,500	970	64.5	87.1
Public 2-year	†	†	†	†
Public 4-year	1,010	760	75.6	87.9
Arizona				
Overall	6,960	5,360	77.1	82.4
Public 2-year	2,920	2,130	72.9	77.8
Public 4-year	1,630	1,540	94.9	97.3
Arkansas				
Overall	5,090	3,880	76.3	80.8
Public 2-year	2,340	2,040	87.3	88.7
Public 4-year	1,650	1,170	70.6	79.8
California				
Overall	24,290	18,040	74.3	76.2
Public 2-year	13,970	9,490	68	66.5
Public 4-year	6,790	6,010	88.5	87.1
Colorado				
Overall	5,650	4,650	82.3	84.6
Public 2-year	800	620	78.1	92.4
Public 4-year	3,580	3,080	86.1	88.1
Connecticut				
Overall	6,100	5,260	86.3	88.8
Public 2-year	2,830	2,420	85.6	75
Public 4-year	840	840	99.9	99.9
Delaware ²				
Overall	2,430	1,130	46.5	34.3
Public 2-year	†	†	†	†
Public 4-year	1,590	890	55.8	40.8
District of Columbia ²				
Overall	3,120	1,630	52.2	73.5
Public 2-year	†	†	†	†
Public 4-year	1,300	360	27.5	27.8
Florida				
Overall	13,400	10,200	76.2	87.4
Public 2-year	2,920	1,460	50.2	76.1
Public 4-year	6,720	5,960	88.7	89.4
Georgia				
Overall	10,350	7,350	71	73.5
Public 2-year	3,480	1,680	48.4	47.4
Public 4-year	4,810	3,800	79	79.4
Hawaii				
Overall	4,150	4,070	98.2	98.3
Public 2-year	1,650	1,610	97.6	98.1
Public 4-year	1,470	1,460	99	99.1

See notes at end of table.

Table 58. Student response rates for undergraduate study students, by state and institution stratum: 2019–20—Continued

State and institution stratum	Number of eligible sampled students	Number of respondents	Response rates	
			Unweighted	Weighted ¹
Idaho				
Overall	3,930	2,000	50.9	39.6
Public 2-year	470	140	30.5	31.8
Public 4-year	1,100	860	78.2	74.9
Illinois				
Overall	9,760	8,160	83.7	87.3
Public 2-year	6,000	4,730	78.9	82.3
Public 4-year	2,010	1,940	96.4	95.6
Indiana				
Overall	6,890	6,450	93.6	96.3
Public 2-year	2,610	2,500	95.8	95.5
Public 4-year	2,590	2,580	99.3	99.6
Iowa				
Overall	3,720	3,530	94.9	95.9
Public 2-year	1,540	1,420	91.8	95.6
Public 4-year	1,040	1,030	99	99.1
Kansas				
Overall	4,130	2,900	70.1	74.1
Public 2-year	1,910	800	42.1	43.2
Public 4-year	1,300	1,260	96.5	97.7
Kentucky				
Overall	4,790	4,320	90.2	92.3
Public 2-year	1,810	1,530	84.8	85.3
Public 4-year	1,350	1,320	97.5	98.2
Louisiana				
Overall	6,060	4,760	78.5	80.3
Public 2-year	2,260	1,320	58.5	57.9
Public 4-year	2,500	2,290	91.4	92.1
Maine				
Overall	4,710	4,240	90	91.2
Public 2-year	910	900	99.3	99.6
Public 4-year	1,560	1,370	88.4	88.8
Maryland				
Overall	5,860	4,570	78	80.4
Public 2-year	2,750	1,710	62.3	61.6
Public 4-year	2,050	1,890	92.3	95.6
Massachusetts				
Overall	7,260	5,990	82.5	85.3
Public 2-year	2,130	1,710	80	83.1
Public 4-year	2,180	1,940	88.9	91.3
Michigan				
Overall	7,800	6,720	86.2	91.1
Public 2-year	3,380	2,850	84.5	87.9
Public 4-year	3,130	2,820	89.9	93.8
Minnesota				
Overall	9,030	8,620	95.4	94.9
Public 2-year	5,110	4,970	97.2	96.8
Public 4-year	2,280	2,250	99	99.3

See notes at end of table.

Table 58. Student response rates for undergraduate study students, by state and institution stratum: 2019–20—Continued

State and institution stratum	Number of eligible sampled students	Number of respondents	Response rates	
			Unweighted	Weighted ¹
Mississippi				
Overall	3,620	3,140	86.7	89.1
Public 2-year	1,480	1,260	84.9	83.5
Public 4-year	1,280	1,270	99.3	99.2
Missouri				
Overall	5,540	4,260	76.9	82.5
Public 2-year	1,730	1,170	67.7	80.3
Public 4-year	1,890	1,550	82	85.3
Montana				
Overall	3,090	1,230	39.9	31
Public 2-year	1,160	410	35.1	29.7
Public 4-year	1,490	430	28.5	24.4
Nebraska				
Overall	3,610	3,260	90.3	93.6
Public 2-year	930	920	99.4	99.9
Public 4-year	1,050	1,030	97.7	97.1
Nevada ²				
Overall	3,040	2,540	83.4	92.2
Public 2-year	†	†	†	†
Public 4-year	1,600	1,480	92.6	93.4
New Hampshire				
Overall	4,140	3,310	80	95.5
Public 2-year	1,040	330	31.4	45.4
Public 4-year	1,450	1,450	99.9	99.9
New Jersey				
Overall	8,780	7,030	80.1	79.6
Public 2-year	3,270	2,470	75.6	73.1
Public 4-year	2,700	2,330	86.5	85.6
New Mexico				
Overall	4,330	4,110	94.8	96.7
Public 2-year	2,190	2,060	94	96.5
Public 4-year	1,390	1,330	95.3	96.8
New York				
Overall	16,110	11,060	68.6	76.3
Public 2-year	5,190	3,400	65.5	64.3
Public 4-year	8,470	5,380	63.5	64.8
North Carolina				
Overall	12,470	11,110	89.1	89.8
Public 2-year	7,980	7,080	88.7	87.7
Public 4-year	2,510	2,310	92.2	94.2
North Dakota				
Overall	1,780	1,430	80.5	87.6
Public 2-year	390	390	100.0	100.0
Public 4-year	730	700	96.1	91.7
Ohio				
Overall	11,490	9,980	86.8	89.6
Public 2-year	4,030	3,890	96.5	98
Public 4-year	5,860	4,590	78.3	81.6

See notes at end of table.

Table 58. Student response rates for undergraduate study students, by state and institution stratum: 2019–20—Continued

State and institution stratum	Number of eligible sampled students	Number of respondents	Response rates	
			Unweighted	Weighted ¹
Oklahoma				
Overall	7,070	5,580	78.9	83
Public 2-year	2,860	2,050	71.6	67
Public 4-year	2,820	2,510	89	93.6
Oregon				
Overall	4,490	4,110	91.6	91.2
Public 2-year	1,850	1,680	90.6	88.4
Public 4-year	1,430	1,410	98.4	95.5
Pennsylvania				
Overall	10,880	7,880	72.5	79.2
Public 2-year	2,430	1,980	81.5	81.3
Public 4-year	6,100	4,140	67.8	74.8
Puerto Rico				
Overall	6,500	3,970	61	84.4
Public 2-year	1,560	#	0	0
Public 4-year	1,820	1,390	76.5	79.6
Rhode Island				
Overall	3,060	1,490	48.5	48.9
Public 2-year	1,640	220	13.1	13.6
Public 4-year ³	†	†	†	†
South Carolina				
Overall	5,530	4,480	81.1	87.8
Public 2-year	1,860	1,090	58.5	70.8
Public 4-year	1,860	1,800	97.1	98.3
South Dakota				
Overall	2,960	2,190	74	77.7
Public 2-year	830	830	99.9	99.9
Public 4-year	1,360	980	71.5	76.5
Tennessee				
Overall	8,870	7,250	81.7	84.9
Public 2-year	5,680	4,620	81.4	81.9
Public 4-year	1,530	1,300	85.1	85.8
Texas				
Overall	17,490	15,050	86.1	86.6
Public 2-year	6,990	5,490	78.6	78.1
Public 4-year	8,420	7,970	94.6	94.5
Utah				
Overall	5,330	4,890	91.7	94.6
Public 2-year	†	†	†	†
Public 4-year	1,270	1,130	89.5	89
Vermont				
Overall	3,470	3,430	98.7	99.4
Public 2-year	1,280	1,280	100.0	100.0
Public 4-year	760	760	100.0	100.0
Virginia				
Overall	6,410	5,610	87.5	90.4
Public 2-year	2,940	2,400	81.5	85.2
Public 4-year	2,570	2,410	93.8	95.2

See notes at end of table.

Table 58. Student response rates for undergraduate study students, by state and institution stratum: 2019–20—Continued

State and institution stratum	Number of eligible sampled students	Number of respondents	Response rates	
			Unweighted	Weighted ¹
Washington				
Overall	7,920	6,070	76.7	76.4
Public 2-year	1,460	940	64.3	61.2
Public 4-year	4,880	3,620	74.1	75.1
West Virginia				
Overall	5,750	4,790	83.3	93.8
Public 2-year	1,160	1,160	100.0	100.0
Public 4-year	2,210	1,880	85.1	92.8
Wisconsin				
Overall	5,300	4,800	90.5	95.1
Public 2-year	2,080	1,650	79.5	84.4
Public 4-year	2,010	2,010	100.0	99.9
Wyoming				
Overall	2,700	2,070	76.5	72.6
Public 2-year	1,320	690	52.1	54.8
Public 4-year	1,290	1,290	100.0	100.0

† Not applicable.

Rounds to zero.

¹ The weighted response rates were calculated using the student base weight (the student sampling weight adjusted for multiplicity and unknown eligibility).

² There were no public 2-year schools in Alaska, Delaware, the District of Columbia, or Nevada.

³ There were no responding public 4-year schools in Rhode Island.

NOTE: Number of eligible sampled students and number of respondents rounded to the nearest 10. Response rates 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

7.1.2 Institution Weights

Institution weights were created for the purpose of constructing student-level weights rather than for analyzing institution data. Institution weights to be used for the student survey weight were computed separately for institutions enrolling undergraduate students and for institutions enrolling graduate students. This is similar to what was done in NPSAS:18-AC. The institution weights were computed for three groups of respondents:

- institution respondents corresponding to the undergraduate student survey weight;
- institution respondents corresponding to the graduate student weight; and
- institution respondents corresponding to the undergraduate student study weight.

Institutions were included in one, two, or all three of these institution weight computations, but each institution weight is associated with the student survey weight for undergraduate students, the student survey weight for graduate students, or the student study weight.

Each 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 in the fourth step, the final institution weight was computed.

7.1.2.1 *Institution sampling weights (WTA_WT1 and WTB_WT1)*

The first weight component is the institution sampling weight, which is associated with the selection process of the NPSAS:20 institution sample described in chapter 2.

The full sample of 3,110 institutions consisted of a census of all public 2-year and all public 4-year institutions, as well as a sample of 1,370 institutions from the “all other sectors” stratum. The sampling weight for each sample institution is the reciprocal of its probability of selection. The probability of selection for institution i was as follows:

$$\pi(i) = \begin{cases} \frac{nS(i)}{S(+)} & \text{for noncertainty selections in the “all other sectors” stratum} \\ 1 & \text{for certainty selections,}^{48} \end{cases}$$

where n = the sample size in the “all other sectors” stratum; $S(i)$ = the measure of size for the i th institution in the “all other sectors” stratum; and $S(+)$ = the total measure of size of all institutions in the “all other sectors” stratum. Thus, the institution sampling weights were computed as follows:

$$WTA_WT1 = WTB_WT1 = 1 / \pi(i).$$

During institution recruitment and student list sampling, a few institutions were identified that had two or more records listed on the IPEDS frame. In some cases, having two or more records was caused by institutions that had recently merged. However, for most cases, the sample institution sent one student list covering multiple campuses, and the sampling team treated the campuses as merged for sampling purposes. If two or more records were sampled for the same institution, then one record was retained as a participant and the other records were classified as ineligible.

⁴⁸ 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.

When an institution had multiple chances of selection (as described above), the institution sampling weight was revised by estimating, as if the selections were independent, the probability that either record could be selected:

$$P(A \text{ or } B \text{ or } C \text{ or } D \dots) = 1 - [1 - P(A)] * [1 - P(B)] * [1 - P(C)] * [1 - P(D)] \dots$$

Then, the revised sampling weight for merged institutions was calculated as the reciprocal of this probability:

$$WTA_WT1 = WTB_WT1 = 1 / P(A \text{ or } B \text{ or } C \text{ or } D \dots).$$

7.1.2.2 Institution nonresponse adjustments (*WTA_adj1* and *WTB_adj1*)

The institution nonresponse adjustment models for both undergraduate- and graduate-enrolling institutions were developed in the following three stages:

1. identification of model predictors;
2. chi-square automatic interaction detection (CHAID) analysis to determine significant interactions between model predictors; and
3. WTADJUST procedure to calculate nonresponse adjustment.

Candidate predictor variables for the institution nonresponse adjustment models were selected because they were thought to predict response status based on knowledge of NPSAS data and were nonmissing for most respondents and nonrespondents. If needed, missing data were recoded into an “unknown” category when there was a small percentage of missing data.

In addition to the candidate predictor variables, potentially important two-way and three-way interactions between predictor variables were included in the nonresponse adjustment models. To identify these interactions, the 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 records and cycles over each predictor, finding for each predictor an optimal partition of the records 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.

The weights used in the institution nonresponse adjustment models were *WTA_WT1* and *WTB_WT1* multiplied by the institution’s undergraduate or

graduate enrollment totals determined by the IPEDS 2019–20 Institutional Characteristics Header (IC-H), Institutional Characteristics (IC), and 12-Month Enrollment (E12) files.

Table 59 lists the predictor variables that were used in the institution nonresponse adjustment models for both undergraduate- and graduate-enrolling institutions. The table indicates whether the predictor variables were included in the nonresponse adjustment model for the undergraduate survey, graduate, and/or undergraduate study institution respondents.

Table 59. Candidate predictor variables for institution nonresponse adjustment models: 2019–20

Nonresponse adjustment model predictor variables ¹	Undergraduate survey	Graduate	Undergraduate study
Geographic variables			
Degree of urbanization ²	✓	✓	✓
Region of institution	✓	✓	✓
State of institution	✓		✓
Institution characteristics			
2015 Carnegie Basic classification	✓	✓	✓
Historically Black College or University (HBCU) status (yes/no)	✓	✓	✓
Hispanic-Serving Institution (HSI) status (yes/no) ³	✓	✓	✓
Control and level of institution	✓	✓	
Public 2-year and public 4-year institution sectors within state of institution			✓
Average net price among first-time, full-time undergraduate students receiving grant or scholarship aid ⁴	✓		✓
Total institution employee count of office and administrative support staff ⁴		✓	
Variables related to the institutions' student population ⁴			
Total undergraduate or graduate student enrollment	✓	✓	✓
Male undergraduate or graduate student enrollment	✓	✓	✓
Female undergraduate or graduate student enrollment	✓	✓	✓
Percentage of undergraduate or graduate students enrolled who were Asian or Pacific Islander, non-Hispanic ⁵	✓	✓	✓
Percentage of undergraduate or graduate students enrolled who were Black, non-Hispanic	✓	✓	✓
Percentage of undergraduate or graduate students enrolled who were Hispanic	✓	✓	✓
Number of first-time, full-time undergraduate students living on campus who were receiving Title IV aid	✓		✓
Number of first-time, full-time undergraduate students with incomes up to \$30,000 who were receiving Title IV aid	✓		✓
Percentage of first-time, full-time undergraduate students receiving federal grant aid	✓		✓
Percentage of first-time, full-time undergraduate students receiving state or local grant aid	✓		✓
Percentage of first-time, full-time undergraduate students receiving institution grant aid	✓		✓
Percentage of first-time, full-time undergraduate students receiving student loan aid	✓		✓
Average amount of grant and scholarship aid received by undergraduate students ⁶	✓		✓
Percentage of undergraduate students receiving any grant aid ⁶	✓		✓

¹ Institution-level nonresponse adjustment model candidate predictor variables came from Integrated Postsecondary Education Data System (IPEDS) 2019–20 files.

² Degree of urbanization is an IPEDS variable representing the urbanicity (city/suburb/rural) by population size of the institution's location.

³ Of the listed variables, only the Hispanic-Serving Institution (HSI) indicator no longer exists in IPEDS. An HSI proxy was created following the definition of HSI provided by the U.S. Department of Education (<https://www2.ed.gov/programs/ideshs/definition.html>) and using IPEDS Hispanic enrollment data.

⁴ Categories were defined by quartiles for continuous variables. Categories of "None" or "Unknown" were excluded from quartile calculations.

⁵ Asian or Pacific Islander, non-Hispanic includes Native Hawaiian.

⁶ These variables were considered for the institution nonresponse weight adjustment models but were excluded from the final models.

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

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.

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 500 was set for undergraduate-enrolling survey institutions. The institution nonresponse weight adjustment factors for the graduate-enrolling institutions and the undergraduate-enrolling study institutions did not have an upper bound.

Table 60 and table 61 show the average weight adjustment and the final model predictor variables used to compute the survey institution nonresponse weight in the undergraduate- and graduate-enrolling institution models, respectively.

Summary statistics of the survey nonresponse weight adjustment factors follow:

- minimum: 1.00 for both undergraduate-enrolling survey and graduate-enrolling institutions;
- median: 1.24 for undergraduate-enrolling survey institutions and 1.08 for graduate-enrolling institutions; and
- maximum: 15.72 for undergraduate-enrolling survey institutions and 32.54 for graduate-enrolling institutions.

Table 60. Weight adjustment factors for institution nonresponse adjustment for undergraduate-enrolling survey institutions, by model predictor variable: 2019–20

Model predictor variable	Number of respondents ¹	Weighted response rate ²	Average weight adjustment factor (WTA_adj1)
Total	2,120	77.8	1.47
Control and level of institution			
Public less-than-2-year	20	68.9	1.81
Public 2-year	690	75.7	1.34
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	130	78.4	1.30
Public 4-year, non-doctorate-granting, primarily baccalaureate	180	84.2	1.19
Public 4-year, doctorate-granting	320	86.3	1.17
Private nonprofit 2-year or less	10	39.9	5.81
Private nonprofit 4-year, non-doctorate-granting	260	76.6	1.46
Private nonprofit 4-year, doctorate-granting	290	76.6	1.35
Private for-profit less-than-2-year	100	48.3	2.37
Private for-profit 2-year	70	46.8	2.53
Private for-profit 4-year	60	64.8	2.62

See notes at end of table.

Table 60. Weight adjustment factors for institution nonresponse adjustment for undergraduate-enrolling survey institutions, by model predictor variable: 2019–20—Continued

Model predictor variable	Number of respondents ¹	Weighted response rate ²	Average weight adjustment factor (WTA_adj1)
Carnegie classification code			
Associate's	780	74.8	1.44
Research and doctoral	250	84.2	1.21
Master's	450	81.7	1.31
Baccalaureate	290	78.8	1.36
Special focus and other	130	66.3	1.64
Unavailable or unknown	230	56.5	2.23
Degree of urbanization			
Large city	420	81.6	1.48
Midsize city	230	75.8	1.58
Small city	290	72.4	1.69
Large suburb	340	77.2	1.46
Midsize suburb	70	80.9	1.26
Small suburb	60	90.9	1.12
Urban area on fringe of town	50	82.9	1.19
Urban area distant from town	220	77.4	1.40
Urban area remote from town	190	86.6	1.18
Rural area on fringe of town	170	66.1	1.62
Rural area distant from town	50	68.3	1.48
Rural area remote from town	30	59.7	1.71
Bureau of Economic Analysis (BEA) region ³			
New England	150	78.9	1.45
Mideast	270	84.3	1.30
Great Lakes	270	78.3	1.39
Plains	220	74.8	1.46
Southeast	600	78.9	1.45
Southwest	220	79.1	1.52
Rocky Mountains	120	78.3	1.40
Far West	250	71.8	1.83
Outlying Areas	40	74.4	1.28
Historically Black College or University			
Yes	70	89.7	1.18
No, unavailable, or unknown	2,050	77.7	1.48
Hispanic-Serving Institution			
Yes	420	76.2	1.50
No	1,710	78.6	1.46
Total undergraduate enrollment			
1–934	390	54.7	2.06
935–2,969	540	71.4	1.47
2,970–8,383	590	78.3	1.29
8,384–162,646	590	79.8	1.27
Total male undergraduate enrollment			
0–353	400	59.5	2.00
354–1,183	550	71.2	1.47
1,184–3,507	590	77.8	1.31
3,508–72,347	590	79.8	1.28

See notes at end of table.

Table 60. Weight adjustment factors for institution nonresponse adjustment for undergraduate-enrolling survey institutions, by model predictor variable: 2019–20—Continued

Model predictor variable	Number of respondents ¹	Weighted response rate ²	Average weight adjustment factor (WTA_adj1)
Total female undergraduate enrollment			
0–531	390	50.6	2.19
532–1,733	540	72.4	1.38
1,734–4,848	590	77.8	1.29
4,849–90,299	600	80.0	1.26
Percentage of undergraduate students enrolled who were Asian or Pacific Islander, non-Hispanic			
0–1	780	75.1	1.48
2	380	79.2	1.42
3–5	470	75.4	1.51
6–73	440	80.6	1.49
Unknown	40	71.0	1.11
Percentage of undergraduate students enrolled who were Black, non-Hispanic			
0–3	540	78.4	1.41
4–8	580	79.1	1.39
9–17	480	77.0	1.63
18–100	490	76.7	1.52
Unknown	40	71.0	1.11
Percentage of undergraduate students enrolled who were Hispanic			
0–4	520	76.7	1.54
5–9	550	76.1	1.48
10–20	480	78.6	1.45
21–100	540	78.8	1.44
Unknown	40	71.0	1.11
Number of first-time, full-time undergraduate students living on campus who were receiving Title IV aid			
0–114	260	77.8	1.39
115–274	270	76.4	1.35
275–655	290	80.8	1.29
656–6,138	300	84.3	1.21
Unknown	1,010	73.9	1.66
Number of first-time, full-time undergraduate students with incomes up to \$30,000 who were receiving Title IV aid			
0–36	410	70.1	1.71
37–89	490	74.2	1.42
90–218	520	73.5	1.42
219–4,610	570	81.7	1.24
Unknown	130	66.9	2.14
Percentage of first-time, full-time undergraduate students receiving federal grant aid			
0–44	550	79.0	1.37
45–61	500	76.6	1.46
62–80	510	78.9	1.40
81–100	460	78.3	1.49
Unknown	110	63.1	2.27

See notes at end of table.

Table 60. Weight adjustment factors for institution nonresponse adjustment for undergraduate-enrolling survey institutions, by model predictor variable: 2019–20—Continued

Model predictor variable	Number of respondents ¹	Weighted response rate ²	Average weight adjustment factor (WTA_adj1)
Percentage of first-time, full-time undergraduate students receiving state or local grant aid			
0–8	450	73.5	1.72
9–28	500	80.1	1.34
29–51	510	79.6	1.33
52–100	550	78.4	1.36
Unknown	110	63.1	2.27
Percentage of first-time, full-time undergraduate students receiving institutional grant aid			
0–10	460	73.2	1.68
11–36	490	78.0	1.39
37–78	550	80.9	1.32
79–100	510	80.8	1.35
Unknown	110	63.1	2.27
Percentage of first-time, full-time undergraduate students receiving student loan aid			
0–14	520	73.3	1.43
15–46	540	82.6	1.29
47–68	520	81.5	1.35
69–100	440	69.3	1.69
Unknown	110	63.1	2.27
Average net price among first-time, full-time undergraduates receiving grant or scholarship aid			
\$–2,010–\$7,920	530	80.4	1.30
\$7,921–\$13,216	510	76.4	1.43
\$13,217–\$20,023	490	80.7	1.41
\$20,024–\$65,896	480	73.5	1.58
Unknown	120	66.2	2.19
CHAID segments			
Public 4-year, doctorate-granting institutions where the percentage of undergraduate students enrolled who were Asian or Pacific Islander, non-Hispanic is between 0 and 2	140	87.5	1.14
Public 4-year, doctorate-granting institutions where the percentage of undergraduate students enrolled who were Asian or Pacific Islander, non-Hispanic is between 3 and 5	70	76.6	1.34
Public 4-year, doctorate-granting institutions the percentage of undergraduate students enrolled who were Asian or Pacific Islander, non-Hispanic is between 6 and 73 or is unknown	120	90.4	1.11
Public 4-year, non-doctorate-granting, primarily baccalaureate or private nonprofit 4-year, doctorate-granting institutions where the percentage of undergraduate students enrolled who were Hispanic is between 0 and 9 and the percentage of undergraduate students enrolled who were Asian or Pacific Islander, non-Hispanic is between 0 and 2	100	80.8	1.27
Public 4-year, non-doctorate-granting, primarily baccalaureate or private nonprofit 4-year, doctorate-granting institutions where the percentage of undergraduate students enrolled who were Hispanic is between 0 and 9 and the percentage of undergraduate students enrolled who were Asian or Pacific Islander, non-Hispanic is between 3 and 73	110	67.6	1.53

See notes at end of table.

Table 60. Weight adjustment factors for institution nonresponse adjustment for undergraduate-enrolling survey institutions, by model predictor variable: 2019–20—Continued

Model predictor variable	Number of respondents ¹	Weighted response rate ²	Average weight adjustment factor (WTA_adj1)
Public 4-year, non-doctorate-granting, primarily baccalaureate or private nonprofit 4-year, doctorate-granting institutions where the percentage of undergraduate students enrolled who were Hispanic is between 10 and 100 or is unknown and the total female undergraduate enrollment is between 0 and 531	40	84.8	1.15
Public 4-year, non-doctorate-granting, primarily baccalaureate or private nonprofit 4-year, doctorate-granting institutions where the percentage of undergraduate students enrolled who were Hispanic is between 10 and 100 or is unknown and the total female undergraduate enrollment is between 532 and 90,299	210	84.0	1.20
Public 2-year or public 4-year, non-doctorate-granting, primarily subbaccalaureate institutions where the total female undergraduate enrollment is between 0 and 531 and the percentage of undergraduate students enrolled who were Black, non-Hispanic is between 0 and 3	30	54.8	1.69
Public 2-year or public 4-year, non-doctorate-granting, primarily subbaccalaureate institutions where the total female undergraduate enrollment is between 0 and 531 and the percentage of undergraduate students enrolled who were Black, non-Hispanic is between 4 and 100 or is unknown	50	69.4	1.35
Public 2-year or public 4-year, non-doctorate-granting, primarily subbaccalaureate institutions where the total female undergraduate enrollment is between 532 and 90,299 and the percentage of first-time, full-time undergraduate students receiving federal grant aid is between 0 and 61	350	73.5	1.41
Public 2-year or public 4-year non-doctorate-granting, primarily subbaccalaureate institutions where the total female undergraduate enrollment is between 532 and 90,299 and the percentage of first-time, full-time undergraduate students receiving federal grant aid is between 62 and 100 or is unknown	400	81.1	1.25
Public less-than-2-year; private nonprofit less-than-4-year; private nonprofit 4-year, non-doctorate-granting; or private for-profit 4-year institutions in the Southeast, Southwest, or Far West BEA region where the total male undergraduate enrollment is between 0 and 353	50	39.3	3.63
Public less-than-2-year; private nonprofit, less-than-4-year; private nonprofit 4-year, non-doctorate-granting; or private for-profit 4-year institutions in the Southeast, Southwest, or Far West BEA regions where the total male undergraduate enrollment is between 354 and 72,347	100	71.4	1.46
Public less-than-2-year; private nonprofit less-than-4-year; private nonprofit 4-year, non-doctorate-granting; or private for-profit 4-year institutions in the Great Lakes, Plains, or Outlying Areas BEA regions where the number of first-time, full-time undergraduate students with incomes up to \$30,000 who were receiving Title IV aid is between 0 and 36	50	53.6	1.68
Public less-than-2-year; private nonprofit less-than-4-year; private nonprofit 4-year, non-doctorate-granting; or private for-profit 4-year institutions in the Great Lakes, Plains, or Outlying Areas BEA regions where the number of first-time, full-time undergraduate students with incomes up to \$30,000 who were receiving Title IV aid is between 37 and 4,610 or is unknown	70	70.1	1.77
Public less-than-2-year; private nonprofit less-than-4-year; private nonprofit 4-year, non-doctorate-granting; or private for-profit 4-year institutions in the New England, Mideast, or Rocky Mountains BEA regions where the number of first-time, full-time undergraduate students with incomes up to \$30,000 who were receiving Title IV aid is between 0 and 36	50	92.5	1.10

See notes at end of table.

Table 60. Weight adjustment factors for institution nonresponse adjustment for undergraduate-enrolling survey institutions, by model predictor variable: 2019–20—Continued

Model predictor variable	Number of respondents ¹	Weighted response rate ²	Average weight adjustment factor (WTA_adj1)
Public less-than-2-year; private nonprofit less-than-4-year; private nonprofit 4-year, non-doctorate-granting; or private for-profit 4-year institutions in the New England, Mideast, or Rocky Mountains BEA regions where the number of first-time, full-time undergraduate students with incomes up to \$30,000 who were receiving Title IV aid is between 37 and 4,610 or is unknown	50	64.5	1.84
Private for-profit less-than-2-year institutions in AL, AR, AZ, CA, GA, IN, KS, MD, MI, MS, MO, NH, ND, NV, OH, OK, PR, RI, TX, UT, VA, WI, or WV	30	30.4	4.30
Private for-profit less-than-2-year institutions in CO, CT, DE, DC, FL, HI, ID, IL, KY, LA, ME, MN, MT, NC, NJ, NM, NY, OR, PA, SC, SD, TN, VT, WA, or WY	70	73.5	1.53
Private for-profit 2-year institutions in AK, AZ, CA, CO, GA, IA, KS, ME, MD, MI, MN, MT, NM, NV, NY, ND, OR, PR, PA, SD, TN, TX, UT, WA, WV, WI, or WY	30	30.7	4.79
Private for-profit 2-year institutions in CT, DC, FL, HI, ID, IN, LA, MS, MO, NE, NJ, NC, OH, OK, SC, or VA	50	76.5	1.22

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 2019–20 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 defined by quartiles for continuous variables. Fewer than four categories will display in cases where the variable's values cause quartiles to overlap. Categories of "None" or "Unknown" were excluded from quartile calculations. 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

Table 61. Weight adjustment factors for institution nonresponse adjustment for graduate-enrolling institutions, by model predictor variable: 2019–20

Model predictor variable	Number of respondents ¹	Weighted response rate ²	Average weight adjustment factor (WTA_adj1)
Total	1,090	80.6	1.37
Control and level of institution			
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	60	44.3	1.40
Public 4-year, non-doctorate-granting, primarily baccalaureate	150	83.9	1.19
Public 4-year, doctorate-granting	340	88.7	1.15
Private nonprofit 4-year, non-doctorate-granting	180	66.1	1.73
Private nonprofit 4-year, doctorate-granting	300	74.2	1.40
Private for-profit 4-year	60	77.8	1.79
Carnegie classification code			
Associate's	50	97.9	1.03
Research and doctoral	250	83.6	1.22
Master's	450	79.7	1.36
Baccalaureate	190	78.7	1.42
Special focus and other	130	69.0	1.68
Unavailable or unknown	10	68.6	2.08

See notes at end of table.

Table 61. Weight adjustment factors for institution nonresponse adjustment for graduate-enrolling institutions, by model predictor variable: 2019–20—Continued

Model predictor variable	Number of respondents ¹	Weighted response rate ²	Average weight adjustment factor (WTA_adj1)
Degree of urbanization			
Large city	280	83.1	1.25
Midsize city	150	77.6	1.36
Small city	150	71.2	1.62
Large suburb	180	79.2	1.35
Midsize suburb	40	96.9	1.01
Small suburb	30	99.4	1.01
Urban area on fringe of town	40	86.0	1.03
Urban area distant from town	100	80.4	1.60
Urban area remote from town	80	95.4	1.02
Rural area on fringe of town	30	54.9	2.47
Rural area distant from town	10	48.6	1.44
Rural area remote from town	10	72.2	2.51
Bureau of Economic Analysis (BEA) region ³			
New England	80	66.4	1.87
Mideast	170	80.6	1.24
Great Lakes	150	81.6	1.26
Plains	120	75.5	1.37
Southeast	280	78.8	1.28
Southwest	100	90.2	1.24
Rocky Mountains	60	89.4	1.20
Far West	120	82.5	1.81
Outlying Areas	30	92.1	1.12
Historically Black College or University			
Yes	50	84.0	1.16
No, unavailable, or unknown	1,040	80.6	1.38
Hispanic-Serving Institution			
Yes	160	87.4	1.20
No	930	79.6	1.40
Total graduate enrollment			
1–218	250	72.4	1.48
219–884	260	75.2	1.50
885–2,812	270	71.8	1.31
2,813–64,922	300	83.4	1.21
Total male graduate enrollment			
0–70	250	74.4	1.47
71–301	270	68.1	1.52
302–1,014	270	76.9	1.29
1,015–22,822	290	83.2	1.21
Total female graduate enrollment			
0–130	250	75.0	1.50
131–565	260	73.6	1.41
566–1,709	280	73.8	1.38
1,710–52,745	290	83.1	1.21

See notes at end of table.

Table 61. Weight adjustment factors for institution nonresponse adjustment for graduate-enrolling institutions, by model predictor variable: 2019–20—Continued

Model predictor variable	Number of respondents ¹	Weighted response rate ²	Average weight adjustment factor (WTA_adj1)
Percentage of graduate students enrolled who were Asian or Pacific Islander, non-Hispanic			
0–1	230	84.3	1.34
2–3	280	68.9	1.54
4–6	250	84.3	1.43
7–68	230	84.1	1.27
Unknown	100	99.5	1.00
Percentage of graduate students enrolled who were Black, non-Hispanic			
0–3	280	86.7	1.44
4–7	260	77.4	1.40
8–14	220	85.9	1.21
15–100	230	74.4	1.55
Unknown	100	99.5	1.00
Percentage of graduate students enrolled who were Hispanic			
0–3	290	79.3	1.35
4–5	200	80.2	1.33
6–10	260	76.1	1.50
11–100	240	86.2	1.44
Unknown	100	99.5	1.00
Total institution employee count of office and administrative support staff			
1–31	230	69.1	1.34
32–75	260	70.6	1.45
76–194	270	85.5	1.21
195–3,743	300	83.7	1.21
Unknown	30	51.2	4.00
CHAID segments			
Public 4-year institutions where the percentage of graduate students enrolled who were Hispanic is between 0 and 5	240	83.9	1.31
Public 4-year institutions where the percentage of graduate students enrolled who were Hispanic is between 6 and 100 or is unknown	300	91.7	1.09
Private nonprofit 4-year, doctorate-granting institutions where the percentage of graduate students enrolled who were Hispanic is between 0 and 10	230	70.8	1.46
Private for-profit 4-year institutions where the percentage of graduate students enrolled who were Hispanic is between 0 and 10	20	60.9	2.96
Private nonprofit 4-year, doctorate-granting or private for-profit 4-year institutions where the percentage of graduate students enrolled who were Hispanic is between 11 and 100 or is unknown and the total male graduate enrollment with is between 0 and 301	50	93.5	1.06
Private nonprofit 4-year, doctorate-granting or private for-profit 4-year institutions where the percentage of graduate students enrolled who were Hispanic is between 11 and 100 or is unknown and the total male graduate enrollment is between 302 and 22,822	50	86.1	1.19
Private nonprofit 4-year, non-doctorate-granting institutions in the Mideast, Plains, or Southeast BEA regions	110	69.5	1.29
Private nonprofit 4-year, non-doctorate-granting institutions in the New England, Great Lakes, Southwest, Rocky Mountains, Far West, or Outlying Areas BEA regions where the percentage of graduate students enrolled who were Asian or Pacific Islander, non-Hispanic is between 0 and 3	40	67.0	1.77

See notes at end of table.

Table 61. Weight adjustment factors for institution nonresponse adjustment for graduate-enrolling institutions, by model predictor variable: 2019–20—Continued

Model predictor variable	Number of respondents ¹	Weighted response rate ²	Average weight adjustment factor (WTA_adj1)
Private nonprofit 4-year, non-doctorate-granting institutions in the New England, Great Lakes, Southwest, Rocky Mountains, Far West, or Outlying Areas BEA regions where the percentage of graduate students enrolled who were Asian or Pacific Islander, non-Hispanic is between 4 and 68 or is unknown	30	63.3	3.07

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 2019–20 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 defined by quartiles for continuous variables. Fewer than four categories will display in cases where the variable's values cause quartiles to overlap. Categories of "None" or "Unknown" were excluded from quartile calculations. Number of respondents 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

Table 62 shows the average study weight adjustment by institution stratum within state for the institution nonresponse study weight adjustments in the undergraduate-enrolling study institution model. Summary statistics of the weight adjustment factor for undergraduate-enrolling study institutions follow:

- minimum: 1.00;
- median: 1.13; and
- maximum: 50.75.

Table 62. Weight adjustment factors for institution nonresponse adjustment for undergraduate-enrolling study institutions, by state and institution stratum: 2019–20

State and institution stratum	Number of respondents ¹	Weighted response rate ²	Average weight adjustment factor (WTB_adj1)
Total	2,150	78.7	1.51
Alabama			
Public 2-year	20	85.3	1.20
Public 4-year	10	53.3	1.95
Other	20	76.0	1.69
Alaska			
Public 4-year	#	98.3	1.01
Other	#	74.2	1.15
Arizona			
Public 2-year	20	81.6	1.26
Public 4-year	10	100.0	1.00
Other	20	78.8	1.18

See notes at end of table.

Table 62. Weight adjustment factors for institution nonresponse adjustment for undergraduate-enrolling study institutions, by state and institution stratum: 2019–20—Continued

State and institution stratum	Number of respondents ¹	Weighted response rate ²	Average weight adjustment factor (WTB_adj1)
Arkansas			
Public 2-year	20	85.0	1.22
Public 4-year	10	92.3	1.07
Other	10	42.3	3.98
California			
Public 2-year	70	69.1	1.61
Public 4-year	30	74.9	1.34
Other	20	67.3	1.57
Colorado			
Public 2-year	10	100.0	1.00
Public 4-year	20	98.1	1.02
Other	10	28.3	3.43
Connecticut			
Public 2-year	10	100.0	1.00
Public 4-year	10	41.1	2.32
Other	20	59.2	1.33
Delaware			
Public 4-year	#	100.0	1.00
Other	10	81.3	1.49
District of Columbia			
Public 4-year	#	100.0	1.00
Other	10	64.9	1.55
Florida			
Public 2-year	20	76.5	1.37
Public 4-year	40	91.2	1.12
Other	20	67.3	1.38
Georgia			
Public 2-year	20	68.0	1.56
Public 4-year	30	100.0	1.00
Other	20	72.0	1.31
Hawaii			
Public 2-year	10	100.0	1.00
Public 4-year	#	100.0	1.00
Other	10	64.3	6.17
Idaho			
Public 2-year	#	70.7	1.41
Public 4-year	#	60.8	1.79
Other	20	96.0	1.10
Illinois			
Public 2-year	40	79.8	1.28
Public 4-year	10	97.8	1.02
Other	20	80.0	1.13
Indiana			
Public 2-year	#	100.0	1.00
Public 4-year	10	67.0	1.63
Other	20	70.8	1.30
Iowa			
Public 2-year	10	81.4	1.47
Public 4-year	#	100.0	1.00
Other	20	51.6	2.09

See notes at end of table.

Table 62. Weight adjustment factors for institution nonresponse adjustment for undergraduate-enrolling study institutions, by state and institution stratum: 2019–20—Continued

State and institution stratum	Number of respondents ¹	Weighted response rate ²	Average weight adjustment factor (WTB_adj1)
Kansas			
Public 2-year	10	33.1	5.13
Public 4-year	10	92.3	1.06
Other	20	56.0	2.08
Kentucky			
Public 2-year	10	56.2	1.77
Public 4-year	10	87.1	1.11
Other	20	87.8	1.11
Louisiana			
Public 2-year	10	74.8	1.45
Public 4-year	10	89.8	1.12
Other	20	57.0	2.18
Maine			
Public 2-year	#	70.0	1.25
Public 4-year	10	100.0	1.00
Other	20	81.8	1.40
Maryland			
Public 2-year	10	69.9	1.59
Public 4-year	10	100.0	1.00
Other	10	61.5	1.94
Massachusetts			
Public 2-year	20	90.5	1.13
Public 4-year	10	100.0	1.00
Other	20	71.4	1.43
Michigan			
Public 2-year	20	63.9	2.37
Public 4-year	20	90.9	1.08
Other	20	57.8	2.36
Minnesota			
Public 2-year	30	99.7	1.00
Public 4-year	10	100.0	1.00
Other	20	53.5	2.11
Mississippi			
Public 2-year	10	74.6	1.28
Public 4-year	10	72.1	1.27
Other	10	35.4	8.46
Missouri			
Public 2-year	10	52.3	1.65
Public 4-year	10	72.2	1.33
Other	20	76.9	1.25
Montana			
Public 2-year	10	50.3	1.64
Public 4-year	10	95.0	1.04
Other	10	57.0	2.09
Nebraska			
Public 2-year	10	76.7	1.32
Public 4-year	#	75.5	1.65
Other	20	94.3	1.02

See notes at end of table.

Table 62. Weight adjustment factors for institution nonresponse adjustment for undergraduate-enrolling study institutions, by state and institution stratum: 2019–20—Continued

State and institution stratum	Number of respondents ¹	Weighted response rate ²	Average weight adjustment factor (WTB_adj1)
Nevada			
Public 4-year	10	100.0	1.00
Other	10	43.4	2.37
New Hampshire			
Public 2-year	#	65.1	1.50
Public 4-year	10	100.0	1.00
Other	10	96.4	1.03
New Jersey			
Public 2-year	20	82.5	1.20
Public 4-year	10	100.0	1.00
Other	20	76.6	1.28
New Mexico			
Public 2-year	20	94.6	1.08
Public 4-year	10	90.2	1.12
Other	10	56.6	1.66
New York			
Public 2-year	30	88.2	1.17
Public 4-year	40	94.0	1.06
Other	20	73.2	1.37
North Carolina			
Public 2-year	50	85.7	1.19
Public 4-year	20	86.3	1.14
Other	30	79.8	1.23
North Dakota			
Public 2-year	#	37.5	2.67
Public 4-year	#	68.8	3.62
Other	10	76.3	1.45
Ohio			
Public 2-year	20	97.5	1.03
Public 4-year	40	85.3	1.21
Other	20	63.9	1.73
Oklahoma			
Public 2-year	20	89.9	1.23
Public 4-year	10	81.8	1.22
Other	20	65.5	1.68
Oregon			
Public 2-year	10	62.3	1.75
Public 4-year	10	95.2	1.05
Other	20	55.9	2.21
Pennsylvania			
Public 2-year	10	95.0	1.04
Public 4-year	40	87.0	1.06
Other	20	78.6	1.27
Puerto Rico			
Public 2-year	10	100.0	1.00
Public 4-year	10	85.0	1.56
Other	20	74.9	1.33
Rhode Island			
Public 2-year	#	100.0	1.00
Public 4-year	#	0.0	0.0
Other	10	62.7	6.44

See notes at end of table.

Table 62. Weight adjustment factors for institution nonresponse adjustment for undergraduate-enrolling study institutions, by state and institution stratum: 2019–20—Continued

State and institution stratum	Number of respondents ¹	Weighted response rate ²	Average weight adjustment factor (WTB_adj1)
South Carolina			
Public 2-year	10	36.0	3.43
Public 4-year	10	74.5	1.32
Other	20	86.8	1.15
South Dakota			
Public 2-year	#	64.1	1.63
Public 4-year	10	87.8	1.18
Other	10	63.7	4.21
Tennessee			
Public 2-year	30	89.5	1.17
Public 4-year	10	82.3	1.25
Other	20	61.8	1.52
Texas			
Public 2-year	40	74.7	1.66
Public 4-year	40	82.0	1.29
Other	20	62.2	1.71
Utah			
Public 2-year	#	0.0	0.0
Public 4-year	10	79.3	1.25
Other	20	96.1	1.06
Vermont			
Public 2-year	#	100.0	1.00
Public 4-year	#	76.8	1.25
Other	10	81.5	1.29
Virginia			
Public 2-year	20	94.3	1.13
Public 4-year	10	97.6	1.08
Other	20	26.3	4.10
Washington			
Public 2-year	10	80.5	1.20
Public 4-year	20	74.1	1.37
Other	20	88.6	1.22
West Virginia			
Public 2-year	10	64.5	2.76
Public 4-year	10	96.0	1.03
Other	10	90.5	1.78
Wisconsin			
Public 2-year	10	57.0	1.60
Public 4-year	10	66.9	1.47
Other	10	35.7	3.27
Wyoming			
Public 2-year	10	79.0	1.23
Public 4-year	#	100.0	1.00
Other	#	17.4	5.76

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 2019–20 Integrated Postsecondary Education Data System files. Response rates are expressed as percentages.

NOTE: Number of respondents 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

7.1.2.3 *Institution poststratification adjustments (WTA_adj2 and WTB_adj2)*

After adjusting for the inverse of the probability of selection into the sample and nonresponse, each nonresponse-adjusted institution weight was further adjusted to meet enrollment totals (control totals) by subgroups.

The enrollment totals were obtained from the IPEDS 2019–20 Institutional Characteristics Header, Institutional Characteristics, and 12-Month Enrollment files. Survey enrollment totals were calculated by institution size⁴⁹ (*small* vs. *large*) within control and level of institution for undergraduate- and graduate-enrolling institutions. Public less-than-2-year, undergraduate-enrolling institutions and public 4-year, non-doctorate-granting, primarily subbaccalaureate graduate-enrolling institutions were not split by institution size due to the small number of institutions in the sample. Undergraduate study enrollment totals were determined for public 2-year, public 4-year, and other institutions within each state.

This adjustment using enrollment ensured that the resultant weight adequately represents the student target population. The weights used in the poststratification models were the institution sampling weights adjusted for nonresponse multiplied by the institution’s full-year undergraduate or graduate enrollment totals.

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, institutions in the “all other sectors” sampling stratum were selected with probability proportional to size, with the size being counts of students. This method of sampling and weighting does not yield an accurate estimate of institutions overall or within control and level of institution.

For the undergraduate-enrolling survey institution and graduate-enrolling institution poststratification weight factors (WTA_adj2), an upper bound of 10 was set on the weight adjustment factors coming out of the poststratification weight adjustment. For the undergraduate-enrolling study institution poststratification weight factor (WTB_adj2), an upper bound of 6 was set on the weight adjustment factors coming out of the poststratification weight adjustment. No lower bounds were applied, and weights were not trimmed for any of the institution poststratification weight adjustment factors.

Table 63 and table 64 show the characteristics associated with the control totals and the average weight adjustments factors by these characteristics for the

⁴⁹ Institution size was determined based on the median total enrollment as a cut point within each control and level of institution.

undergraduate-enrolling survey institutions and graduate-enrolling institutions, respectively. Summary statistics of the poststratification survey weight adjustment factors follow:

- minimum: 0.30 for undergraduate-enrolling survey institutions and 0.91 for graduate-enrolling institutions;
- median: 1.01 for undergraduate-enrolling survey institutions and 0.99 for graduate-enrolling institutions; and
- maximum: 1.38 for undergraduate-enrolling survey institutions and 1.23 for graduate-enrolling institutions.

Table 63. Weight adjustment factors for institution poststratification adjustment for undergraduate-enrolling survey institutions, by poststratification category: 2019–20

Control and level of institution by size ¹	Control total ²	Average weight adjustment factor (WTA_adj2)
Total	22,424,817	1.00
Public less-than-2-year, all	80,813	1.09
Public 2-year, small	1,078,922	0.95
Public 2-year, large	6,930,856	1.00
Public 4-year, non-doctorate-granting, primarily subbaccalaureate, small	262,017	0.97
Public 4-year, non-doctorate-granting, primarily subbaccalaureate, large	1,703,819	0.99
Public 4-year, non-doctorate-granting, primarily baccalaureate, small	224,377	1.07
Public 4-year, non-doctorate-granting, primarily baccalaureate, large	917,033	0.99
Public 4-year, doctorate-granting, small	1,281,048	0.96
Public 4-year, doctorate-granting, large	4,754,356	1.00
Private nonprofit 2-year or less, small	8,005	0.30
Private nonprofit 2-year or less, large	72,378	0.70
Private nonprofit 4-year, non-doctorate-granting, small	152,943	1.34
Private nonprofit 4-year, non-doctorate-granting, large	1,112,203	0.95
Private nonprofit 4-year, doctorate-granting, small	294,592	0.95
Private nonprofit 4-year, doctorate-granting, large	1,862,020	1.02
Private for-profit less-than-2-year, small	57,869	1.33
Private for-profit less-than-2-year, large	295,314	1.12
Private for-profit 2-year, small	44,087	1.26
Private for-profit 2-year, large	273,163	0.87
Private for-profit 4-year, small	39,331	0.75
Private for-profit 4-year, large	979,671	1.04

¹ Size for poststratification weighting classes was based on the median enrollment within control and level for the institutions on the sampling frame.

² Control totals are the sum of enrollment across institutions based on Integrated Postsecondary Education Data System 2019–20 enrollment data.

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

Table 64. Weight adjustment factors for institution poststratification adjustment for graduate-enrolling institutions, by control and level of institution by size: 2019–20

Control and level of institution by size ¹	Control total ²	Average weight adjustment factor (WTA_adj2)
Total	3,955,983	0.99
Public 4-year, non-doctorate-granting, primarily subbaccalaureate, all	1,337	0.92
Public 4-year, non-doctorate-granting, primarily baccalaureate, small	18,948	1.21
Public 4-year, non-doctorate-granting, primarily baccalaureate, large	124,414	0.97
Public 4-year, doctorate-granting, small	297,430	1.00
Public 4-year, doctorate-granting, large	1,402,692	1.00
Private nonprofit 4-year, non-doctorate-granting, small	14,087	0.85
Private nonprofit 4-year, non-doctorate-granting, large	203,687	0.93
Private nonprofit 4-year, doctorate-granting, small	138,245	0.99
Private nonprofit 4-year, doctorate-granting, large	1,372,500	0.99
Private for-profit 4-year, small	5,864	1.06
Private for-profit 4-year, large	376,779	0.91

¹ Size for poststratification weighting classes was based on the median enrollment within control and level for the institutions on the sampling frame.

² Control totals are the sum of enrollment across institutions based on Integrated Postsecondary Education Data System 2019–20 enrollment data.

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

Table 65 shows the characteristics associated with the control totals and the average weight adjustments factors by these characteristics for the undergraduate-enrolling study institutions. Summary statistics of the institution poststratification study weight adjustment factor follow:

- minimum: 0.90;
- median: 1.00; and
- maximum: 2.30.

Table 65. Weight adjustment factors for institution poststratification adjustment for undergraduate-enrolling study institutions, by state and institution stratum: 2019–20

State and institution stratum	Control total ¹	Average weight adjustment factor (WTB_adj2)
Total	22,424,817	1.00
Alabama		
Public 2-year	111,541	1.00
Public 4-year	152,189	1.00
Other	56,669	1.00
Alaska		
Public 4-year	33,817	1.00
Other	2,299	1.00
Arizona		
Public 2-year	288,711	1.00
Public 4-year	192,329	1.00
Other	277,050	0.98

See notes at end of table.

Table 65. Weight adjustment factors for institution poststratification adjustment for undergraduate-enrolling study institutions, by state and institution stratum: 2019–20—
Continued

State and institution stratum	Control total ¹	Average weight adjustment factor (WTB_adj2)
Arkansas		
Public 2-year	63,304	1.00
Public 4-year	88,649	1.01
Other	19,866	1.01
California		
Public 2-year	1,815,110	0.96
Public 4-year	1,074,740	1.00
Other	450,994	1.05
Colorado		
Public 2-year	30,455	1.00
Public 4-year	283,623	1.00
Other	113,068	1.02
Connecticut		
Public 2-year	63,886	1.00
Public 4-year	58,774	1.00
Other	91,981	1.03
Delaware		
Public 4-year	44,743	1.00
Other	19,054	1.03
District of Columbia		
Public 4-year	4,872	1.00
Other	59,542	1.00
Florida		
Public 2-year	78,049	0.97
Public 4-year	922,966	1.00
Other	336,211	1.00
Georgia		
Public 2-year	173,663	1.17
Public 4-year	314,115	1.00
Other	113,306	0.98
Hawaii		
Public 2-year	33,903	1.00
Public 4-year	26,822	1.00
Other	11,796	0.97
Idaho		
Public 2-year	30,691	1.00
Public 4-year	68,132	1.00
Other	71,060	1.00
Illinois		
Public 2-year	472,479	1.00
Public 4-year	144,989	1.00
Other	217,517	1.00
Indiana		
Public 2-year	162,646	1.00
Public 4-year	249,702	1.00
Other	89,430	1.00
Iowa		
Public 2-year	127,596	1.00
Public 4-year	64,873	1.00
Other	49,413	0.99

See notes at end of table.

Table 65. Weight adjustment factors for institution poststratification adjustment for undergraduate-enrolling study institutions, by state and institution stratum: 2019–20—
Continued

State and institution stratum	Control total ¹	Average weight adjustment factor (WTB_adj2)
Kansas		
Public 2-year	120,402	1.00
Public 4-year	86,292	1.00
Other	38,154	0.99
Kentucky		
Public 2-year	107,892	1.00
Public 4-year	110,687	1.00
Other	50,393	1.01
Louisiana		
Public 2-year	91,046	1.00
Public 4-year	134,628	1.00
Other	44,346	1.03
Maine		
Public 2-year	23,416	1.00
Public 4-year	32,714	1.00
Other	25,865	1.00
Maryland		
Public 2-year	159,058	1.00
Public 4-year	176,058	1.00
Other	43,046	1.01
Massachusetts		
Public 2-year	112,432	1.00
Public 4-year	112,879	1.00
Other	214,329	0.94
Michigan		
Public 2-year	200,213	1.00
Public 4-year	291,114	1.00
Other	83,366	1.13
Minnesota		
Public 2-year	159,854	1.00
Public 4-year	126,182	1.00
Other	100,318	1.01
Mississippi		
Public 2-year	98,227	1.00
Public 4-year	71,134	1.00
Other	20,240	1.01
Missouri		
Public 2-year	113,963	1.00
Public 4-year	132,549	1.00
Other	115,299	1.00
Montana		
Public 2-year	12,169	1.00
Public 4-year	37,326	1.00
Other	4,302	1.00
Nebraska		
Public 2-year	57,771	1.00
Public 4-year	55,142	1.00
Other	34,863	0.98

See notes at end of table.

Table 65. Weight adjustment factors for institution poststratification adjustment for undergraduate-enrolling study institutions, by state and institution stratum: 2019–20—
Continued

State and institution stratum	Control total ¹	Average weight adjustment factor (WTB_adj2)
Nevada		
Public 2-year	†	†
Public 4-year	128,876	1.00
Other	14,151	0.93
New Hampshire		
Public 2-year	23,543	1.00
Public 4-year	24,667	1.00
Other	153,022	1.00
New Jersey		
Public 2-year	200,913	1.00
Public 4-year	175,057	1.00
Other	95,937	0.92
New Mexico		
Public 2-year	93,086	1.00
Public 4-year	48,298	1.00
Other	5,626	1.02
New York		
Public 2-year	383,652	1.00
Public 4-year	397,942	1.00
Other	492,144	1.05
North Carolina		
Public 2-year	314,373	1.00
Public 4-year	214,144	1.00
Other	100,134	0.99
North Dakota		
Public 2-year	9,795	1.00
Public 4-year	38,520	1.00
Other	6,926	0.99
Ohio		
Public 2-year	216,307	1.00
Public 4-year	383,257	1.00
Other	159,517	0.99
Oklahoma		
Public 2-year	86,287	1.00
Public 4-year	106,917	1.00
Other	41,841	0.97
Oregon		
Public 2-year	144,378	1.00
Public 4-year	101,060	1.00
Other	30,613	1.00
Pennsylvania		
Public 2-year	172,034	1.00
Public 4-year	243,284	1.00
Other	280,911	1.06
Puerto Rico		
Public 2-year	1,971	1.00
Public 4-year	50,904	1.00
Other	143,607	0.90

See notes at end of table.

Table 65. Weight adjustment factors for institution poststratification adjustment for undergraduate-enrolling study institutions, by state and institution stratum: 2019–20—
Continued

State and institution stratum	Control total [†]	Average weight adjustment factor (WTB_adj2)
Rhode Island		
Public 2-year and Public 4-year	44,266	2.30
Other	39,367	1.01
South Carolina		
Public 2-year	92,997	1.00
Public 4-year	118,559	1.00
Other	49,167	1.02
South Dakota		
Public 2-year	9,024	1.00
Public 4-year	39,832	1.00
Other	9,487	1.00
Tennessee		
Public 2-year	135,058	1.00
Public 4-year	126,550	1.00
Other	97,207	1.03
Texas		
Public 2-year	893,113	1.00
Public 4-year	835,128	1.00
Other	213,090	1.06
Utah		
Public 2-year and Other	265,771	1.20
Public 4-year	172,553	1.00
Vermont		
Public 2-year	9,251	1.00
Public 4-year	20,406	1.00
Other	16,614	1.00
Virginia		
Public 2-year	221,071	1.00
Public 4-year	198,878	1.00
Other	156,512	0.96
Washington		
Public 2-year	45,792	1.00
Public 4-year	348,964	1.00
Other	46,160	0.99
West Virginia		
Public 2-year	21,951	1.00
Public 4-year	61,316	1.00
Other	90,719	1.00
Wisconsin		
Public 2-year	135,599	1.00
Public 4-year	179,679	1.00
Other	53,812	0.94
Wyoming		
Public 2-year	27,691	1.00
Public 4-year	10,836	1.00
Other	409	1.00

† Not applicable.

[†] Control totals are the sum of enrollment across institutions based on Integrated Postsecondary Education Data System 2019–20 enrollment data.

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

7.1.2.4 *Final institution weights (WTA_INSTWT and WTB_INSTWT)*

The final institution weights were calculated by multiplying the institution sampling weights by the nonresponse and poststratification adjustments:

$WTA_INSTWT = WTA_WT1 * WTA_adj1 * WTA_adj2$ for both undergraduate-enrolling survey responding institutions and graduate-enrolling responding institutions.

$WTB_INSTWT = WTB_WT1 * WTB_adj1 * WTB_adj2$ for undergraduate-enrolling study responding institutions.

The final institution weights were used to develop student analysis weights, described in section 7.1.3.

7.1.3 *Student Weights*

A student survey analysis weight was computed for each student survey respondent, and a student study analysis weight was computed for each study respondent as defined in section 7.1.1.2. The final analysis weights were calculated through a series of steps: (1) create the student within-institution sampling weight; (2) multiply by the institution sampling weight to obtain the student sampling weight; (3) adjust for multiplicity; (4) adjust for unknown eligibility; (5) adjust for nonresponse; (6) poststratify; and (7) compute final analysis weight. Each step is described below.

7.1.3.1 *Student within-institution sampling weight (WTA_WT4 and WTB_WT4)*

As described in chapter 2, the overall student sampling strata were defined for undergraduate and graduate students by crossing the institution sampling strata and the student strata. 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 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 for the s th student stratum and $\pi_r(i)$ = the institution's probability of selection. This sampling rate accounts for the subsampling of students selected to receive the survey. That is, there is one sampling rate for all students used for computing the study weight and another

sampling rate for the students who were selected to receive the survey used for computing the survey weight.

Initial student sampling rates were set using IPEDS data before enrollment lists were received. If an institution's enrollment list contained more students than expected based on the IPEDS data, the initial student sampling rates across strata would yield a larger-than-expected sample size for that institution. Likewise, if the enrollment list count was smaller than expected based on the IPEDS data, the initial student sampling rates across strata would yield a smaller-than-expected sample size for that institution. To maintain control over the sample sizes and meet stratum yield targets, the sampling rates were adjusted, when necessary, so that the number of students selected within an institution usually did not exceed 600, unless exceeding 600 was necessary to have a sufficient sample size for state representation. A minimum sample size constraint of 30 students was imposed to ensure sufficient yield for variance estimation. The sample for some institutions was less than 30 because there were fewer than 30 students on the enrollment list.

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

$$\text{WTA_WT4} = 1/f_{s|i}$$

where $f_{s|i}$ are the sampling rates for students who were selected to receive the survey.

Similarly, the student within-institution sampling survey weight was calculated as the reciprocal of the adjusted institution-specific student stratum study sampling rates, or

$$\text{WTB_WT4} = 1/f_{s|i}$$

where $f_{s|i}$ are the sampling rates for students selected for the study.

7.1.3.2 Student sampling weights (WTA_WT5 and WTB_WT5)

The student sampling weights, denoted WTA_WT5 and WTB_WT5, were defined as the product of the institution weights (WTA_INSTWT and WTB_INSTWT) and student within-institution sampling weights (WTA_WT4 and WTB_WT4):

- $\text{WTA_WT5} = \text{WTA_INSTWT} * \text{WTA_WT4}$ for undergraduate survey students and graduate students; and
- $\text{WTB_WT5} = \text{WTB_INSTWT} * \text{WTB_WT4}$ for undergraduate study students.

7.1.3.3 *Student multiplicity adjustment (WTA_adj3 and WTB_adj3)*

Students who attended more than one eligible institution during the 2019–20 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 sampling weights (WTA_WT5 and WTB_WT5) by the number of institutions attended that were eligible for sample selection. Specifically, the student multiplicity weight adjustment factors were defined as

$$WTA_adj3 = WTB_adj3 = 1 / M,$$

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 were set to 1 if there was no evidence to the contrary.

Summary statistics of the student multiplicity weight adjustment factors follow:

- minimum: 0.09 for undergraduate students and 0.14 for graduate students;
- median: 1.00 for both undergraduate and graduate students; and
- maximum: 1.00 for both undergraduate and graduate students.

7.1.3.4 *Student unknown eligibility status adjustment (WTA_adj4 and WTB_adj4)*

It was not possible to determine the final eligibility status of some nonresponding students. These students were treated as eligible, and their weights were adjusted to compensate for the small portion of these students who were likely ineligible. Unknown eligibility weight adjustment factors were calculated using the rate of eligibility among students with known eligibility status within weighting classes defined by the intersection of institution control and level with the students' matching status to financial aid files (matched to NSLDS, matched to CPS only, and no match). For the students with known eligibility status, the weight adjustment factor is equal to 1.

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

- minimum: 0.93 for undergraduate students and 0.96 for graduate students;
- median: 1.00 for both undergraduate and graduate students; and
- maximum: 1.00 for both undergraduate and graduate students.

Table 66 shows the weight adjustment factors applied to the students with unknown eligibility for undergraduate and graduate students.

Table 66. Weight adjustment factors for unknown student eligibility status adjustment for undergraduate and graduate students, by control and level of institution: 2019–20

Control and level of institution	Undergraduate survey students		Graduate students		Undergraduate study students	
	Number adjusted for unknown eligibility	Weight adjustment factor (WTA_adj4)	Number adjusted for unknown eligibility	Weight adjustment factor (WTA_adj4)	Number adjusted for unknown eligibility	Weight adjustment factor (WTA_adj4)
Public less-than-2-year	50	0.96	†	†	60	0.98
Public 2-year	9,220	0.94	†	†	3,920	0.96
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	1,000	0.95	#	0.98	360	0.98
Public 4-year, non-doctorate-granting, primarily baccalaureate	1,050	0.98	100	0.96	400	1.00
Public 4-year, doctorate-granting	1,600	0.98	370	0.98	1,090	1.00
Private nonprofit 2-year or less	250	0.97	†	†	150	0.98
Private nonprofit 4-year, non-doctorate-granting	800	0.98	110	0.99	500	1.00
Private nonprofit 4-year, doctorate-granting	1,050	0.98	490	0.97	410	1.00
Private for-profit less-than-2-year	1,540	0.93	†	†	580	0.95
Private for-profit 2-year	2,440	0.95	†	†	1,600	0.93
Private for-profit 4-year	1,250	0.97	300	0.99	390	1.00

† Not applicable.

Rounds to zero.

NOTE: Number adjusted for unknown eligibility rounded to the nearest 10. Detail may not sum to totals because of rounding.

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

7.1.3.5 *Student base weights (WTA_WT7 and WTB_WT7)*

Student base weights were calculated as the product of the student sampling weights (WTA_WT5 and WTB_WT5), student multiplicity adjustment factors (WTA_adj3 and WTB_adj3), and unknown eligibility status adjustment factors (WTA_adj4 and WTB_adj4):

$$\text{WTA_WT7} = \text{WTA_WT5} * \text{WTA_adj3} * \text{WTA_adj4} \text{ and}$$

$$\text{WTB_WT7} = \text{WTB_WT5} * \text{WTB_adj3} * \text{WTB_adj4}.$$

7.1.3.6 *Student nonresponse adjustment (WTA_adj5 and WTB_adj5)*

The student survey weight for undergraduate students was adjusted for nonresponse in three stages. These stages include inability to locate the student, survey refusal, and other nonresponse (student located, did not refuse, did not respond) because the predictors of response propensity were different for each of these nonresponse outcomes. Adjusting for different types of nonresponse separately helps reduce nonresponse bias when the significant predictors of response propensity are different for each type of nonresponse. For graduate students, characteristics were compared between respondents and nonrespondents to determine that only two stages of nonresponse adjustment were necessary. These two stages are inability to locate the student and other nonresponse (student located, did not respond) because the predictors of response propensity were different for each of these nonresponse outcomes. The student study weight was adjusted for nonresponse in one step.

The student survey nonresponse adjustment was calculated as the product of the student survey not located adjustment (WTA_adj5a), the student survey refusal nonresponse adjustment (WTA_adj5b), and the student survey other nonresponse adjustment (WTA_adj5c), as follows:

$$\text{WTA_adj5} = \text{WTA_adj5a} * \text{WTA_adj5b} * \text{WTA_adj5c},$$

where $\text{WTA_adj5b} = 1$ for graduate students.

The models for each nonresponse adjustment were developed in the following three stages similar to the nonresponse weight adjustments for institutions described above:

1. identification of model predictors;
2. CHAID analysis to determine significant interactions between model predictors (see section 7.1.2.2 for information about CHAID); and
3. WTADJUST procedure to calculate nonresponse adjustment.

Candidate predictor variables for the student nonresponse adjustment models were selected because they were thought to predict response status and be correlated with one or more outcomes, based on knowledge of past NPSAS data collections, and were nonmissing for most respondents and nonrespondents.

Table 67 lists the predictor variables that were used in the student nonresponse adjustment models and indicates whether the predictor variables were included in the nonresponse adjustment model for the undergraduate survey, graduate, and/or undergraduate study weight.

Table 67. Candidate predictor variables for student nonresponse adjustment models: 2019–20

Nonresponse adjustment model predictor variable ¹	Undergraduate survey	Graduate	Undergraduate study
Institution characteristics			
Undergraduate or graduate full-year enrollment ²	✓	✓	✓
Control and level of institution	✓	✓	
Region of institution	✓	✓	
State of institution			✓
Public 2-year and public 4-year institution sectors within state of institution			✓
Student characteristics			
Student age as of December 31, 2019 ²	✓	✓	✓
Gender	✓	✓	✓
Race/ethnicity	✓	✓	✓
Veteran status	✓	✓	✓
CPS parents' highest education	✓	✓	
CPS has dependents (yes/no)	✓	✓	
CPS marital status	✓	✓	
CPS income ²	✓	✓	
FTB status (yes/no)	✓		✓
Graduate student type (graduate [excluding doctoral–professional practice] or doctoral–professional practice)		✓	
Student aid			
Federal aid recipient (yes/no)	✓	✓	✓
Direct Loan amount disbursed ²	✓	✓	✓
State aid recipient (yes/no)	✓	✓	✓
Institution aid recipient (yes/no)	✓	✓	✓
Any aid receipt excluding private loans (yes/no)	✓	✓	
Federal Pell Grant amount awarded ²	✓		✓
PLUS Loan amount disbursed ²	✓		

¹ Institution-level predictor variables came from Integrated Postsecondary Education Data System (IPEDS) 2019–20 files, and student-level predictor variables came from NPSAS:20 institution enrollment lists and CPS for aid applicants.

² Categories were defined by quartiles for continuous variables. Categories of "None" or "Unknown" were excluded from quartile calculations.

NOTE: CPS = Central Processing System; FTB = first-time beginning student.

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

Similar to what was done for the institution nonresponse adjustment, the CHAID algorithm was used to identify two-way and three-way interaction terms for the student nonresponse adjustment models (see section 7.1.3.6.1).

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 selection method until converging models were attained. Categories of candidate predictor variables that impeded the creation of a convergent model were first collapsed to attempt to achieve model convergence. If the model still failed to converge after collapsing categories, the variable was dropped from the final model.

Different bounds were used on the nonresponse weight adjustments, depending on whether the weight was classified as high extreme,⁵⁰ low extreme,⁵¹ or nonextreme, to accomplish nonresponse adjustment, truncation, and smoothing in one step. In addition, the data were inspected for extreme weights by checking for outliers by state and institution stratum within state for the undergraduate study weight and by institution control and level for the undergraduate survey weight and the graduate weight.

7.1.3.6.1 *Undergraduate student survey nonresponse adjustment (WTA_adj5)*

There were three stages of weight adjustments for nonresponse: first, inability to locate the student, then survey refusal, and finally, other nonresponse (student located, did not refuse, did not respond)—because the predictors of response propensity were potentially different for each of these nonresponse outcomes, as described in section 7.1.3.6.

Undergraduate student survey not located nonresponse adjustment (WTA_adj5a). The first type of adjustment for undergraduate student survey nonresponse was an adjustment for the inability to locate the student⁵² (“not located”).

Table 68 shows the final predictor variables used in the model to determine the undergraduate student survey not located nonresponse weight adjustments and the average weight adjustment factors resulting from these variables. Summary statistics of the weight adjustment factors were as follows:

⁵⁰ High extreme weights were identified as those greater than the median weight plus three times the interquartile range for weights.

⁵¹ Low extreme weights were identified as those less than the median weight minus three times the interquartile range for weights.

⁵² Refer to section 4.2.3 for further details on the process of locating sample members.

- minimum: 1.02;
- median: 1.21; and
- maximum: 4.23.

The final lower bound was 1.0, and the final upper bound was 15.0 to achieve convergence for this weight adjustment.

Table 68. Weight adjustment factors for undergraduate survey student not located nonresponse adjustment, by model predictor variable: 2019–20

Model predictor variable	Number of students who were located	Weighted response rate ¹	Average weight adjustment factor (WTA_adj5a)
Total	106,850	79.70	1.28
Control and level of institution			
Public less-than-2-year	980	75.40	1.38
Public 2-year	36,510	75.40	1.33
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	6,880	78.50	1.28
Public 4-year, non-doctorate-granting, primarily baccalaureate	7,200	85.20	1.18
Public 4-year, doctorate-granting	21,340	82.50	1.23
Private nonprofit 2-year or less	920	74.00	1.39
Private nonprofit 4-year, non-doctorate-granting	7,490	85.40	1.18
Private nonprofit 4-year, doctorate-granting	9,210	81.90	1.24
Private for-profit less-than-2-year	3,630	69.70	1.43
Private for-profit 2-year	5,040	69.40	1.39
Private for-profit 4-year	7,660	79.60	1.28
Bureau of Economic Analysis region ²			
New England	6,460	82.20	1.25
Mideast	15,330	79.80	1.28
Great Lakes	12,620	79.90	1.28
Plains	8,050	81.80	1.26
Southeast	27,070	81.30	1.26
Southwest	12,640	78.60	1.31
Rocky Mountains	6,050	83.60	1.25
Far West	16,640	75.70	1.34
Outlying Areas	1,990	78.50	1.29
Institution total undergraduate enrollment ³			
1–2,882	26,440	79.70	1.30
2,883–9,097	26,840	80.70	1.27
9,098–21,262	26,550	79.40	1.28
21,263 or more	27,020	79.40	1.28
Age as of December 31, 2019			
15–23	72,920	80.30	1.28
24–29	15,240	77.90	1.31
30 or more	18,690	79.10	1.29
Veteran status			
Yes	3,530	79.30	1.29
No	103,320	79.70	1.28

See notes at end of table.

Table 68. Weight adjustment factors for undergraduate survey student not located nonresponse adjustment, by model predictor variable: 2019–20—Continued

Model predictor variable	Number of students who were located	Weighted response rate ¹	Average weight adjustment factor (WTA_adj5a)
Collapsed race/ethnicity			
Asian, not Hispanic or Latino	6,860	81.30	1.23
Black, not Hispanic or Latino	15,230	83.50	1.21
Hispanic or Latino of any race	23,220	82.70	1.22
White, not Hispanic or Latino	49,670	83.40	1.21
Other, not Hispanic or Latino or unknown	11,880	56.60	1.83
Gender			
Male	43,700	77.30	1.32
Female	63,150	81.50	1.25
Total amount of Direct Loans disbursed to undergraduate students ³			
\$0	67,220	76.90	1.33
\$1–\$4,500	9,640	83.80	1.24
\$4,501–\$5,500	10,630	86.20	1.20
\$5,501–\$8,617	9,400	87.00	1.19
\$8,618 or more	9,970	85.90	1.2
Institutional aid status			
Received	27,500	86.50	1.18
Did not receive	57,240	78.40	1.29
Unknown	22,110	73.00	1.39
State aid status			
Received	19,960	85.90	1.19
Did not receive	64,250	79.50	1.28
Unknown	22,640	73.10	1.39
Federal aid status			
Received	65,260	84.70	1.22
Did not receive	31,610	75.50	1.34
Unknown	9,980	68.10	1.51
Any aid status excluding private loans			
Received	79,440	84.50	1.22
Did not receive	18,380	71.40	1.42
Unknown	9,040	65.40	1.57
Total amount of PLUS Loans disbursed to undergraduate students ³			
\$0	101,270	79.40	1.29
\$1–\$7,059	1,340	84.70	1.25
\$7,060–\$12,414	1,390	85.20	1.24
\$12,415–\$20,000	1,410	86.60	1.21
\$20,001 or more	1,450	88.30	1.16
Total amount of federal Pell Grants awarded to undergraduate students ⁴			
\$0	60,790	77.20	1.33
\$1–\$3,097	13,230	80.90	1.29
\$3,098–\$6,194	17,440	84.50	1.23
\$6,195 or more	15,390	87.40	1.18

See notes at end of table.

Table 68. Weight adjustment factors for undergraduate survey student not located nonresponse adjustment, by model predictor variable: 2019–20—Continued

Model predictor variable	Number of students who were located	Weighted response rate ¹	Average weight adjustment factor (WTA_adj5a)
CPS—Parents' highest education			
Middle school/junior high	3,480	84.20	1.22
High school	21,230	84.20	1.22
College or beyond	34,590	87.20	1.17
Other/unknown	4,170	81.80	1.27
Missing	730	74.70	1.4
Not matched in CPS:20	42,660	71.20	1.42
CPS—Has dependents			
Yes	11,730	84.30	1.23
No	52,460	86.00	1.19
Not matched in CPS:20	42,660	71.20	1.42
CPS—Marital status			
Single	55,890	85.60	1.2
Married/remarried	6,080	87.10	1.17
Separated	800	82.60	1.25
Divorced/widowed	1,420	85.50	1.2
Not matched in CPS:20	42,660	71.20	1.42
CPS—Total income ³			
-\$1,017,256–\$14,162	15,080	83.40	1.24
\$14,163–\$34,248	15,640	84.40	1.21
\$34,249–\$81,192	16,210	86.60	1.18
\$81,193 or more	16,720	88.00	1.15
Missing	550	77.40	1.32
Not matched in CPS:20	42,660	71.20	1.42
First-time beginner status (sampled)			
FTB	38,430	79.70	1.29
Not FTB	68,420	79.70	1.28
CHAID segments			
Undergraduate students with Black, not Hispanic or Latino or White, not Hispanic or Latino race/ethnicity whose parents' highest education in CPS:20 was college or beyond and the institution's control and level was private nonprofit 4-year, non-doctorate-granting or private nonprofit 4-year, doctorate-granting	4,980	91.50	1.09
Undergraduate students with Black, not Hispanic or Latino or White, not Hispanic or Latino race/ethnicity whose parents' highest education in CPS:20 was college or beyond and the institution's control and level was public 4-year, non-doctorate-granting, primarily baccalaureate; public 5-year, doctorate-granting; or private for-profit 2-year	8,600	89.80	1.11
Undergraduate students with Black, not Hispanic or Latino or White, not Hispanic or Latino race/ethnicity whose parents' highest education in CPS:20 was college or beyond and the institution's control and level was public less-than-2-year; public 2-year; public 4-year, non-doctorate-granting, primarily subbaccalaureate; private nonprofit less-than-4-year; or private for-profit 4-year	9,790	85.30	1.18
Undergraduate students with Black, not Hispanic or Latino or White, not Hispanic or Latino race/ethnicity whose parents' highest education in CPS:20 was college or beyond and the institution's control and level was private for-profit less-than-2-year	440	78.60	1.26

See notes at end of table.

Table 68. Weight adjustment factors for undergraduate survey student not located nonresponse adjustment, by model predictor variable: 2019–20—Continued

Model predictor variable	Number of students who were located	Weighted response rate ¹	Average weight adjustment factor (WTA_adj5a)
Undergraduate students with Black, not Hispanic or Latino or White, not Hispanic or Latino race/ethnicity whose parents' highest education in CPS:20 was missing or did not match in CPS:20 and who received any aid excluding private loans	14,150	81.70	1.23
Undergraduate students with Black, not Hispanic or Latino or White, not Hispanic or Latino race/ethnicity whose parents' highest education in CPS:20 was missing or did not match in CPS:20 and for whom it is unknown if they received any aid excluding private loans	3,290	79.20	1.27
Undergraduate students with Black, not Hispanic or Latino or White, not Hispanic or Latino race/ethnicity whose parents' highest education in CPS:20 was missing or did not match in CPS:20 and who did not receive any aid excluding private loans	8,560	73.00	1.39
Undergraduate students with Black, not Hispanic or Latino or White, not Hispanic or Latino race/ethnicity whose parents' highest education in CPS:20 was middle school/junior high or high school and for whom it is unknown if they received institution aid	1,780	90.80	1.11
Undergraduate students with Black, not Hispanic or Latino or White, not Hispanic or Latino race/ethnicity whose parents' highest education in CPS:20 was middle school/junior high or high school and who did not receive institution aid	8,350	83.10	1.22
Undergraduate students with Black, not Hispanic or Latino or White, not Hispanic or Latino race/ethnicity whose parents' highest education in CPS:20 was middle school/junior high or high school and who received institution aid	3,390	88.90	1.13
Undergraduate students with Black, not Hispanic or Latino or White, not Hispanic or Latino race/ethnicity whose parents' highest education in CPS:20 was other/unknown and whose marital status was single or separated in CPS:20	1,210	79.50	1.28
Undergraduate students with Black, not Hispanic or Latino or White, not Hispanic or Latino race/ethnicity whose parents' highest education in CPS:20 was other/unknown and whose marital status was married/remarried or divorced/widowed in CPS:20	370	91.40	1.1
Undergraduate students with Other, not Hispanic or Latino or unknown race/ethnicity whose total income in CPS:20 was – \$1,017,256 to \$34,248 and for whom it is unknown if they received institution aid	1,400	54.30	1.91
Undergraduate students with Other, not Hispanic or Latino or unknown race/ethnicity whose total income in CPS:20 was –\$1,017,256 to \$34,248 and who did not receive institution aid	1,620	73.60	1.4
Undergraduate students with Other, not Hispanic or Latino or unknown race/ethnicity whose total income in CPS:20 was –\$1,017,256 to \$34,248 and who received institution aid	530	82.20	1.24
Undergraduate students with Other, not Hispanic or Latino or unknown race/ethnicity whose total income in CPS:20 was \$34,249 to \$81,192 and for whom it is unknown if they received institution aid	580	59.00	1.74
Undergraduate students with Other, not Hispanic or Latino or unknown race/ethnicity whose total income in CPS:20 was \$34,249 to \$81,192 and who did not receive institution aid	600	77.40	1.32
Undergraduate students with Other, not Hispanic or Latino or unknown race/ethnicity whose total income in CPS:20 was \$34,249 to \$81,192 and who received institution aid	360	86.40	1.17
Undergraduate students with Other, not Hispanic or Latino or unknown race/ethnicity whose total income in CPS:20 was \$81,193 or more and for whom it is unknown if they received institution aid	370	63.40	1.61

See notes at end of table.

Table 68. Weight adjustment factors for undergraduate survey student not located nonresponse adjustment, by model predictor variable: 2019–20—Continued

Model predictor variable	Number of students who were located	Weighted response rate ¹	Average weight adjustment factor (WTA_adj5a)
Undergraduate students with Other, not Hispanic or Latino or unknown race/ethnicity whose total income in CPS:20 was \$81,193 or more and who did not receive institution aid	500	80.90	1.26
Undergraduate students with Other, not Hispanic or Latino or unknown race/ethnicity whose total income in CPS:20 was \$81,193 or more and who received institution aid	390	87.60	1.15
Undergraduate students with Other, not Hispanic or Latino or unknown race/ethnicity who were not matched in CPS:20 and who received any aid excluding private loans	2,470	53.40	1.89
Undergraduate students with Other, not Hispanic or Latino or unknown race/ethnicity who were not matched in CPS:20 and for whom it is unknown if they received any aid excluding private loans	1,930	38.90	2.64
Undergraduate students with Other, not Hispanic or Latino or unknown race/ethnicity who were not matched in CPS:20 and did not receive any aid excluding private loans	1,070	46.30	2.22
Undergraduate students with Other, not Hispanic or Latino or unknown race/ethnicity whose total income in CPS:20 was missing	60	73.90	1.39
Undergraduate students with Hispanic or Latino ethnicity of any race whose parents' highest education in CPS:20 was high school or college or beyond and for whom it is unknown if they received institution aid	1,520	93.50	1.07
Undergraduate students with Hispanic or Latino ethnicity of any race whose parents' highest education in CPS:20 was high school or college or beyond and who did not receive institution aid	6,500	84.20	1.2
Undergraduate students with Hispanic or Latino ethnicity of any race whose parents' highest education in CPS:20 was high school or college or beyond and who received institution aid	2,940	90.50	1.11
Undergraduate students with Hispanic or Latino ethnicity of any race whose parents' highest education in CPS:20 was missing or did not match in CPS:20 and who received any aid excluding private loans or for whom it is unknown if they received any aid excluding private loans	6,080	79.70	1.27
Undergraduate students with Hispanic or Latino ethnicity of any race whose parents' highest education in CPS:20 was missing or did not match in CPS:20 and who did not receive any aid excluding private loans	2,500	70.50	1.44
Undergraduate students with Hispanic or Latino ethnicity of any race whose parents' highest education in CPS:20 was middle school/junior high or other/unknown and who received institution aid or for whom it is unknown if they received institution aid	1,340	89.50	1.12
Undergraduate students with Hispanic or Latino ethnicity of any race whose parents' highest education in CPS:20 was middle school/junior high or other/unknown and who did not receive institution aid	2,350	83.50	1.21
Undergraduate students with Asian, not Hispanic or Latino race/ethnicity who received any aid excluding private loans and whose parents' highest education from CPS:20 was middle school/junior high, high school, college or beyond, or other/unknown	3,120	88.70	1.13
Undergraduate students with Asian, not Hispanic or Latino race/ethnicity who received any aid excluding private loans and whose parents' highest education from CPS:20 was missing or not matched in CPS:20	1,460	82.20	1.22
Undergraduate students with Asian, not Hispanic or Latino race/ethnicity for whom it is unknown if they received any aid excluding private loans and whose institution's undergraduate enrollment was 1–21,262	370	77.10	1.32

See notes at end of table.

Table 68. Weight adjustment factors for undergraduate survey student not located nonresponse adjustment, by model predictor variable: 2019–20—Continued

Model predictor variable	Number of students who were located	Weighted response rate ¹	Average weight adjustment factor (WTA_adj5a)
Undergraduate students with Asian, not Hispanic or Latino race/ethnicity for whom it is unknown if they received any aid excluding private loans and whose institution's undergraduate enrollment was 21,263 or more	300	88.60	1.13
Undergraduate students with Asian, not Hispanic or Latino race/ethnicity who did not receive any aid excluding private loans and whose total income in CPS:20 was -\$1,017,256 to \$34,248	80	76.30	1.33
Undergraduate students with Asian, not Hispanic or Latino race/ethnicity who did not receive any aid excluding private loans and whose total income in CPS:20 was \$34,249 or more	270	86.90	1.16
Undergraduate students with Asian, not Hispanic or Latino race/ethnicity who did not receive any aid excluding private loans and whose total income in CPS:20 was missing or did not match in CPS:20	1,260	67.50	1.5

¹ The weighted response rates were calculated using the student base weight (the student sampling 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 Categories were defined by quartiles. Fewer than four categories will display in cases where the variable's values cause quartiles to overlap. Categories of "None" or "Unknown" were excluded from quartile calculations.

⁴ For Total amount of federal Pell Grants awarded to undergraduate students, the second and third categories were defined by the median of Pell amounts greater than \$0 and less than \$6,195. Categories were defined by quartiles. Fewer than four categories will display in cases where the variable's values cause quartiles to overlap. Categories of "None" or "Unknown" were excluded from quartile calculations.

NOTE: CHAID = chi-square automatic interaction detection; CPS = Central Processing System; FTB = first-time beginning student. Number of respondents rounded to the nearest 10. Detail may not sum to totals because of rounding.

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

Undergraduate student survey refusal nonresponse adjustment

(WTA_adj5b). The second stage of the undergraduate student survey nonresponse adjustment was an adjustment for refusal, given that the student was located. This additional type of nonresponse adjustment was made to compensate further for the potential student nonresponse bias.

Table 69 shows the final predictor variables used in the model to determine the undergraduate student survey refusal nonresponse weight adjustments and the average weight adjustment factors resulting from these variables. Summary statistics of the weight adjustment factors were

- minimum: 1.00;
- median: 1.04; and
- maximum: 1.58.

The final lower bound was 1.0 and the final upper bound was 15.0 to achieve convergence for this weight adjustment.

Table 69. Weight adjustment factors for undergraduate survey student refusal nonresponse adjustment, by model predictor variable: 2019–20

Model predictor variable	Number of students who were located but refused to complete the survey	Weighted response rate ¹	Average weight adjustment factor (WTA_adj5b)
Total	102,240	95.10	1.05
Control and level of institution			
Public less-than-2-year	920	94.60	1.06
Public 2-year	34,650	94.20	1.05
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	6,510	93.90	1.06
Public 4-year, non-doctorate-granting, primarily baccalaureate	6,900	95.30	1.04
Public 4-year, doctorate-granting	20,540	95.70	1.04
Private nonprofit 2-year or less	870	96.10	1.04
Private nonprofit 4-year, non-doctorate-granting	7,260	96.10	1.03
Private nonprofit 4-year, doctorate-granting	8,890	96.30	1.03
Private for-profit less-than-2-year	3,470	94.20	1.05
Private for-profit 2-year	4,830	92.80	1.06
Private for-profit 4-year	7,410	96.00	1.03
Bureau of Economic Analysis region ²			
New England	6,150	95.80	1.04
Mideast	14,640	95.20	1.05
Great Lakes	12,070	95.20	1.05
Plains	7,750	95.80	1.04
Southeast	25,960	95.10	1.05
Southwest	12,080	95.00	1.05
Rocky Mountains	5,770	94.80	1.05
Far West	15,900	94.70	1.05
Outlying Areas	1,930	97.10	1.03
Institution total undergraduate enrollment ³			
1–2,882	25,480	95.90	1.04
2,883–9,097	25,690	95.10	1.05
9,098–21,262	25,310	95.00	1.05
21,263 or more	25,760	95.00	1.05
Age as of December 31, 2019			
15–23	70,000	95.70	1.04
24–29	14,550	94.70	1.05
30 or older	17,690	93.70	1.06
Veteran status			
Yes	3,290	92.30	1.07
No	98,950	95.20	1.04
Collapsed race/ethnicity			
Asian, not Hispanic or Latino	6,640	96.20	1.04
Black, not Hispanic or Latino	14,800	96.70	1.03
Hispanic or Latino of any race	22,520	96.80	1.03
White, not Hispanic or Latino	47,520	95.30	1.05
Other, not Hispanic or Latino or unknown	10,770	89.60	1.11
Gender			
Male	41,200	93.70	1.06
Female	61,040	96.20	1.04

See notes at end of table.

Table 69. Weight adjustment factors for undergraduate survey student refusal nonresponse adjustment, by model predictor variable: 2019–20—Continued

Model predictor variable	Number of students who were located but refused to complete the survey	Weighted response rate ¹	Average weight adjustment factor (WTA_adj5b)
Total amount of Direct Loans disbursed to undergraduate students ³			
\$0	63,850	94.40	1.05
\$1–\$4,500	9,310	96.20	1.04
\$4,501–\$5,500	10,280	96.90	1.03
\$5,501–\$8,617	9,100	96.70	1.03
\$8,618 or more	9,690	96.90	1.03
Institutional aid status			
Received	26,550	96.20	1.04
Did not receive	54,570	94.90	1.05
Unknown	21,120	94.10	1.05
State aid status			
Received	19,290	96.60	1.03
Did not receive	61,340	94.90	1.05
Unknown	21,620	94.00	1.05
Federal aid status			
Received	62,980	96.30	1.04
Did not receive	29,830	94.00	1.06
Unknown	9,440	92.80	1.06
Any aid status excluding private loans			
Received	76,630	96.30	1.04
Did not receive	17,130	92.90	1.07
Unknown	8,480	92.20	1.06
Total amount of PLUS Loans disbursed to undergraduate students ³			
\$0	96,870	95.10	1.05
\$1–\$7,059	1,290	96.00	1.04
\$7,060–\$12,414	1,340	95.20	1.05
\$12,415–\$20,000	1,350	96.40	1.04
\$20,001 or more	1,390	95.50	1.05
Total amount of federal Pell Grants awarded to undergraduate students ⁴			
\$0	57,740	94.40	1.05
\$1–\$3,097	12,680	95.70	1.04
\$3,098–\$6,194	16,850	96.50	1.04
\$6,195 or more	14,970	97.10	1.03
CPS—Parents' highest education			
Middle school/junior high	3,350	96.30	1.04
High school	20,420	96.20	1.04
College or beyond	33,270	96.20	1.04
Other/unknown	3,970	95.30	1.05
Missing	690	95.80	1.04
Not matched in CPS:20	40,540	93.70	1.06

See notes at end of table.

Table 69. Weight adjustment factors for undergraduate survey student refusal nonresponse adjustment, by model predictor variable: 2019–20—Continued

Model predictor variable	Number of students who were located but refused to complete the survey	Weighted response rate ¹	Average weight adjustment factor (WTA_adj5b)
CPS—Has dependents			
Yes	11,300	96.10	1.04
No	50,400	96.20	1.04
Not matched in CPS:20	40,540	93.70	1.06
CPS—Marital status			
Single	53,730	96.20	1.04
Married/remarried	5,820	95.50	1.05
Separated	780	98.00	1.02
Divorced/widowed	1,370	96.20	1.04
Not matched in CPS:20	40,540	93.70	1.06
CPS—Total income ³			
-\$1,017,256–\$14,162	14,510	96.20	1.04
\$14,163–\$34,248	15,090	96.10	1.04
\$34,249–\$81,192	15,560	96.10	1.04
\$81,193 or more	16,040	96.20	1.04
Missing	520	95.60	1.05
Not matched in CPS:20	40,540	93.70	1.06
First-time beginner status (sampled)			
FTB	36,760	95.60	1.04
Not FTB	65,480	95.00	1.05
CHAID segments			
Undergraduate students with Black, not Hispanic or Latino race/ethnicity or Hispanic or Latino of any race who received any aid excluding private loans or for whom it is unknown if they received any aid excluding private loans and whose gender is female	20,700	97.80	1.02
Undergraduate students with Black, not Hispanic or Latino race/ethnicity or Hispanic or Latino of any race who received any aid excluding private loans or for whom it is unknown if they received any aid excluding private loans and whose gender is male	11,470	96.60	1.03
Undergraduate students with Black, not Hispanic or Latino race/ethnicity or Hispanic or Latino of any race who did not receive any aid excluding private loans and whose gender is female	2,930	95.60	1.05
Undergraduate students with Black, not Hispanic or Latino race/ethnicity or Hispanic or Latino of any race who did not receive any aid excluding private loans and whose gender is male	2,220	93.10	1.07
Undergraduate students with White, not Hispanic or Latino race/ethnicity who received any aid excluding private loans or for whom it is unknown if they received any aid excluding private loans and whose gender is female	22,810	97.10	1.03
Undergraduate students with White, not Hispanic or Latino race/ethnicity who received any aid excluding private loans or for whom it is unknown if they received any aid excluding private loans and whose gender is male	15,570	95.20	1.05
Undergraduate students with White, not Hispanic or Latino race/ethnicity who did not receive any aid excluding private loans and who matched in CPS:20	2,300	95.00	1.05

See notes at end of table.

Table 69. Weight adjustment factors for undergraduate survey student refusal nonresponse adjustment, by model predictor variable: 2019–20—Continued

Model predictor variable	Number of students who were located but refused to complete the survey	Weighted response rate ¹	Average weight adjustment factor (WTA_adj5b)
Undergraduate students with White, not Hispanic or Latino race/ethnicity who did not receive any aid excluding private loans and who did not match in CPS:20	6,840	91.70	1.09
Undergraduate students with Other, not Hispanic or unknown race/ethnicity for whom it is unknown if they received state aid and whose institution's control and level was public less-than-2-year; private nonprofit 4-year, non-doctorate-granting; or private for-profit 2-year	840	83.10	1.19
Undergraduate students with Other, not Hispanic or unknown race/ethnicity for whom it is unknown if they received state aid and whose institution's control and level was public 4-year, non-doctorate-granting, primarily baccalaureate or private for-profit less-than-2-year	680	89.90	1.11
Undergraduate students with Other, not Hispanic or unknown race/ethnicity for whom it is unknown if they received state aid and whose institution's control and level was public 2-year; public 4-year, non-doctorate-granting, primarily subbaccalaureate; public 4-year, doctorate-granting; private nonprofit less-than-4-year; or private nonprofit 4-year, doctorate-granting	2,960	86.60	1.14
Undergraduate students with Other, not Hispanic or unknown race/ethnicity for whom it is unknown if they received state aid and whose institution's control and level was private for-profit 4-year	290	96.40	1.03
Undergraduate students with Other, not Hispanic or unknown race/ethnicity who received state aid and whose parents' highest education from CPS:20 was middle school/junior high, high school, or college or beyond	860	96.20	1.04
Undergraduate students with Other, not Hispanic or unknown race/ethnicity who received state aid and whose parents' highest education from CPS:20 was other/unknown or missing	60	100.0	1
Undergraduate students with Other, not Hispanic or Latino or unknown race/ethnicity who received state aid and who did not match into CPS:20	260	87.80	1.13
Undergraduate students with Other, not Hispanic or Latino or unknown race/ethnicity who did not receive state aid and whose age as of December 31, 2019, was 15–23	3,040	93.70	1.06
Undergraduate students with Other, not Hispanic or Latino or unknown race/ethnicity who did not receive state aid and whose age as of December 31, 2019, was 24–29	850	90.20	1.1
Undergraduate students with Other, not Hispanic or Latino or unknown race/ethnicity who did not receive state aid and whose age as of December 31, 2019, was 30 or older	940	86.20	1.15
Undergraduate students with Asian, not Hispanic or Latino race/ethnicity whose total amount of PLUS Loans disbursed was \$0 and whose institution's total undergraduate enrollment was 1–2,882	980	99.20	1.01
Undergraduate students with Asian, not Hispanic or Latino race/ethnicity whose total amount of PLUS Loans disbursed was \$0 and whose institution's total undergraduate enrollment was 2,883 or more	5,410	96.00	1.04
Undergraduate students with Asian, not Hispanic or Latino race/ethnicity whose total amount of PLUS Loans disbursed was \$1–\$12,414	80	100.0	1

See notes at end of table.

Table 69. Weight adjustment factors for undergraduate survey student refusal nonresponse adjustment, by model predictor variable: 2019–20—Continued

Model predictor variable	Number of students who were located but refused to complete the survey	Weighted response rate ¹	Average weight adjustment factor (WTA_adj5b)
Undergraduate students with Asian, not Hispanic or Latino race/ethnicity whose total amount of PLUS Loans disbursed was \$12,415 or more and whose gender is female	100	99.90	1
Undergraduate students with Asian, not Hispanic or Latino race/ethnicity whose total amount of PLUS Loans disbursed was \$12,415 or more and whose gender is male	70	85.10	1.15

¹ The weighted response rates were calculated using the student base weight adjusted for not located nonresponse.

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 Categories were defined by quartiles. Fewer than four categories will display in cases where the variable's values cause quartiles to overlap. Categories of "None" or "Unknown" were excluded from quartile calculations.

⁴ For Total amount of federal Pell Grants awarded to undergraduate students, the second and third categories were defined by the median of Pell amounts greater than \$0 and less than \$6,195. Categories were defined by quartiles. Fewer than four categories will display in cases where the variable's values cause quartiles to overlap. Categories of "None" or "Unknown" were excluded from quartile calculations.

NOTE: CHAID = chi-square automatic interaction detection; CPS = Central Processing System; FTB = first-time beginning student. Number of respondents rounded to the nearest 10. Detail may not sum to totals because of rounding.

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

Undergraduate student survey other nonresponse adjustment (WTA_adj5c).

The third, and final, stage of adjustment for undergraduate student survey nonresponse was an adjustment for other nonresponse (e.g., contacted, but not interviewed before the end of the data collection period), given that the student was located and did not explicitly refuse to participate.

Table 70 shows the final predictor variables used in the model to determine the undergraduate student survey other nonresponse weight adjustments and the average weight adjustment factor resulting from these variables. Summary statistics of the weight adjustment factors were

- minimum: 1.01;
- median: 1.18; and
- maximum: 11.39.

The final lower bound was 1.0 and the final upper bound was 15.0 to achieve convergence for this weight adjustment.

Table 70. Weight adjustment factors for undergraduate survey student other nonresponse adjustment, by model predictor variable: 2019–20

Model predictor variable	Number of respondents	Weighted response rate ¹	Average weight adjustment factor (WTA_adj5c)
Total	80,760	78.20	1.26
Control and level of institution			
Public less-than-2-year	660	75.50	1.35
Public 2-year	25,940	71.70	1.33
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	5,020	75.10	1.29
Public 4-year, non-doctorate-granting, primarily baccalaureate	5,650	81.20	1.22
Public 4-year, doctorate-granting	17,240	82.90	1.19
Private nonprofit 2-year or less	650	72.60	1.35
Private nonprofit 4-year, non-doctorate-granting	6,230	84.70	1.16
Private nonprofit 4-year, doctorate-granting	7,620	84.10	1.18
Private for-profit less-than-2-year	2,470	65.10	1.43
Private for-profit 2-year	3,390	64.90	1.34
Private for-profit 4-year	5,890	78.20	1.26
Bureau of Economic Analysis (BEA) region ²			
New England	4,800	79.30	1.27
Mideast	11,630	78.40	1.26
Great Lakes	9,610	79.50	1.25
Plains	6,280	80.60	1.24
Southeast	20,430	78.80	1.26
Southwest	9,360	77.20	1.28
Rocky Mountains	4,570	80.60	1.25
Far West	12,550	75.00	1.28
Outlying Areas	1,520	78.90	1.26
Institution total undergraduate enrollment ³			
1–2,882	20,130	79.00	1.26
2,883–9,097	20,300	78.30	1.27
9,098–21,262	19,870	77.40	1.27
21,263 or more	20,460	78.30	1.25
Age as of December 31, 2019			
15–23	55,770	79.60	1.25
24–29	11,380	76.40	1.28
30 or more	13,610	74.90	1.31
Veteran status			
Yes	2,550	75.30	1.31
No	78,210	78.30	1.26
Collapsed race/ethnicity			
Asian, not Hispanic or Latino	5,840	87.20	1.14
Black, not Hispanic or Latino	11,930	81.20	1.23
Hispanic or Latino of any race	18,500	82.10	1.22
White, not Hispanic or Latino	39,020	81.90	1.22
Other, not Hispanic or Latino or unknown	5,470	49.90	1.95
Gender			
Male	30,850	73.80	1.33
Female	49,910	81.40	1.22

See notes at end of table.

Table 70. Weight adjustment factors for undergraduate survey student other nonresponse adjustment, by model predictor variable: 2019–20—Continued

Model predictor variable	Number of respondents	Weighted response rate ¹	Average weight adjustment factor (WTA_adj5c)
Total amount of Direct Loans disbursed to undergraduate students ³			
\$0	49,040	75.20	1.30
\$1–\$4,500	7,460	82.50	1.23
\$4,501–\$5,500	8,490	83.80	1.21
\$5,501–\$8,617	7,690	85.60	1.18
\$8,618 or more	8,080	85.80	1.18
Institutional aid status			
Received	22,640	85.90	1.17
Did not receive	42,030	77.20	1.29
Unknown	16,090	69.00	1.3
State aid status			
Received	16,070	84.70	1.2
Did not receive	48,260	78.50	1.27
Unknown	16,430	68.90	1.31
Federal aid status			
Received	51,290	83.00	1.22
Did not receive	22,680	75.30	1.32
Unknown	6,790	62.90	1.4
Any aid status excluding private loans			
Received	62,930	83.40	1.21
Did not receive	12,030	70.00	1.44
Unknown	5,800	59.90	1.46
Total amount of PLUS Loans disbursed to undergraduate students ³			
\$0	76,310	77.90	1.27
\$1–\$7,059	1,030	83.70	1.24
\$7,060–\$12,414	1,100	83.60	1.22
\$12,415–\$20,000	1,130	83.80	1.21
\$20,001 or more	1,200	86.80	1.17
Total amount of federal Pell Grants awarded to undergraduate students ⁴			
\$0	44,870	75.90	1.29
\$1–\$3,097	9,800	78.70	1.28
\$3,098–\$6,194	13,650	83.30	1.22
\$6,195 or more	12,450	84.90	1.19
CPS—Parents' highest education			
Middle school/junior high	2,660	81.10	1.24
High school	16,330	81.80	1.23
College or beyond	27,530	84.10	1.19
Other/unknown	2,970	76.50	1.32
Missing	490	74.90	1.37
Not matched in CPS:20	30,790	71.70	1.33
CPS—Has dependents			
Yes	8,970	81.50	1.24
No	40,990	83.00	1.21
Not matched in CPS:20	30,790	71.70	1.33

See notes at end of table.

Table 70. Weight adjustment factors for undergraduate survey student other nonresponse adjustment, by model predictor variable: 2019–20—Continued

Model predictor variable	Number of respondents	Weighted response rate ¹	Average weight adjustment factor (WTA_adj5c)
CPS—Marital status			
Single	43,430	82.60	1.22
Married/remarried	4,780	83.40	1.2
Separated	640	85.00	1.19
Divorced/widowed	1,120	84.50	1.19
Not matched in CPS:20	30,790	71.70	1.33
CPS—Total income ³			
-\$1,017,256–\$14,162	11,260	80.10	1.26
\$14,163–\$34,248	12,140	82.00	1.22
\$34,249–\$81,192	12,790	83.90	1.2
\$81,193 or more	13,420	84.80	1.19
Missing	360	69.70	1.48
Not matched in CPS:20	30,790	71.70	1.33
First-time beginner status (sampled)			
FTB	28,330	78.00	1.28
Not FTB	52,430	78.20	1.25
CHAID segments			
Undergraduate students with Black, not Hispanic or Latino race/ethnicity for whom it is unknown if they received institution aid and whose institution's BEA region was Mideast, Southeast, or Southwest	1,730	91.10	1.1
Undergraduate students with Black, not Hispanic or Latino race/ethnicity for whom it is unknown if they received institution aid and whose institution's BEA region was New England or Rocky Mountains	130	82.90	1.21
Undergraduate students with Black, not Hispanic or Latino race/ethnicity for whom it is unknown if they received institution aid and whose institution's BEA region was Great Lakes, Plains, Far West, or Outlying Areas	470	95.90	1.04
Undergraduate students with Black, not Hispanic or Latino race/ethnicity who did not receive institution aid and who received any other aid excluding private loans	5,620	81.30	1.25
Undergraduate students with Black, not Hispanic or Latino race/ethnicity who did not receive institution aid or any aid excluding private loans or for whom it is unknown if they received any aid excluding private loans	1,210	69.20	1.47
Undergraduate students with Black, not Hispanic or Latino race/ethnicity who received institution aid and whose institution's total undergraduate enrollment was 1–21,262	2,250	86.80	1.16
Undergraduate students with Black, not Hispanic or Latino race/ethnicity who received institution aid and whose institution's total undergraduate enrollment was 21,263 or more	510	80.60	1.25
Undergraduate students with Hispanic or Latino of any race or White, not Hispanic or Latino race/ethnicity who received any aid excluding private loans and for whom it is unknown if they received institution aid	7,280	93.00	1.08
Undergraduate students with Hispanic or Latino of any race or White, not Hispanic or Latino race/ethnicity who received any aid excluding private loans but who did not receive institution aid	20,530	83.80	1.21
Undergraduate students with Hispanic or Latino of any race or White, not Hispanic or Latino race/ethnicity who received any aid excluding private loans and who received institution aid	16,670	86.60	1.16

See notes at end of table.

Table 70. Weight adjustment factors for undergraduate survey student other nonresponse adjustment, by model predictor variable: 2019–20—Continued

Model predictor variable	Number of respondents	Weighted response rate ¹	Average weight adjustment factor (WTA_adj5c)
Undergraduate students with Hispanic or Latino ethnicity of any race or White, not Hispanic or Latino race/ethnicity for whom it is unknown if they received any aid excluding private loans and whose institution's BEA region was Southeast or Southwest	1,220	76.90	1.31
Undergraduate students with Hispanic or Latino ethnicity of any race or White, not Hispanic or Latino race/ethnicity for whom it is unknown if they received any aid excluding private loans and whose institution's BEA region was Plains or Rocky Mountains	790	89.60	1.13
Undergraduate students with Hispanic or Latino ethnicity of any race or White, not Hispanic or Latino race/ethnicity for whom it is unknown if they received any aid excluding private loans and whose institution's BEA region was Mideast, Great Lakes, or Far West	1,910	85.40	1.17
Undergraduate students with Hispanic or Latino ethnicity of any race or White, not Hispanic or Latino race/ethnicity for whom it is unknown if they received any aid excluding private loans and whose institution's BEA region was New England	190	67.40	1.44
Undergraduate students with Hispanic or Latino ethnicity of any race or White, not Hispanic or Latino race/ethnicity for whom it is unknown if they received any aid excluding private loans and whose institution's BEA region was Outlying Areas	70	97.50	1.02
Undergraduate students with Hispanic or Latino ethnicity of any race or White, not Hispanic or Latino race/ethnicity who did not receive any aid excluding private loans and whose institution's control and level was private nonprofit 4-year, non-doctorate-granting or private nonprofit 4-year, doctorate-granting	850	80.00	1.25
Undergraduate students with Hispanic or Latino ethnicity of any race or White, not Hispanic or Latino race/ethnicity who did not receive any aid excluding private loans and whose institution's control and level was public 4-year, non-doctorate-granting, primarily baccalaureate; public 4-year, doctorate-granting; private nonprofit less-than-4-year; private for-profit less-than-2-year; or private for-profit 4-year	2,840	74.90	1.35
Undergraduate students with Hispanic or Latino ethnicity of any race or White, not Hispanic or Latino race/ethnicity who did not receive any aid excluding private loans and whose institution's control and level was public less-than-2-year; public 2-year; public 4-year, non-doctorate-granting, primarily subbaccalaureate; or private for-profit 2-year	5,190	66.50	1.51
Undergraduate students with Other, not Hispanic or Latino or unknown race/ethnicity for whom it is unknown if they received institution aid and whose institution's control and level was public less-than-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 4-year, non-doctorate-granting; or private nonprofit 4-year, doctorate-granting	480	37.20	2.46
Undergraduate students with Other, not Hispanic or Latino or unknown race/ethnicity for whom it is unknown if they received institution aid and whose institution's control and level was private nonprofit less-than-4-year, private for-profit less-than-2-year, or private for-profit 2-year	260	21.00	4.04
Undergraduate students with Other, not Hispanic or Latino or unknown race/ethnicity for whom it is unknown if they received institution aid and whose institution's control and level was public 2-year	640	25.00	3.73
Undergraduate students with Other, not Hispanic or Latino or unknown race/ethnicity for whom it is unknown if they received institution aid and whose institution's control and level was private for-profit 4-year	120	68.60	1.4
Undergraduate students with Other, not Hispanic or Latino or unknown race/ethnicity who did not receive institution aid and whose marital status from CPS:20 was single or separated	1,490	72.60	1.4

See notes at end of table.

Table 70. Weight adjustment factors for undergraduate survey student other nonresponse adjustment, by model predictor variable: 2019–20—Continued

Model predictor variable	Number of respondents	Weighted response rate ¹	Average weight adjustment factor (WTA_adj5c)
Undergraduate students with Other, not Hispanic or Latino or unknown race/ethnicity who did not receive institution aid and who did not match into CPS:20	890	54.20	1.82
Undergraduate students with Other, not Hispanic or Latino or unknown race/ethnicity who did not receive institution aid and whose marital status from CPS:20 was married/remarried	160	64.80	1.55
Undergraduate students with Other, not Hispanic or Latino or unknown race/ethnicity who did not receive institution aid and whose marital status from CPS:20 was divorced/widowed	60	89.10	1.12
Undergraduate students with Other, not Hispanic or Latino or unknown race/ethnicity who received institution aid and who did not have dependents in CPS:20	860	85.50	1.18
Undergraduate students with Other, not Hispanic or Latino or unknown race/ethnicity who received institution aid and who did not match into CPS:20	410	59.90	1.66
Undergraduate students with Other, not Hispanic or Latino or unknown race/ethnicity who received institution aid and who had dependents in CPS:20	110	79.60	1.27
Undergraduate students with Asian, not Hispanic or Latino race/ethnicity whose institution's control and level was public less-than-2-year; private nonprofit 4-year, non-doctorate-granting; private for-profit less-than-2-year; or private for-profit 2-year and whose total income in CPS:20 was -\$1,017,256–\$14,162	50	80.60	1.25
Undergraduate students with Asian, not Hispanic or Latino race/ethnicity whose institution's control and level was public less-than-2-year; private nonprofit 4-year, non-doctorate-granting; private for-profit less-than-2-year; or private for-profit 2-year and whose total income in CPS:20 was \$14,163 or more, missing, or did not match into CPS:20	600	94.80	1.06
Undergraduate students with Asian, not Hispanic or Latino race/ethnicity whose institution's control and level was public 2-year or public 4-year, non-doctorate-granting, primarily baccalaureate and for whom it is unknown if they received institution aid	540	93.50	1.07
Undergraduate students with Asian, not Hispanic or Latino race/ethnicity whose institution's control and level was public 2-year or public 4-year, non-doctorate-granting, primarily baccalaureate and for whom it is known if they received institution aid	1,540	80.00	1.24
Undergraduate students with Asian, not Hispanic or Latino race/ethnicity whose institution's control and level was public 4-year, doctorate-granting; private nonprofit less-than-4-year; private nonprofit 4-year, doctorate-granting; or private for-profit 4-year and whose gender is female	1,630	92.00	1.08
Undergraduate students with Asian, not Hispanic or Latino race/ethnicity whose institution's control and level was public 4-year, doctorate-granting; private nonprofit less-than-4-year; private nonprofit 4-year, doctorate-granting; or private for-profit 4-year and whose gender is male	1,220	87.50	1.14
Undergraduate students with Asian, not Hispanic or Latino race/ethnicity whose institution's control and level was public 4-year, non-doctorate-granting, primarily subbaccalaureate and whose gender is female	170	89.00	1.12

See notes at end of table.

Table 70. Weight adjustment factors for undergraduate survey student other nonresponse adjustment, by model predictor variable: 2019–20—Continued

Model predictor variable	Number of respondents	Weighted response rate ¹	Average weight adjustment factor (WTA_adj5c)
Undergraduate students with Asian, not Hispanic or Latino race/ethnicity whose institution's control and level was public 4-year, non-doctorate-granting, primarily subbaccalaureate and whose gender is male	100	59.50	1.71

¹ The weighted response rates were calculated using the student base weight adjusted for not located nonresponse and refusal nonresponse.

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 Categories were defined by quartiles. Fewer than four categories will display in cases where the variable's values cause quartiles to overlap. Categories of "None" or "Unknown" were excluded from quartile calculations.

⁴ For Total amount of federal Pell Grants awarded to undergraduate students, the second and third categories were defined by the median of Pell amounts greater than \$0 and less than \$6,195. Categories were defined by quartiles. Fewer than four categories will display in cases where the variable's values cause quartiles to overlap. Categories of "None" or "Unknown" were excluded from quartile calculations.

NOTE: CHAID = chi-square automatic interaction detection; CPS = Central Processing System; FTB = first-time beginning student. Number of respondents rounded to the nearest 10. Detail may not sum to totals because of rounding.

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

7.1.3.6.2 Graduate student nonresponse adjustment (WTA_adj5)

The weighting team adjusted the graduate student weights for nonresponse in two stages: first, inability to locate the student, and then, other nonresponse (student located, did not respond)—because the predictors of response propensity were potentially different for each of these nonresponse outcomes for graduate students. Graduate students were not separately adjusted for refusal because the response propensity for refusal was not different from the response propensity for other nonresponse. As with the undergraduate student survey weight, adjusting for these two types of nonresponse separately achieved greater reduction in nonresponse bias (section 7.2.3) because different characteristics were significant predictors of response propensity at each stage.

Graduate student not located nonresponse adjustment (WTA_adj5a). As with the undergraduate student survey nonresponse adjustments, the first type of adjustment for graduate student nonresponse was an adjustment for the inability to locate the student⁵³ (“not located”).

Table 71 shows the final predictor variables used to determine the graduate student not located nonresponse weight adjustments and the average weight adjustment factors resulting from the graduate student model. Summary statistics of the weight adjustment factors follow:

⁵³ Refer to section 4.2.3 for further details on the process of locating sample members.

- minimum: 1.00;
- median: 1.11; and
- maximum: 6.42.

For the graduate student nonresponse not located weight adjustment (WTA_adj5a), a lower bound of 1 was set on the weight adjustment factors coming out of the nonresponse weight adjustment. No upper bound was set on the weight adjustment factors coming out of the nonresponse weight adjustment.

Table 71. Weight adjustment factors for graduate student not located nonresponse adjustment, by model predictor variable: 2019–20

Model predictor variable	Number of students who were located	Weighted response rate ¹	Average weight adjustment factor (WTA_adj5a)
Total	23,380	85.90	1.17
Institution stratum			
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	120	91.00	1.11
Public 4-year, non-doctorate-granting, primarily baccalaureate	1,570	87.60	1.17
Public 4-year, doctorate-granting	8,270	86.00	1.17
Private nonprofit 4-year, non-doctorate-granting	1,580	89.10	1.13
Private nonprofit 4-year, doctorate-granting	7,630	84.60	1.20
Private for-profit 4-year	4,200	88.50	1.16
Bureau of Economic Analysis (BEA) region ²			
New England	1,680	86.10	1.14
Mideast	4,400	82.00	1.24
Great Lakes	2,960	85.90	1.16
Plains	2,330	90.20	1.13
Southeast	4,790	87.30	1.15
Southwest	2,770	86.90	1.15
Rocky Mountains	1,250	88.10	1.16
Far West	2,930	84.40	1.20
Outlying Areas	280	90.60	1.17
Graduate full-year enrollment ³			
1–1,615	5,870	86.00	1.17
1,616–4,775	5,860	88.20	1.15
4,776–11,064	5,990	84.00	1.19
11,065–64,922	5,650	86.00	1.17
Graduate student type			
Graduate [excluding doctoral–professional practice]	19,790	85.70	1.18
Doctoral–professional practice	3,590	88.10	1.16
Age as of December 31, 2019			
15–23	2,630	83.70	1.23
24–29	8,420	86.50	1.16
30 or more	12,330	86.20	1.17
Veteran status			
Yes	2,450	89.50	1.11
No	20,920	85.80	1.18

See notes at end of table.

Table 71. Weight adjustment factors for graduate student not located nonresponse adjustment, by model predictor variable: 2019–20—Continued

Model predictor variable	Number of students who were located	Weighted response rate ¹	Average weight adjustment factor (WTA_adj5a)
Collapsed race/ethnicity			
Asian, not Hispanic or Latino	3,270	87.70	1.13
Black, not Hispanic or Latino	3,250	89.40	1.12
Hispanic or Latino of any race	2,580	88.20	1.13
White, not Hispanic or Latino	12,260	89.70	1.11
Other, not Hispanic or Latino or unknown	2,030	58.60	1.74
Gender			
Male	9,870	85.80	1.17
Female	13,510	86.00	1.18
Total amount of unsubsidized Direct Loans disbursed to graduate students ³			
\$0	15,050	83.90	1.20
\$16–\$10,250	2,050	86.90	1.16
\$10,251–\$20,500	2,020	88.40	1.14
\$20,501–\$20,500	2,410	91.10	1.11
\$20,501–\$78,611	1,850	91.80	1.11
Institutional aid status			
Received	7,780	89.10	1.13
Did not receive	10,360	100.0	1.15
Unknown	5,230	62.50	1.29
State aid status			
Received	450	90.80	1.1
Did not receive	19,710	86.70	1.16
Unknown	3,220	78.80	1.28
Federal aid status			
Received	9,580	89.40	1.13
Did not receive	11,680	84.50	1.19
Unknown	2,120	78.30	1.29
Any aid status excluding private loans			
Received	16,300	89.70	1.12
Did not receive	4,110	100.0	1.22
Unknown	2,970	59.40	1.38
CPS—Parents' highest education			
Middle school/junior high	350	88.10	1.15
High school	2,670	89.30	1.12
College or beyond	5,650	91.40	1.1
Other/unknown	390	86.50	1.15
Missing	90	86.50	1.15
Not matched in CPS:20	14,220	82.30	1.21
CPS—Has dependents			
Yes	3,250	89.80	1.12
No	5,900	90.70	1.11
Not matched in CPS:20	14,220	82.30	1.21

See notes at end of table.

Table 71. Weight adjustment factors for graduate student not located nonresponse adjustment, by model predictor variable: 2019–20—Continued

Model predictor variable	Number of students who were located	Weighted response rate ¹	Average weight adjustment factor (WTA_adj5a)
CPS—Marital status			
Single	5,390	89.90	1.12
Married/remarried	3,050	91.40	1.1
Separated	170	92.90	1.07
Divorced/widowed	540	91.00	1.09
Not matched in CPS:20	14,220	82.30	1.21
CPS—Total income ³			
\$-14,016–\$9,174	2,270	91.70	1.09
\$9,175–\$30,742	2,290	92.00	1.09
\$30,743–\$66,838	2,270	87.30	1.15
\$66,839–\$7,867,100	2,300	90.80	1.1
Missing	20	99.60	1
Not matched in CPS:20	14,220	82.30	1.21
CHAID segments			
Graduate students with Black or White, not Hispanic or Latino race/ethnicity who did not receive any aid excluding private loans and who did not match in CPS:20	3,770	83.80	1.19
Graduate students with Black or White, not Hispanic or Latino race/ethnicity who did not receive any aid excluding private loans and whose CPS - Parents' highest education is High School, College or beyond, or Missing	620	92.60	1.07
Graduate students with Black or White, not Hispanic or Latino race/ethnicity who did not receive any aid excluding private loans and who did not match in CPS:20	50	98.30	1.01
Graduate students with Black or White, not Hispanic or Latino race/ethnicity who did not receive any aid excluding private loans and whose CPS - Parents' highest education is high school, college or beyond, or missing	4,150	89.20	1.12
Graduate students with Black or White, not Hispanic or Latino race/ethnicity who did not receive any aid excluding private loans and whose CPS - Parents' highest education is middle school/junior high or other/unknown	580	99.00	1.01
Graduate students in the Great Lakes or Southeast BEA regions with Black or White, not Hispanic or Latino race/ethnicity who received any aid excluding private loans or has unknown aid status	2,680	91.40	1.1
Graduate students in the Rocky Mountains or Outlying BEA regions with Black or White, not Hispanic or Latino race/ethnicity who received any aid excluding private loans or has unknown aid status	2,430	93.30	1.07
Graduate students in the New England or Mideast BEA regions with Black or White, not Hispanic or Latino race/ethnicity who received any aid excluding private loans or has unknown aid status	1,230	94.90	1.05
Graduate students in the Southwest or Far West BEA regions with Black or White, not Hispanic or Latino race/ethnicity who received any aid excluding private loans or has unknown aid status	100	24.40	3.52
Graduate students in the Plains BEA region with Black or White, not Hispanic or Latino race/ethnicity who received any aid excluding private loans or has unknown aid status	1,230	54.80	1.83
Graduate students with Other, not Hispanic or Latino or unknown race/ethnicity who did not match to CPS:20 and are 15–23 years old	180	91.80	1.09
Graduate students with Other, not Hispanic or Latino or unknown race/ethnicity who did not match to CPS:20 and are 15–23 years old	140	69.90	1.4
Graduate students in the New England, Great Lakes, Southeast and Southwest BEA regions with Other, not Hispanic or Latino or unknown race/ethnicity whose CPS - Marital status is single	70	45.40	2.39

See notes at end of table.

Table 71. Weight adjustment factors for graduate student not located nonresponse adjustment, by model predictor variable: 2019–20—Continued

Model predictor variable	Number of students who were located	Weighted response rate ¹	Average weight adjustment factor (WTA_adj5a)
Graduate students in the Plains, Rocky Mountains, Far West or Outlying Areas BEA regions with Other, not Hispanic or Latino or unknown race/ethnicity whose CPS - Marital status is single	230	79.20	1.25
Graduate students in the Mideast BEA region with Other, not Hispanic or Latino or unknown race/ethnicity whose CPS - Marital status is single	70	96.10	1.04
Graduate students with Other, not Hispanic or Latino or unknown race/ethnicity whose CPS - Marital status is married/remarried, separated, or divorced/widowed who did not receive institutional aid or whose institutional aid status is unknown	250	99.90	1
Graduate students with Other, not Hispanic or Latino or unknown race/ethnicity whose CPS - Marital status is married/remarried, separated, or divorced/widowed and who received institutional aid	110	97.20	1.02
Graduate students in the New England, Great Lakes, Southeast, Far West or Outlying BEA regions with Hispanic or Latino of any race or Asian, not Hispanic race/ethnicity	220	89.90	1.1
Graduate students in the Plains and Rocky Mountains BEA regions whose race/ethnicity is Latino or Hispanic of any race or Asian, not Hispanic and whose federal aid status is unknown	1,370	80.10	1.25
Graduate students in the Mideast or Southwest BEA regions with Hispanic or Latino of any race or Asian, not Hispanic race/ethnicity whose federal aid status is unknown	1,930	91.40	1.09
Graduate students with Hispanic or Latino of any race or Asian, not Hispanic race/ethnicity who did not receive any aid excluding private loans	380	95.40	1.05
Graduate students with Hispanic or Latino of any race or Asian, not Hispanic race/ethnicity in the Mid East, Great Lakes, Plains, Rocky Mountains or Far West OBE regions who received any federal aid and who received aid excluding private loans or for whom it is unknown if they received aid excluding private loans	1,300	86.20	1.16
Graduate students with Hispanic or Latino of any race or Asian, not Hispanic race/ethnicity in the Southeast or Southwest BEA regions who received federal aid	290	98.40	1.02

¹ The weighted response rates were calculated using the student base weight (the student sampling 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 Categories were defined by quartiles. Fewer than four categories will display in cases where the variable's values cause quartiles to overlap. Categories of "None" or "Unknown" were excluded from quartile calculations.

NOTE: CHAID = chi-square automatic interaction detection; CPS = Central Processing System. Number of respondents rounded to the nearest 10. Detail may not sum to totals because of rounding.

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

Graduate student refusal nonresponse adjustment (WTA_adj5b). Since the graduate weights were not adjusted for refusal nonresponse, the graduate refusal nonresponse adjustment (WTA_adj5b) was set equal to 1. The refusal nonresponse was accounted for in the adjustment of other nonresponse described below.

Graduate student other nonresponse adjustment (WTA_adj5c). The second, and final, stage of adjustment for graduate student nonresponse was an adjustment for other nonresponse (e.g., contacted, but refused or not interviewed before the end of the data collection period), given that the student was located.

Table 72 shows the final predictor variables used to determine the graduate student other nonresponse weight adjustment (WTA_adj5c), and the average weight adjustment factors resulting from the graduate student model. Summary statistics of the weight adjustment factors follow:

- minimum: 1.00;
- median: 1.14; and
- maximum: 13.49.

For the graduate student other nonresponse weight adjustment (WTA_adj5c), a lower bound of 1 was set on the weight adjustment factors coming out of the nonresponse weight adjustment. No upper bound was set on the weight adjustment factors coming out of the nonresponse weight adjustment.

Table 72. Weight adjustment factors for graduate student other nonresponse adjustment, by model predictor variable: 2019–20

Model predictor variable	Number of respondents	Weighted response rate ¹	Average weight adjustment factor (WTA_adj5c)
Total	19,650	81.00	1.22
Institution stratum			
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	110	96.40	1.03
Public 4-year, non-doctorate-granting, primarily baccalaureate	1,260	80.00	1.21
Public 4-year, doctorate-granting	7,050	82.10	1.21
Private nonprofit 4-year, non-doctorate-granting	1,360	84.40	1.16
Private nonprofit 4-year, doctorate-granting	6,420	79.20	1.24
Private for-profit 4-year	3,450	81.30	1.22
Bureau of Economic Analysis (BEA) region ²			
New England	1,360	74.30	1.26
Mideast	3,660	79.50	1.21
Great Lakes	2,530	82.60	1.20
Plains	1,950	82.60	1.24
Southeast	4,000	81.20	1.24
Southwest	2,340	79.10	1.27
Rocky Mountains	1,050	85.30	1.15
Far West	2,520	83.80	1.15
Outlying Areas	240	86.40	1.16
Graduate full-year enrollment ³			
1–1,615	4,970	82.50	1.16
1,616–4,775	4,890	81.60	1.24
4,776–11,064	5,060	79.30	1.26
11,065–64,922	4,740	81.50	1.21

See notes at end of table.

Table 72. Weight adjustment factors for graduate student other nonresponse adjustment, by model predictor variable: 2019–20—Continued

Model predictor variable	Number of respondents	Weighted response rate ¹	Average weight adjustment factor (WTA_adj5c)
Graduate student type			
Graduate [excluding doctoral–professional practice]	16,520	80.50	1.22
Doctoral–professional practice	3,130	84.80	1.19
Age as of December 31, 2019			
15–23	2,240	84.40	1.14
24–29	7,210	81.50	1.22
30 or more	10,200	79.60	1.23
Veteran status			
Yes	1,990	80.60	1.21
No	17,660	81.00	1.22
Collapsed race/ethnicity			
Asian, not Hispanic or Latino	2,960	87.70	1.13
Black, not Hispanic or Latino	2,790	84.50	1.18
Hispanic or Latino of any race	2,240	87.40	1.15
White, not Hispanic or Latino	10,650	84.80	1.17
Other, not Hispanic or Latino or unknown	1,010	43.10	2.26
Gender			
Male	8,170	78.80	1.24
Female	11,480	82.40	1.20
Total amount of unsubsidized Direct Loans disbursed to graduate students ³			
\$0	12,450	78.80	1.24
\$16–\$10,250	1,720	77.80	1.28
\$10,251–\$20,499	1,750	88.20	1.13
\$20,500	2,120	86.40	1.16
\$20,501–\$78,611	1,610	87.80	1.17
Institutional aid status			
Received	6,820	84.50	1.18
Did not receive	10,360	100.0	1.25
Unknown	2,460	37.20	1.18
State aid status			
Received	400	88.50	1.17
Did not receive	16,580	80.90	1.23
Unknown	2,670	80.20	1.17
Federal aid status			
Received	8,270	85.00	1.18
Did not receive	9,670	78.60	1.26
Unknown	1,720	77.00	1.2
Any aid status excluding private loans			
Received	14,330	86.20	1.16
Did not receive	4,110	100.0	1.4
Unknown	1,210	31.20	1.27

See notes at end of table.

Table 72. Weight adjustment factors for graduate student other nonresponse adjustment, by model predictor variable: 2019–20—Continued

Model predictor variable	Number of respondents	Weighted response rate ¹	Average weight adjustment factor (WTA_adj5c)
CPS—Parents' highest education			
Middle school/junior high	300	86.70	1.15
High school	2,290	83.80	1.19
College or beyond	4,950	86.70	1.15
Other/unknown	330	80.40	1.27
Missing	80	92.80	1.1
Not matched in CPS:20	11,710	77.30	1.25
CPS—Has dependents			
Yes	2,770	83.80	1.2
No	5,170	86.60	1.15
Not matched in CPS:20	11,710	77.30	1.25
CPS—Marital status			
Single	4,710	86.20	1.15
Married/remarried	2,650	86.30	1.16
Separated	130	75.80	1.32
Divorced/widowed	450	78.00	1.31
Not matched in CPS:20	11,710	77.30	1.25
CPS—Total income ³			
\$-14,016–\$9,174	1,970	87.70	1.13
\$9,175–\$30,742	2,010	83.50	1.19
\$30,743–\$66,838	2,000	87.70	1.14
\$66,839–\$7,867,100	1,950	83.90	1.19
Missing	20	86.60	1.2
Not matched in CPS:20	11,710	77.30	1.25
CHAID segments			
Graduate students with White, not Hispanic or Latino race/ethnicity whose state aid status is unknown	1,360	97.50	1.02
Graduate students with Black, not Hispanic or Latino race/ethnicity whose state aid status is unknown	410	90.60	1.1
Graduate students with Black or White, not Hispanic or Latino race/ethnicity who did not receive state aid or any other aid excluding private loans	2,850	75.00	1.33
Graduate students with Black or White, not Hispanic or Latino race/ethnicity who did not receive state aid and received aid excluding private loans or whose aid status excluding private loans is unknown	8,550	87.30	1.15
Graduate students with Black or White, not Hispanic or Latino race/ethnicity who received state aid and whose CPS Total Income is between \$-14,016 and \$9,174	70	98.50	1.02
Graduate students with Black or White, not Hispanic or Latino race/ethnicity who received state aid and whose CPS Total Income is between \$9,175 and \$7,867,100	120	93.20	1.07
Graduate students with Black or White, not Hispanic or Latino race/ethnicity who received state aid and whose CPS Total Income is missing or who did not match to CPS:20	90	70.40	1.5
Graduate students in the Mideast, Southeast, Rocky Mountains, or Outlying BEA regions with Other, not Hispanic or Latino or unknown race/ethnicity who did not receive any aid excluding private loans	100	21.30	5.1
Graduate students in the New England, Great Lakes, Southwest or Far West BEA regions with Other, not Hispanic or Latino or unknown race/ethnicity who did not receive any aid excluding private loans	110	39.00	2.19

See notes at end of table.

Table 72. Weight adjustment factors for graduate student other nonresponse adjustment, by model predictor variable: 2019–20—Continued

Model predictor variable	Number of respondents	Weighted response rate ¹	Average weight adjustment factor (WTA_adj5c)
Graduate students in the Plains BEA region with Other, not Hispanic or Latino or unknown race/ethnicity who did not receive any aid excluding private loans	30	18.40	5.43
Graduate students in the Great Lakes, Plain, Southeast, Rocky Mountains or Far West BEA regions with Other, not Hispanic or Latino or unknown race/ethnicity who received aid excluding private loans or whose aid status excluding private loans is unknown	480	68.30	1.44
Graduate students in the New England, Mideast, Southwest, and Outlying BEA regions with Other, not Hispanic or Latino or unknown race/ethnicity who received aid excluding private loans or whose aid status excluding private loans is unknown	280	45.20	2.28
Graduate students with Hispanic or Latino of any race or Asian, not Hispanic race/ethnicity whose state aid status is unknown and who attended a school with the graduate full-year enrollment between 1 and 1,615	200	98.50	1.01
Graduate students with Hispanic or Latino of any race or Asian, not Hispanic race/ethnicity whose state aid status is unknown and who attended a school with the graduate full-year enrollment between 1,616 and 4,775	350	95.00	1.05
Graduate students with Hispanic or Latino of any race or Asian, not Hispanic race/ethnicity whose state aid status is unknown and who attended a school with the graduate full-year enrollment between 4,776 and 64,922	180	99.90	1
Graduate students with Hispanic or Latino of any race or Asian, not Hispanic race/ethnicity whose state aid status is known and who did not receive any aid excluding private loans	1,150	79.60	1.26
Graduate students with Hispanic or Latino of any race or Asian, not Hispanic race/ethnicity whose state aid status is known and who received aid excluding private loans or whose aid status excluding private loans is unknown	3,330	89.60	1.12

¹ The weighted response rates were calculated using the student base weight adjusted for not located nonresponse.

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 Categories were defined by quartiles. Fewer than four categories will display in cases where the variable's values cause quartiles to overlap. Categories of "None" or "Unknown" were excluded from quartile calculations.

NOTE: CHAID = chi-square automatic interaction detection; CPS = Central Processing System. Number of respondents rounded to the nearest 10. Detail may not sum to totals because of rounding.

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

7.1.3.6.3 Undergraduate student study nonresponse adjustment (WTB_adj5)

The weighting team adjusted the undergraduate study weight for nonresponse in only one stage. The initial undergraduate study nonresponse model included variables that were defined by public 2-year, public 4-year, and other institutions within each state for representative sectors. 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.⁵⁴ The institution strata used in the final model included the following:

- 15 states with variables defined by public 2-year, public 4-year, and other institutions within state;
- 16 states with variables defined by public 2-year and public 4-year combined and other institutions within state; and
- 21 states with variables defined by state overall.

Table 73 and table 74 show, respectively, the final predictor variables used to determine the student study nonresponse weight adjustments and the average weight adjustment factors resulting from the undergraduate student study nonresponse 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 student study nonresponse adjustment were created separately within each state and institution stratum listed in table 73 and were too numerous and varied to include in table 74. 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 study nonresponse adjustment. For the undergraduate student study nonresponse weight adjustment, a lower bound of 1 and an upper bound of 5,000 were set on the weight adjustment factors coming out of the undergraduate student study nonresponse weight adjustment. Summary statistics of the undergraduate student study nonresponse weight adjustment factors for undergraduate students follow:

- minimum: 1.00;
- median: 1.04; and
- maximum: 57.05.

Table 73. Weight adjustment factors for undergraduate student study nonresponse adjustment, by state and institution stratum: 2019–20

State and institution stratum	Number of respondents	Weighted response rate ¹	Average weight adjustment factor (WTB_adj5)
Total	276,350	83.70	1.22
Alabama			
Public 2-year and Public 4-year	4,360	85.40	1.27
Other	1,120	92.70	1.12

See notes at end of table.

⁵⁴ Final representativeness at the public 2-year, public 4-year, and state levels is discussed in section 1.3.1.

Table 73. Weight adjustment factors for undergraduate student study nonresponse adjustment, by state and institution stratum: 2019–20—Continued

State and institution stratum	Number of respondents	Weighted response rate ¹	Average weight adjustment factor (WTB_adj5)
Alaska			
All	970	87.10	1.32
Arizona			
Public 2-year and Public 4-year	3,670	86.70	1.18
Other	1,690	76.70	1.23
Arkansas			
All	3,880	80.80	1.23
California			
Public 2-year	9,490	66.50	1.50
Public 4-year	6,010	87.10	1.14
Other	2,530	83.80	1.26
Colorado			
Public 2-year and Public 4-year	3,700	88.50	1.10
Other	950	74.10	1.33
Connecticut			
Public 2-year and Public 4-year	3,260	87.10	1.24
Other	2,000	91.20	1.17
Delaware			
All	1,130	34.30	2.13
District of Columbia			
All	1,630	73.50	1.79
Florida			
Public 2-year and Public 4-year	7,420	88.40	1.14
Other	2,780	84.00	1.24
Georgia			
Public 2-year	1,680	47.40	2.08
Public 4-year	3,800	79.40	1.26
Other	1,870	91.70	1.11
Hawaii			
Public 2-year	1,610	98.10	1.02
Public 4-year	1,460	99.10	1.01
Other	1,010	96.70	1.03
Idaho			
Public 2-year	140	31.80	3.21
Public 4-year	860	74.90	1.40
Other	1,000	18.10	2.35
Illinois			
Public 2-year	4,730	82.30	1.21
Public 4-year	1,940	95.60	1.05
Other	1,490	89.20	1.17
Indiana			
All	6,450	96.30	1.05
Iowa			
Public 2-year	1,420	95.60	1.05
Public 4-year	1,030	99.10	1.01
Other	1,090	92.40	1.07
Kansas			
Public 2-year and Public 4-year	2,060	70.70	1.69
Other	830	94.90	1.06
Kentucky			
All	4,320	92.30	1.09

See notes at end of table.

Table 73. Weight adjustment factors for undergraduate student study nonresponse adjustment, by state and institution stratum: 2019–20—Continued

State and institution stratum	Number of respondents	Weighted response rate ¹	Average weight adjustment factor (WTB_adj5)
Louisiana			
Public 2-year and Public 4-year	3,610	78.60	1.33
Other	1,150	89.90	1.12
Maine			
Public 2-year and Public 4-year	2,280	92.60	1.08
Other	1,970	88.10	1.15
Maryland			
Public 2-year	1,710	61.60	1.58
Public 4-year	1,890	95.60	1.04
Other	970	82.40	1.19
Massachusetts			
Public 2-year	1,710	83.10	1.21
Public 4-year	1,940	91.30	1.10
Other	2,340	83.20	1.20
Michigan			
All	6,720	91.10	1.12
Minnesota			
Public 2-year	4,970	96.80	1.03
Public 4-year	2,250	99.30	1.01
Other	1,400	87.30	1.12
Mississippi			
All	3,140	89.10	1.15
Missouri			
All	4,260	82.50	1.29
Montana			
All	1,230	31.00	3.14
Nebraska			
All	3,260	93.60	1.08
Nevada			
All	2,540	92.20	1.14
New Hampshire			
All	3,310	95.50	1.05
New Jersey			
Public 2-year and Public 4-year	4,800	79.50	1.27
Other	2,220	80.10	1.30
New Mexico			
Public 2-year and Public 4-year	3,390	96.60	1.05
Other	720	97.80	1.02
New York			
Public 2-year	3,400	64.30	1.54
Public 4-year	5,380	64.80	1.55
Other	2,290	95.00	1.07
North Carolina			
All	11,110	89.80	1.14
North Dakota			
Public 2-year	390	100.0	1.00
Public 4-year	700	91.70	1.08
Other	350	45.80	2.13

See notes at end of table.

Table 73. Weight adjustment factors for undergraduate student study nonresponse adjustment, by state and institution stratum: 2019–20—Continued

State and institution stratum	Number of respondents	Weighted response rate ¹	Average weight adjustment factor (WTB_adj5)
Ohio			
Public 2-year	3,890	98.00	1.03
Public 4-year	4,590	81.60	1.24
Other	1,500	95.20	1.06
Oklahoma			
All	5,580	83.00	1.23
Oregon			
All	4,110	91.20	1.09
Pennsylvania			
Public 2-year and Public 4-year	6,120	77.40	1.31
Other	1,770	81.80	1.23
Puerto Rico			
Public 2-year and Public 4-year	1,390	77.10	1.29
Other	2,580	87.50	1.17
Rhode Island			
All	1,490	48.90	2.01
South Carolina			
Public 2-year and Public 4-year	2,890	87.70	1.21
Other	1,590	88.20	1.15
South Dakota			
All	2,190	77.70	1.29
Tennessee			
Public 2-year and Public 4-year	5,920	84.00	1.22
Other	1,330	87.60	1.20
Texas			
Public 2-year	5,490	78.10	1.26
Public 4-year	7,970	94.50	1.07
Other	1,590	85.60	1.27
Utah			
All	4,890	94.60	1.08
Vermont			
Public 2-year	1,280	100.0	1.00
Public 4-year	760	100.0	1.00
Other	1,380	98.50	1.02
Virginia			
Public 2-year and Public 4-year	4,810	90.70	1.15
Other	800	89.70	1.17
Washington			
All	6,070	76.40	1.32
West Virginia			
Public 2-year and Public 4-year	3,030	94.80	1.09
Other	1,760	93.00	1.08
Wisconsin			
All	4,800	95.10	1.06
Wyoming			
Public 2-year	690	54.80	1.86
Public 4-year	1,290	100.0	1.00
Other	80	100.0	1.00

¹ The weighted response rates were calculated using the student base weight (the student sampling weight adjusted for multiplicity and unknown eligibility).

NOTE: Number of respondents rounded to the nearest 10. Detail may not sum to totals because of rounding.

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

Table 74. Weight adjustment factors for undergraduate student study nonresponse adjustment, by model predictor variable: 2019–20

Model predictor variable ¹	Number of respondents	Weighted response rate ²	Average weight adjustment factor (WTB_adj5)
Total	276,350	83.70	1.22
Age as of December 31, 2019			
15–23	183,700	85.20	1.20
24–29	41,690	80.90	1.27
30 or more	50,960	80.80	1.27
Gender			
Male	116,430	83.50	1.22
Female	159,920	83.80	1.22
Veteran status			
Yes	9,840	82.50	1.24
No	266,520	83.70	1.22
Sampled as first-time beginning student			
Yes	54,910	90	1.14
No	221,440	82.6	1.24
Total amount of federal Pell Grants awarded to undergraduate students ³			
\$0	175,070	84.4	1.21
\$1–\$3,097	31,510	81.3	1.27
\$3,098–\$6,194	38,170	82.6	1.25
\$6,195 or more	31,600	83	1.25
Institutional aid status			
Received	79,450	93.4	1.08
Did not receive	196,910	80.2	1.28
State aid status			
Received	55,310	92.2	1.1
Did not receive	221,040	81.70	1.26
Federal aid status			
Received	151,440	83.50	1.24
Did not receive	124,910	83.90	1.21
Total amount of Direct Loans disbursed to undergraduate students ⁴			
\$0	182,830	83.10	1.23
\$1–\$4,500	23,340	84.80	1.23
\$4,501–\$6,250	24,490	85.90	1.21
\$6,251–\$7,500	23,730	85.8	1.21
\$7,501 or more	21,960	83.6	1.24
Institution total undergraduate enrollment ⁴			
1–3,669	68,200	82.2	1.24
3,670–9,668	69,570	83.6	1.23
9,669–20,011	66,610	81.40	1.23
20,012 or more	71,980	85.80	1.18

¹ Not all model predictor variables were retained for every institution stratum due to removal of variables to resolve convergence issues. CHAID segments were created separately within each state and institution stratum listed in Table 73 and were too numerous and varied to include in Table 74.

² The weighted response rates were calculated using the student base weight (the student sampling weight adjusted for multiplicity and unknown eligibility).

³ For Total amount of federal Pell Grants awarded to undergraduate students, the second and third categories were defined by the median of Pell amounts greater than \$0 and less than \$6,195. Categories were defined by quartiles. Fewer than four categories will display in cases where the variable's values cause quartiles to overlap. Categories of "None" or "Unknown" were excluded from quartile calculations.

⁴ Categories were defined by quartiles. Fewer than four categories will display in cases where the variable's values cause quartiles to overlap. Categories of "None" or "Unknown" were excluded from quartile calculations.

NOTE: Number of respondents rounded to the nearest 10. Detail may not sum to totals because of rounding.

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

7.1.3.7 *Student poststratification adjustments (WTA_adj6 and WTB_adj6)*

To ensure that the weighted student sample adequately represents the student population, the nonresponse-adjusted student weights were further adjusted so that weighted sums will equal known population control totals for key characteristics. This adjustment also helped increase the precision of the estimates for these key characteristics and any related characteristics.

For undergraduate student study respondents, student enrollment and federal financial aid control totals included state-level totals such 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 “all other sectors” within state if the control totals were available by those three sectors. State aid control totals included state-level totals. Control totals that were unavailable at the state level but were available at the national level were included in the model.

For undergraduate student survey and graduate respondents, control totals included enrollment and federal financial aid totals by control and level of institution.

Control totals for undergraduate students were established when available, by institution control and level for undergraduate survey respondents and by state and by public 2-year, public 4-year, and “all other sectors” within state for undergraduate study respondents for the following:

- full-year undergraduate student enrollment, by race/ethnicity;
- fall undergraduate student enrollment;
- number of Pell Grant undergraduate recipients;
- total amount of Pell Grants awarded to undergraduate students;
- total amount of Parent PLUS Loans disbursed to undergraduate students;
- number of Direct Loan undergraduate student recipients; and
- total amount of Direct Loans disbursed to undergraduate students.

Additional state-level control totals for study undergraduate respondents were established when available by state for the following:

- number of state aid undergraduate recipients, by type of aid (merit, need, non-need, other, and veteran); and
- total amount of state aid awarded to undergraduate students, by type of aid (i.e., merit, need, non-need other, and veteran).

Control totals for graduate students were established for the following:

- full-year graduate student enrollment, by race/ethnicity and by control and level of institution;
- fall graduate student enrollment, 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; and
- PLUS Loan amounts disbursed to graduate students, by control and level of institution.

Direct Loan, Pell Grant, and PLUS control totals were obtained from the U.S. Department of Education's FSA Office. The Direct Loan program is the largest single student loan program in terms of the number of students affected and the dollars involved. Therefore, accurate control total data on Direct Loans by loan type (subsidized or unsubsidized), control and level of institution, and student level (undergraduate or graduate) was crucial for poststratification. State aid control totals were obtained from the National Association of State Student Grant and Aid Programs.

Student enrollment control totals were determined using IPEDS data, which was 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:

1. EF2019A: 2019 Fall Enrollment—includes data on race/ethnicity, attendance status, and level of student for fall 2019;
2. EFFY2020: 2019–20 12-month Enrollment—includes 12-month unduplicated head count for 2019–20;
3. HD2019: 2019–20 Institutional Characteristics Header—includes directory information for 2019–20;
4. IC2019: 2019–20 Institutional Characteristics—includes data on educational offerings, organization, admissions, services, and athletic associations for 2019–20; and
5. IC2019_PY: 2019–20 Institutional Characteristics—includes data on student charges by program (vocational programs) for 2019–20.

The HD2019, IC2019, and IC2019_PY files were used in determining which institutions were in the NPSAS population of institutions and were also used to create the control and level of institution variable. The EF2019A and EFFY2020 files were used to determine the undergraduate and graduate enrollment totals for fall and the full year, respectively. Missing enrollment totals were imputed for about 80 institutions. 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:

$$\text{NPSAS control total} = (\text{IPEDS enrollment total}) * (1 / \text{mean student multiplicity count}),$$

where NPSAS control total, IPEDS enrollment total, and mean student multiplicity count (from NPSAS:20 data) are 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 counts of dual-enrolled students included on the student enrollment lists were used to estimate dual enrollment in the population and adjust the control totals accordingly. 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, and within state, the following formula was used (with Black students as an example):

$$\text{NPSAS Black control total} = \text{NPSAS full-year control total} * \text{Proportion Black}$$

where:

- Proportion Black = IPEDS Black full-year enrollment total/IPEDS full-year enrollment total;
- NPSAS Black control total, NPSAS full-year control total, Proportion Black, IPEDS Black full-year enrollment total, and IPEDS full-year enrollment total are by undergraduate and graduate student level within the given control and level of institution or within the state; and
- all control totals used in the calculations are adjusted for multiplicity and dual enrollment.

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 weights.

As with the nonresponse weight adjustments, different bounds to trim input weights were used during the poststratification weight adjustments, depending on whether the input weight was classified as high extreme,⁵⁰ low extreme,⁵¹ or nonextreme, to accomplish the poststratification adjustment, truncation, and smoothing in one step. In addition, the data were inspected for extreme weights by checking for outliers by key domains.

Undergraduate student survey poststratification adjustment (WTA_adj6).

Before the undergraduate student poststratification survey weight adjustment, bounds were set to trim the extreme weights going into the adjustment. To trim small weights, lower bounds were set by student sampling strata within each category of institution control and level; the lower bounds ranged from 1 for all students in public less-than-2-year; public 4-year, non-doctorate-granting, primarily baccalaureate; and private nonprofit institutions to 700 for students sampled as graduate students in private not-for-profit 4-year, doctorate-granting institutions. To trim large weights, an upper bound was set equal to the median input weight (the nonresponse-adjusted student weight) plus three times the interquartile range by student sampling strata within control and level of institution for all other students.

For the undergraduate student poststratification survey weight adjustment, lower bounds and upper bounds on the weight adjustment factors coming out of the poststratification weight adjustment were used to control the magnitude of the poststratified weights. The adjustment lower bound was set to 0.1 for students in public 2-year institutions and was unbounded (equal to 0) for all other students. The adjustment upper bound was set to 1.7 for one student in a private nonprofit 4-year, doctorate-granting institution and was unbounded for all other students.

Table 75 shows the control totals associated with the undergraduate poststratification weight adjustment and the average weight adjustment factors resulting from the undergraduate survey student model. Summary statistics of the undergraduate poststratification student survey weight adjustment factors follow:

- minimum: 0.0015;
- median: 1.03; and
- maximum: 18.59.

Table 75. Weight adjustment factors for undergraduate survey student poststratification, by poststratification category: 2019–20

Poststratification category	Control total	Average weight adjustment factor (WTA_adj6)
Total, full-year undergraduate student enrollment	17,103,820	1.17
Full-year undergraduate student enrollment, by race, within control and level of institution ¹		
Black, not Hispanic or Latino		
Public less-than-2-year	7,071	0.98
Public 2-year	743,767	1.11
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	166,601	1.02
Public 4-year, non-doctorate-granting, primarily baccalaureate	102,821	1.03
Public 4-year, doctorate-granting	517,064	0.97
Private nonprofit 2-year or less	16,606	1.77
Private nonprofit 4-year, non-doctorate-granting	131,605	1.12
Private nonprofit 4-year, doctorate-granting	175,782	1.04
Private for-profit less-than-2-year	71,654	1.41
Private for-profit 2-year	54,849	2.59
Private for-profit 4-year	198,449	1.63
Hispanic or Latino of any race		
Public less-than-2-year	10,177	2.27
Public 2-year	1,397,708	1.16
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	416,408	1.16
Public 4-year, non-doctorate-granting, primarily baccalaureate	166,882	1.09
Public 4-year, doctorate-granting	831,894	0.98
Private nonprofit 2-year or less	8,750	0.74
Private nonprofit 4-year, non-doctorate-granting	142,453	1.1
Private nonprofit 4-year, doctorate-granting	259,804	0.98
Private for-profit less-than-2-year	71,118	1.13
Private for-profit 2-year	62,561	1.57
Private for-profit 4-year	143,036	1.12
White, not Hispanic or Latino		
Public less-than-2-year	38,396	1.84
Public 2-year	2,544,287	1.2
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	589,052	1.12
Public 4-year, non-doctorate-granting, primarily baccalaureate	462,943	1.01
Public 4-year, doctorate-granting	2,515,722	1
Private nonprofit 2-year or less	22,049	1.09
Private nonprofit 4-year, non-doctorate-granting	586,168	0.98
Private nonprofit 4-year, doctorate-granting	922,432	1.09
Private for-profit less-than-2-year	86,350	1.41
Private for-profit 2-year	74,809	1.69
Private for-profit 4-year	297,090	1.49
Other, not Hispanic or Latino		
Public less-than-2-year	6,558	1.72
Public 2-year	938,613	1
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	261,368	1.31
Public 4-year, non-doctorate-granting, primarily baccalaureate	143,585	0.96
Public 4-year, doctorate-granting	991,169	1.11
Private nonprofit 2-year or less	12,179	2.16
Private nonprofit 4-year, non-doctorate-granting	200,180	1.34
Private nonprofit 4-year, doctorate-granting	429,542	1.36
Private for-profit less-than-2-year	39,343	1.38
Private for-profit 2-year	38,302	1.2
Private for-profit 4-year	206,623	2.61

See notes at end of table.

Table 75. Weight adjustment factors for undergraduate survey student poststratification, by poststratification category: 2019–20—Continued

Poststratification category	Control total	Average weight adjustment factor (WTA_adj6)
Fall undergraduate student enrollment, by control and level of institution ¹		
Public less-than-2-year	44,870	1.61
Public 2-year	3,826,076	0.97
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	1,003,466	1
Public 4-year, non-doctorate-granting, primarily baccalaureate	729,652	0.95
Public 4-year, doctorate-granting	4,282,541	0.95
Private nonprofit 2-year or less	43,330	1.15
Private nonprofit 4-year, non-doctorate-granting	883,166	1
Private nonprofit 4-year, doctorate-granting	1,494,867	1.02
Private for-profit less-than-2-year	159,296	1.24
Private for-profit 2-year	146,591	1.43
Private for-profit 4-year	512,586	1.2
Number of Pell Grant undergraduate recipients, by control and level of institution		
Public less-than-2-year	18,123	1.65
Public 2-year	2,040,225	1.14
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	503,600	1.09
Public 4-year, non-doctorate-granting, primarily baccalaureate	381,584	1.11
Public 4-year, doctorate-granting	1,917,496	1.2
Private nonprofit 2-year or less	49,495	2.08
Private nonprofit 4-year, non-doctorate-granting	436,460	1.27
Private nonprofit 4-year, doctorate-granting	609,663	1.26
Private for-profit less-than-2-year	211,441	1.56
Private for-profit 2-year	192,305	2.35
Private for-profit 4-year	519,400	1.95
Total amount of Pell Grants awarded to undergraduate students, by control and level of institution		
Public less-than-2-year	\$73,353,222	1.65
Public 2-year	7,363,244,519	1.14
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	1,858,282,536	1.09
Public 4-year, non-doctorate-granting, primarily baccalaureate	1,709,182,215	1.11
Public 4-year, doctorate-granting	8,781,170,371	1.2
Private nonprofit 2-year or less	234,676,098	2.08
Private nonprofit 4-year, non-doctorate-granting	1,942,990,357	1.27
Private nonprofit 4-year, doctorate-granting	2,591,366,471	1.26
Private for-profit less-than-2-year	814,253,077	1.56
Private for-profit 2-year	805,177,838	2.35
Private for-profit 4-year	2,170,328,036	1.95
Total amount of PLUS Loans disbursed to undergraduate student recipients, by control and level of institution		
Public less-than-2-year	\$1,980,971	18.59
Public 2-year	78,251,823	0.38
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	73,006,738	0.57
Public 4-year, non-doctorate-granting, primarily baccalaureate	577,103,927	1.05
Public 4-year, doctorate-granting	5,637,677,253	1.15
Private nonprofit 2-year or less	25,486,139	2.06
Private nonprofit 4-year, non-doctorate-granting	1,634,442,063	0.99
Private nonprofit 4-year, doctorate-granting	3,479,855,502	1.25
Private for-profit less-than-2-year	124,083,972	1.34
Private for-profit 2-year	186,456,193	2.22
Private for-profit 4-year	369,709,970	1.4

See notes at end of table.

Table 75. Weight adjustment factors for undergraduate survey student poststratification, by poststratification category: 2019–20—Continued

Poststratification category	Control total	Average weight adjustment factor (WTA_adj6)
Number of subsidized Direct Loan undergraduate student recipients, by control and level of institution		
Public less-than-2-year	7,068	1.51
Public 2-year	552,792	0.95
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	155,940	0.91
Public 4-year, non-doctorate-granting, primarily baccalaureate	309,722	1.11
Public 4-year, doctorate-granting	1,815,599	1.21
Private nonprofit 2-year or less	41,629	2.17
Private nonprofit 4-year, non-doctorate-granting	481,938	1.18
Private nonprofit 4-year, doctorate-granting	777,017	1.25
Private for-profit less-than-2-year	152,602	1.29
Private for-profit 2-year	159,574	2.24
Private for-profit 4-year	486,846	1.83
Total amount of subsidized Direct Loans disbursed to undergraduate students, by control and level of institution		
Public less-than-2-year	\$21,474,392	1.51
Public 2-year	1,598,762,396	0.95
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	478,073,743	0.91
Public 4-year, non-doctorate-granting, primarily baccalaureate	1,203,551,704	1.11
Public 4-year, doctorate-granting	7,329,025,707	1.21
Private nonprofit 2-year or less	151,866,778	2.17
Private nonprofit 4-year, non-doctorate-granting	1,902,111,026	1.18
Private nonprofit 4-year, doctorate-granting	3,178,257,704	1.25
Private for-profit less-than-2-year	436,188,416	1.29
Private for-profit 2-year	555,926,162	2.24
Private for-profit 4-year	1,871,349,179	1.83
Number of unsubsidized Direct Loan undergraduate student recipients, by control and level of institution		
Public less-than-2-year	8,288	1.16
Public 2-year	501,525	0.93
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	130,758	0.89
Public 4-year, non-doctorate-granting, primarily baccalaureate	309,765	1.14
Public 4-year, doctorate-granting	1,838,980	1.19
Private nonprofit 2-year or less	42,769	2.04
Private nonprofit 4-year, non-doctorate-granting	502,685	1.17
Private nonprofit 4-year, doctorate-granting	803,104	1.3
Private for-profit less-than-2-year	164,155	1.31
Private for-profit 2-year	165,251	2.23
Private for-profit 4-year	496,438	1.81

See notes at end of table.

Table 75. Weight adjustment factors for undergraduate survey student poststratification, by poststratification category: 2019–20—Continued

Poststratification category	Control total	Average weight adjustment factor (WTA_adj6)
Total amount of unsubsidized Direct Loans disbursed to undergraduate students, by control and level of institution		
Public less-than-2-year	\$40,065,968	1.16
Public 2-year	1,834,596,926	0.93
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	490,525,422	0.89
Public 4-year, non-doctorate-granting, primarily baccalaureate	1,226,308,998	1.14
Public 4-year, doctorate-granting	7,248,156,905	1.19
Private nonprofit 2-year or less	182,303,975	2.04
Private nonprofit 4-year, non-doctorate-granting	1,943,115,489	1.17
Private nonprofit 4-year, doctorate-granting	3,141,937,441	1.3
Private for-profit less-than-2-year	664,533,881	1.31
Private for-profit 2-year	780,005,343	2.23
Private for-profit 4-year	2,568,250,120	1.81

¹ Enrollment control totals are the sum of enrollment (full year and fall) across institutions based on 2019–20 Integrated Postsecondary Education Data System enrollment data adjusted for student multiplicity and dual enrollment.

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

Graduate student poststratification adjustment (WTA_adj6). 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 institution control and level; a lower bound of 2 and 4 was set for students in public 4-year, non-doctorate-granting, primarily subbaccalaureate institutions and private for-profit 4-year institutions, respectively, and was set equal to the median input weight minus three times the interquartile range by control and level of institution for all other students. To trim large weights, the following upper bounds were set:

- 325 for students in public 4-year, non-doctorate-granting, primarily subbaccalaureate institutions;
- 500 for students in public 4-year, doctorate-granting institutions and private not-for-profit 4-year, doctorate-granting institutions; and
- the median incoming weight plus three times the interquartile range by control and level of institution for all other students.

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. For students in public 2-year institutions, the lower bounds were set to 0.03 and 0.4 for students with input weights less than 10 and between 10 and 100, respectively. For students in private for-profit 4-year

institutions, the lower bound was set to 0.05 for students with input weights less than 5. For all graduate students, the upper bound on the poststratification adjustment factor was unbounded.

Table 76 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: 0.039;
- median: 1.36; and
- maximum: 15.02.

Table 76. Weight adjustment factors for graduate student poststratification, by poststratification category: 2019–20

Poststratification category	Control total	Average weight adjustment factor (WTA_adj6)
Total, full-year graduate student enrollment¹	3,636,340	1.53
Full-year graduate student enrollment, by race, within control and level of institution ¹		
Black		
Public		
4-year, non-doctorate-granting	12,407	1.03
4-year, doctorate-granting	145,270	1.42
Private nonprofit		
4-year, non-doctorate-granting	21,504	1.23
4-year, doctorate-granting	147,066	1.49
Private for-profit 4-year	90,124	2.05
Hispanic or Latino		
Public		
4-year, non-doctorate-granting	13,430	0.88
4-year, doctorate-granting	154,358	1.48
Private nonprofit		
4-year, non-doctorate-granting	20,246	1.21
4-year, doctorate-granting	140,552	1.40
Private for-profit 4-year	36,990	1.09
White		
Public		
4-year, non-doctorate-granting	77,340	1.05
4-year, doctorate-granting	821,887	1.48
Private nonprofit		
4-year, non-doctorate-granting	109,040	1.26
4-year, doctorate-granting	652,329	1.49
Private for-profit 4-year	133,347	1.29

See notes at end of table.

Table 76. Weight adjustment factors for graduate student poststratification, by poststratification category: 2019–20—Continued

Poststratification category	Control total	Average weight adjustment factor (WTA_adj6)
Other race		
Public		
4-year, non-doctorate-granting	22,938	1.24
4-year, doctorate-granting	447,678	1.71
Private nonprofit		
4-year, non-doctorate-granting	40,345	2.33
4-year, doctorate-granting	465,272	1.94
Private for-profit 4-year	84,217	2.29
Fall undergraduate student enrollment, by control and level of institution ¹		
Public		
4-year, non-doctorate-granting	86,859	0.95
4-year, doctorate-granting	1,297,709	1.42
Private nonprofit		
4-year, non-doctorate-granting	134,749	1.23
4-year, doctorate-granting	1,128,261	1.50
Private for-profit 4-year	217,941	1.22
Number of Unsubsidized Direct Loans graduate student recipients, by control and level of institution		
Public		
4-year, non-doctorate-granting	43,496	1.17
4-year, doctorate-granting	531,497	1.59
Private nonprofit		
4-year, non-doctorate-granting	81,448	1.20
4-year, doctorate-granting	574,734	1.80
Private for-profit 4-year	189,909	1.93
Total amount of Unsubsidized Direct Loans disbursed to graduate students, by control and level of institution		
Public		
4-year, non-doctorate-granting	\$573,680,294	1.17
4-year, doctorate-granting	9,936,536,122	1.59
Private nonprofit		
4-year, non-doctorate-granting	1,085,977,773	1.20
4-year, doctorate-granting	11,918,268,538	1.80
Private for-profit 4-year	3,016,374,325	1.93
PLUS Loan amounts disbursed to graduate students, by control and level of institution		
Public		
4-year, non-doctorate-granting	\$24,630,742	1.22
4-year, doctorate-granting	2,652,840,341	1.41
Private nonprofit		
4-year, non-doctorate-granting	187,000,251	1.17
4-year, doctorate-granting	6,717,115,639	1.86
Private for-profit 4-year	578,063,762	1.15

¹ Enrollment control totals are the sum of enrollment (full year and fall) across institutions based on 2019–20 Integrated Postsecondary Education Data System enrollment data adjusted for student multiplicity and dual enrollment.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

Undergraduate student study poststratification adjustment (WTB_adj6). The initial undergraduate student study poststratification model included variables that were defined by three institution strata (public 2-year, public 4-year, and “all other sectors”) within each state for a total of 152 institution strata.⁵⁵ 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 a state and which institution strata to retain in the model.⁵⁶ The institution strata used in the final undergraduate student study poststratification model included the following:

- 36 potentially state-representative states with variables defined within three institution strata: public 2-year, public 4-year, and “all other sectors” within the state;
- 2 public 2-year strata within nonrepresentative states, which were determined to be potentially representative at the public 2-year level;
- 8 public 4-year strata within nonrepresentative states, which were determined to be potentially representative at the public 4-year level; and
- 34 nonrepresentative institution strata that were collapsed into an “all other sectors” stratum.

Bounds for the undergraduate student study poststratification were determined at the institution strata level. Before the undergraduate student study 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 2,851 low extreme weights that were trimmed ranged from 1 to 30, with an average lower bound of 4.36. The upper bounds for the 2,970 high extreme weights that were trimmed ranged from 9.52 to 1,800, with an average upper bound of 141.08.

For the undergraduate student study 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 undergraduate student study poststratification weight adjustment in order to control the magnitude of the poststratified undergraduate student study weights. The lower bounds ranged from 0.01 to 0.8, with an average lower bound of 0.087. The upper bounds ranged from

⁵⁵ There were no public 2-year institutions in Alaska, Delaware, the District of Columbia, or Nevada.

⁵⁶ Final representativeness at the public 2-year, public 4-year, and state levels is discussed in section 1.3.1.

4.8 to 400, with an average upper bound of 71.9. For institution control and level categories for which no upper bound was specified, the upper bound on the undergraduate student study poststratification adjustment was unbounded.

Table 77 shows state-level full-year enrollment control totals and the average undergraduate student study poststratification weight adjustment factors, by the institution strata used in the final model for undergraduate study students.

Table 77. Weight adjustment factors for undergraduate student study poststratification, by state and institution stratum: 2019–20

State and institution stratum ¹	Full-year enrollment control total ²	Average weight adjustment factor (WTB_adj6)
Alabama		
Public 2-year	76,988	0.97
Alaska		
Public 4-year	28,372	1.41
Arizona		
Public 2-year	205,717	1.28
Public 4-year	123,133	0.94
Other	237,230	1.78
Arkansas		
Public 2-year	43,152	1.12
Public 4-year	68,796	1.01
Other	16,076	1.29
California		
Public 2-year	1,260,111	1.02
Public 4-year	835,068	1.00
Other	357,915	1.29
Colorado		
Public 2-year	13,818	0.81
Public 4-year	204,047	0.96
Other	96,107	1.34
Connecticut		
Public 2-year	46,615	0.89
Delaware		
Public 4-year	37,086	1.04
Other	16,465	2.10
Florida		
Public 2-year	59,703	1.22
Public 4-year	747,009	1.06
Other	276,713	1.43
Georgia		
Public 2-year	142,020	1.43
Public 4-year	242,457	0.82
Other	71,643	0.98
Hawaii		
Public 2-year	21,653	0.93
Public 4-year	21,509	0.98
Other	8,978	1.26

See notes at end of table.

Table 77. Weight adjustment factors for undergraduate student study poststratification, by state and institution stratum: 2019–20—Continued

State and institution stratum ¹	Full-year enrollment control total ²	Average weight adjustment factor (WTB_adj6)
Illinois		
Public 2-year	339,869	1.20
Public 4-year	110,682	0.93
Other	178,040	1.11
Indiana		
Public 2-year	128,183	1.33
Iowa		
Public 2-year	56,564	0.91
Public 4-year	51,503	0.97
Other	40,279	0.96
Kansas		
Public 4-year	66,606	0.99
Kentucky		
Public 2-year	79,825	1.43
Public 4-year	71,566	1.01
Other	34,769	2.35
Louisiana		
Public 2-year	71,755	1.02
Public 4-year	104,735	0.97
Other	33,665	1.15
Maine		
Public 2-year	19,951	1.92
Public 4-year	22,587	0.93
Other	22,987	1.37
Maryland		
Public 4-year	121,152	1.21
Public 4-year	150,571	1.01
Other	37,763	1.28
Massachusetts		
Public 2-year	92,443	1.14
Public 4-year	97,702	0.98
Other	194,640	1.00
Michigan		
Public 2-year	149,401	0.99
Public 4-year	235,024	0.99
Other	70,332	1.28
Minnesota		
Public 2-year	94,894	0.86
Public 4-year	94,616	1.10
Other	84,181	1.36
Mississippi		
Public 2-year	72,349	1.09
Public 4-year	57,239	0.96
Other	15,637	1.33
Missouri		
Public 4-year	100,604	0.99
Nebraska		
Public 2-year	38,975	1.02
Public 4-year	40,599	0.98
Other	25,229	2.27

See notes at end of table.

Table 77. Weight adjustment factors for undergraduate student study poststratification, by state and institution stratum: 2019–20—Continued

State and institution stratum ¹	Full-year enrollment control total ²	Average weight adjustment factor (WTB_adj6)
Nevada		
Public 4-year	99,108	0.97
Other	8,890	1.10
New Hampshire		
Public 4-year	18,842	0.90
New Jersey		
Public 2-year	154,292	1.00
Public 4-year	155,396	1.05
Other	77,444	1.05
New Mexico		
Public 2-year	55,692	0.92
Public 4-year	38,931	1.03
Other	3,974	0.90
New York		
Public 2-year	298,328	1.20
Public 4-year	348,153	1.03
Other	402,320	1.19
North Carolina		
Public 2-year	215,707	1.05
Public 4-year	182,639	0.98
Other	78,301	1.03
Ohio		
Public 2-year	144,795	0.99
Public 4-year	274,563	1.07
Other	141,439	1.04
Oklahoma		
Public 2-year	58,664	1.16
Public 4-year	82,775	0.96
Other	30,670	1.57
Oregon		
Public 2-year	113,238	1.07
Public 4-year	78,212	1.23
Other	24,234	1.02
Pennsylvania		
Public 2-year	132,574	1.05
Public 4-year	185,774	1.00
Other	231,314	1.19
Puerto Rico		
Public 4-year	47,330	1.03
South Carolina		
Public 2-year	75,538	1.23
Public 4-year	97,363	0.95
Other	37,558	1.05
South Dakota		
Public 4-year	28,675	0.96
Tennessee		
Public 2-year	89,397	1.02
Public 4-year	105,063	1.01
Other	76,411	1.22

See notes at end of table.

Table 77. Weight adjustment factors for undergraduate student study poststratification, by state and institution stratum: 2019–20—Continued

State and institution stratum ¹	Full-year enrollment control total ²	Average weight adjustment factor (WTB_adj6)
Texas		
Public 2-year	549,046	1.08
Public 4-year	632,197	1.07
Other	150,376	1.23
Utah		
Public 4-year	130,016	1.02
Vermont		
Public 2-year	5,563	1.17
Public 4-year	17,602	1.03
Other	14,315	1.00
Virginia		
Public 2-year	142,166	1.01
Public 4-year	166,055	0.98
Other	127,926	0.95
Washington		
Public 2-year	28,839	0.88
Public 4-year	266,503	1.04
Other	40,528	1.06
West Virginia		
Public 2-year	17,715	1.16
Public 4-year	43,747	0.94
Other	81,424	1.38
Wisconsin		
Public 2-year	100,569	1.32
Public 4-year	156,409	1.01
Other	43,996	1.91
Wyoming		
Public 2-year	18,879	1.17
Public 4-year	9,119	0.93
All other institutions ³	1,701,923	1.38

¹ There were no public 2-year institutions in Alaska, Delaware, the District of Columbia, or Nevada.

² Enrollment control totals are the sum of full-year enrollment across institutions based on 2019–20 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 Idaho, Montana, North Dakota, and Rhode Island; public 2-year and other institutions in Kansas, Missouri, New Hampshire, Puerto Rico, South Dakota, and Utah; public 4-year and other institutions in Alabama, Connecticut, the District of Columbia and Indiana; and other institutions in Alaska and Wyoming.

NOTE: 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

There were 1,693 control totals in the final undergraduate student study poststratification model, which included control totals by public 2-year, public 4-year, and “all other sectors” within each state and state-level and national control totals. For the sake of brevity, table 78 shows only the national control totals that were included in the model as well as the average weight adjustment factors resulting from the undergraduate student study poststratification student model. Summary statistics of the undergraduate student study poststratification student weight adjustment factors follow:

- minimum: 0.0085;
- median: 0.89; and
- maximum: 276.89.

Table 78. Weight adjustment factors for undergraduate student study poststratification, by national poststratification category: 2019–20

National poststratification category ¹	Control total	Average weight adjustment factor (WTB_adj6)
Total, full-year undergraduate student enrollment	17,103,820	1.13
Full-year undergraduate student enrollment, by race, within control and level of institution ²		
Black, non-Hispanic		
Public less-than-2-year	7,071	0.69
Public 2-year	743,767	1.03
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	166,601	0.96
Public 4-year, non-doctorate-granting, primarily baccalaureate	102,821	0.98
Public 4-year, doctorate-granting	517,064	1.01
Private nonprofit 2-year or less	16,606	1.91
Private nonprofit 4-year, non-doctorate-granting	131,605	0.90
Private nonprofit 4-year, doctorate-granting	175,782	1.11
Private for-profit less-than-2-year	71,654	1.59
Private for-profit 2-year	54,849	2.13
Private for-profit 4-year	198,449	1.76
Hispanic		
Public less-than-2-year	10,177	1.65
Public 2-year	1,397,708	0.99
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	416,408	0.92
Public 4-year, non-doctorate-granting, primarily baccalaureate	166,882	1.03
Public 4-year, doctorate-granting	831,894	0.97
Private nonprofit 2-year or less	8,750	0.91
Private nonprofit 4-year, non-doctorate-granting	142,453	1.03
Private nonprofit 4-year, doctorate-granting	259,804	0.99
Private for-profit less-than-2-year	71,118	1.30
Private for-profit 2-year	62,561	1.64
Private for-profit 4-year	143,036	1.02
White, non-Hispanic		
Public less-than-2-year	38,396	1.52
Public 2-year	2,544,287	1.07
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	589,052	1.02
Public 4-year, non-doctorate-granting, primarily baccalaureate	462,943	0.97
Public 4-year, doctorate-granting	2,515,722	0.98
Private nonprofit 2-year or less	22,049	0.99
Private nonprofit 4-year, non-doctorate-granting	586,168	0.98
Private nonprofit 4-year, doctorate-granting	922,432	1.12
Private for-profit less-than-2-year	86,350	1.75
Private for-profit 2-year	74,809	2.58
Private for-profit 4-year	297,090	1.79

See notes at end of table.

Table 78. Weight adjustment factors for undergraduate student study poststratification, by national poststratification category: 2019–20—Continued

National poststratification category ¹	Control total	Average weight adjustment factor (WTB_adj6)
Other		
Public less-than-2-year	6,558	2.15
Public 2-year	938,613	1.37
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	261,368	1.46
Public 4-year, non-doctorate-granting, primarily baccalaureate	143,585	1.20
Public 4-year, doctorate-granting	991,169	1.13
Private nonprofit 2-year or less	12,179	1.74
Private nonprofit 4-year, non-doctorate-granting	200,180	1.51
Private nonprofit 4-year, doctorate-granting	429,542	1.47
Private for-profit less-than-2-year	39,343	2.39
Private for-profit 2-year	38,302	2.96
Private for-profit 4-year	206,623	3.08
Fall undergraduate student enrollment, by control and level of institution ²		
Public less-than-2-year	44,870	1.52
Public 2-year	3,826,076	0.93
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	1,003,466	0.94
Public 4-year, non-doctorate-granting, primarily baccalaureate	729,652	0.95
Public 4-year, doctorate-granting	4,282,541	0.95
Private nonprofit 2-year or less	43,330	1.10
Private nonprofit 4-year, non-doctorate-granting	883,166	0.98
Private nonprofit 4-year, doctorate-granting	1,494,867	1.07
Private for-profit less-than-2-year	159,296	1.55
Private for-profit 2-year	146,591	1.93
Private for-profit 4-year	512,586	1.44
Number of Pell Grant undergraduate recipients, by control and level of institution		
Public less-than-2-year	18,123	1.72
Public 2-year	2,040,225	1.12
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	503,600	1.00
Public 4-year, non-doctorate-granting, primarily baccalaureate	381,584	1.11
Public 4-year, doctorate-granting	1,917,496	1.25
Private nonprofit 2-year or less	49,495	2.63
Private nonprofit 4-year, non-doctorate-granting	436,460	1.21
Private nonprofit 4-year, doctorate-granting	609,663	1.44
Private for-profit less-than-2-year	211,441	2.02
Private for-profit 2-year	192,305	3.15
Private for-profit 4-year	519,400	2.53
Total amount of Pell Grants awarded to undergraduate students, by control and level of institution		
Public less-than-2-year	\$73,353,222	1.72
Public 2-year	7,363,244,519	1.12
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	1,858,282,536	1.00
Public 4-year, non-doctorate-granting, primarily baccalaureate	1,709,182,215	1.11
Public 4-year, doctorate-granting	8,781,170,371	1.25
Private nonprofit 2-year or less	234,676,098	2.63
Private nonprofit 4-year, non-doctorate-granting	1,942,990,357	1.21
Private nonprofit 4-year, doctorate-granting	2,591,366,471	1.44
Private for-profit less-than-2-year	814,253,077	2.02
Private for-profit 2-year	805,177,838	3.15
Private for-profit 4-year	2,170,328,036	2.53

See notes at end of table.

Table 78. Weight adjustment factors for undergraduate student study poststratification, by national poststratification category: 2019–20—Continued

National poststratification category¹	Control total	Average weight adjustment factor (WTB_adj6)
Total amount of PLUS Loans disbursed to undergraduate student recipients, by control and level of institution		
Public less-than-2-year	\$1,980,971	30.76
Public 2-year	78,251,823	0.34
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	73,006,738	0.60
Public 4-year, non-doctorate-granting, primarily baccalaureate	577,103,927	1.08
Public 4-year, doctorate-granting	5,637,677,253	1.19
Private nonprofit 2-year or less	25,486,139	1.76
Private nonprofit 4-year, non-doctorate-granting	1,634,442,063	1.08
Private nonprofit 4-year, doctorate-granting	3,479,855,502	1.32
Private for-profit less-than-2-year	124,083,972	1.27
Private for-profit 2-year	186,456,193	2.66
Private for-profit 4-year	369,709,970	2.80
Number of subsidized Direct Loan undergraduate student recipients, by control and level of institution		
Public less-than-2-year	7,068	1.71
Public 2-year	552,792	0.93
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	155,940	0.84
Public 4-year, non-doctorate-granting, primarily baccalaureate	309,722	1.15
Public 4-year, doctorate-granting	1,815,599	1.25
Private nonprofit 2-year or less	41,629	2.94
Private nonprofit 4-year, non-doctorate-granting	481,938	1.17
Private nonprofit 4-year, doctorate-granting	777,017	1.46
Private for-profit less-than-2-year	152,602	1.73
Private for-profit 2-year	159,574	3.05
Private for-profit 4-year	486,846	2.44
Total amount of subsidized Direct Loans disbursed to undergraduate students, by control and level of institution		
Public less-than-2-year	\$21,474,392	1.71
Public 2-year	1,598,762,396	0.93
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	478,073,743	0.84
Public 4-year, non-doctorate-granting, primarily baccalaureate	1,203,551,704	1.15
Public 4-year, doctorate-granting	7,329,025,707	1.25
Private nonprofit 2-year or less	151,866,778	2.94
Private nonprofit 4-year, non-doctorate-granting	1,902,111,026	1.17
Private nonprofit 4-year, doctorate-granting	3,178,257,704	1.46
Private for-profit less-than-2-year	436,188,416	1.73
Private for-profit 2-year	555,926,162	3.05
Private for-profit 4-year	1,871,349,179	2.44

See notes at end of table.

Table 78. Weight adjustment factors for undergraduate student study poststratification, by national poststratification category: 2019–20—Continued

National poststratification category ¹	Control total	Average weight adjustment factor (WTB_adj6)
Number of unsubsidized Direct Loans undergraduate student recipients, by control and level of institution		
Public less-than-2-year	8,288	1.66
Public 2-year	501,525	0.89
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	130,758	0.84
Public 4-year, non-doctorate-granting, primarily baccalaureate	309,765	1.14
Public 4-year, doctorate-granting	1,838,980	1.22
Private nonprofit 2-year or less	42,769	2.69
Private nonprofit 4-year, non-doctorate-granting	502,685	1.16
Private nonprofit 4-year, doctorate-granting	803,104	1.47
Private for-profit less-than-2-year	164,155	1.71
Private for-profit 2-year	165,251	3.07
Private for-profit 4-year	496,438	2.39
Total amount of unsubsidized Direct Loans disbursed to undergraduate students, by control and level of institution		
Public less-than-2-year	\$40,065,968	1.66
Public 2-year	1,834,596,926	0.89
Public 4-year, non-doctorate-granting, primarily subbaccalaureate	490,525,422	0.84
Public 4-year, non-doctorate-granting, primarily baccalaureate	1,226,308,998	1.14
Public 4-year, doctorate-granting	7,248,156,905	1.22
Private nonprofit 2-year or less	182,303,975	2.69
Private nonprofit 4-year, non-doctorate-granting	1,943,115,489	1.16
Private nonprofit 4-year, doctorate-granting	3,141,937,441	1.47
Private for-profit less-than-2-year	664,533,881	1.71
Private for-profit 2-year	780,005,343	3.07
Private for-profit 4-year	2,568,250,120	2.39

¹ There were 1,693 control totals in the final undergraduate poststratification model. For the sake of brevity, the table shows only the national control totals that were included in the model.

² Enrollment (full year and fall) control totals are the sum of enrollment across institutions based on 2019–20 Integrated Postsecondary Education Data System enrollment data adjusted for student multiplicity and dual enrollment.

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

7.1.3.8 Final analysis weights (WTA000 and WTB000)

As described above, the final analysis survey and study weights, denoted WTA000 and WTB000, respectively, were computed as the product of the weight components described in this section:

Final survey analysis weight (WTA000) = WTA_INSWT * WTA_WT4 * WTA_adj3 * WTA_adj4 * WTA_adj5 * WTA_adj6; and

Final study analysis weight for undergraduate students (WTB000) = WTB_INSTWT * WTB_WT4 * WTB_adj3 * WTB_adj4 * WTB_adj5 * WTB_adj6.

All weight adjustments, along with the final analysis weights, are included on the NPSAS:20 weights history files found in the NPSAS:20 RUF.

Weighted estimates were compared for key variables from the NPSAS:20 data with estimates from other sources—such as estimates from NPSAS:18-AC, NPSAS:16, FSA, and VBA. NPSAS:20 estimates were found to be reasonable considering differences in time frame, population, and other factors.⁵⁷

7.1.4 Weighting Adjustment Summary and Evaluation

This section summarizes the institution and student weight distributions and UWEs (see section 7.1 for more information about UWEs). The overall predictive ability of the nonresponse models are also assessed through receiver operating characteristic (ROC) curves.

7.1.4.1 Institution weighting adjustment summary and evaluation

Table 79 and table 80 summarize the institution weight distributions and UWEs by control and level of institution for undergraduate-enrolling survey and graduate-enrolling institutions, respectively. Table 81 summarizes the institution weight distributions and the UWEs by control and level of institution for undergraduate-enrolling study institutions.

The UWE for undergraduate-enrolling survey institutions is 8.25 overall and ranges from 1.02 for public 4-year, non-doctorate-granting, primarily baccalaureate institutions and public 4-year, doctorate-granting institutions to 5.97 for public less-than-2-year institutions, with most UWEs less than 4.00.

The UWE for graduate-enrolling institutions is 3.87 overall and ranges from 1.04 for public 4-year, doctorate-granting institutions to 4.08 for private nonprofit 4-year, doctorate-granting institutions, with most UWEs less than 4.00.

The UWE for undergraduate-enrolling study institutions is 10.29 overall and ranges from 1.05 for public 4-year, doctorate-granting institutions to 8.54 for private 2-year institutions, with most UWEs less than 4.00.

⁵⁷ When comparing NPSAS:20 estimates with those from other sources, there are potential differences between the data because of NPSAS poststratification to full-year enrollment, adjusted for multiplicity and dual enrollment, and NPSAS inclusion of institutions in Puerto Rico and degree-granting institutions.

Table 79. Institution weight distribution and unequal weighting effects for undergraduate-enrolling survey institutions, by control and level of institution: 2019–20

Control and level of institution	Minimum	First quartile	Median	Third quartile	Maximum	Mean	Unequal weighting effect ¹
Total	0.38	1.14	1.28	1.62	150.11	2.59	8.25
Public							
Less-than-2-year	1.23	1.47	3.04	13.81	150.11	14.52	5.97
2-year	0.99	1.15	1.26	1.42	4.88	1.34	1.05
4-year, non-doctorate-granting, primarily subbaccalaureate	0.97	1.14	1.23	1.36	2.18	1.28	1.03
4-year, non-doctorate-granting, primarily baccalaureate	1.00	1.10	1.17	1.29	2.19	1.22	1.02
4-year, doctorate-granting	0.97	1.06	1.13	1.20	2.12	1.16	1.02
Private nonprofit							
2-year or less	0.38	3.18	6.26	22.52	84.12	15.31	2.96
4-year, non-doctorate-granting	0.99	1.26	1.68	3.21	43.60	2.92	2.57
4-year, doctorate-granting	0.97	1.22	1.40	1.91	51.65	2.07	3.59
Private for-profit							
Less-than-2-year	1.28	2.04	4.48	13.54	135.57	11.20	3.57
2-year	0.87	1.40	2.28	7.26	85.25	7.95	4.95
4-year	0.81	1.48	2.57	7.08	29.17	5.31	2.44

¹ Unequal weighting effect calculated as sample size multiplied by the sum of the squared weights, divided by the sum of the weights squared.

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

Table 80. Institution weight distribution and unequal weighting effects for graduate-enrolling institutions, by control and level of institution: 2019–20

Control and level of institution	Minimum	First quartile	Median	Third quartile	Maximum	Mean	Unequal weighting effect ¹
Total	0.92	1.03	1.19	1.51	54.43	1.92	3.87
Public							
4-year, non-doctorate-granting, primarily subbaccalaureate	0.96	0.96	0.97	0.97	4.83	1.35	1.57
4-year, non-doctorate-granting, primarily baccalaureate	0.97	1.03	1.23	1.25	13.96	1.30	1.68
4-year, doctorate-granting	1.00	1.02	1.07	1.19	2.87	1.15	1.04
Private nonprofit							
4-year, non-doctorate-granting	0.95	1.03	1.41	2.89	51.74	2.87	3.68
4-year, doctorate-granting	0.99	1.12	1.36	1.99	54.43	2.25	4.08
Private for-profit							
4-year	0.92	1.10	1.67	4.61	31.76	4.01	3.04

¹ Unequal weighting effect calculated as sample size multiplied by the sum of the squared weights, divided by the sum of the weights squared.

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

Table 81. Institution weight distribution and unequal weighting effects for undergraduate-enrolling study institutions, by control and level of institution: 2019–20

Control and level of institution	Minimum	First quartile	Median	Third quartile	Maximum	Mean	Unequal weighting effect ¹
Total	0.92	1.04	1.19	1.67	165.24	2.63	10.29
Public							
Less-than-2-year	1.03	1.21	2.79	9.66	133.30	15.96	5.61
2-year	0.99	1.05	1.15	1.43	26.01	1.44	1.84
4-year, non-doctorate-granting, primarily subbaccalaureate	1.00	1.03	1.14	1.36	10.60	1.36	1.53
4-year, non-doctorate-granting, primarily baccalaureate	1.00	1.00	1.03	1.12	3.82	1.16	1.12
4-year, doctorate-granting	1.00	1.01	1.06	1.19	2.64	1.16	1.05
Private nonprofit							
2-year or less	1.00	1.07	4.07	14.60	145.67	17.62	5.31
4-year, non-doctorate-granting	0.94	1.10	1.42	2.74	61.62	2.83	3.75
4-year, doctorate-granting	0.92	1.07	1.35	2.05	53.22	2.16	3.66
Private for-profit							
Less-than-2-year	1.01	1.66	4.01	11.43	92.88	11.11	3.39
2-year	0.99	1.20	2.25	3.62	165.24	7.91	8.54
4-year	1.00	1.35	2.15	5.21	37.24	4.68	3.12

¹ Unequal weighting effect was calculated as sample size multiplied by the sum of the squared weights, divided by the sum of the weights squared.

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

The UWEs for undergraduate-enrolling study institutions shown in table 81 are inflated because the sample selection and weight adjustments were done within state by the three-level institution sampling strata (public 2-year, public 4-year, and “all other sectors”). The UWEs for the public 2-year and public 4-year study institutions range from 1.05 to 1.84, which are relatively small. Since “all other sectors” were grouped together for sampling and weighting, the UWEs for these sectors reflect that variation in the weights when the weights are reported at the control and level of the “other” institutions.

The UWEs indicate that for most institution types, the inflation of the variance of estimates due to unequal weighting is relatively small, and even for the institution types with higher UWEs, there is little concern about the effects unequal weighting could have on estimation. The sample design and sample sizes were determined to account for UWEs in this range to ensure precision of estimates.

To assess the overall predictive ability of the model used to adjust for institution nonresponse, NPSAS staff used an ROC curve (Hanley and McNeil 1982). The ROC curve provides a measure of how well the model correctly classified individuals of known response type—in other words, how well the model predicts

an institution's response propensity.⁵⁸ NPSAS staff developed the ROC curve in the following manner. The predicted probabilities of response (c) are derived from the model used to adjust for institution nonresponse. For any specified probability of response, c , two proportions were calculated:

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

The plot of the first probability against the second, for c from zero to 1, resulted in the ROC curves shown in figure 8 for undergraduate-enrolling survey institutions and in figure 9 for graduate-enrolling institutions and shown in figure 10 for undergraduate-enrolling study institutions. The area under the curve equals the probability that the fitted model correctly classifies two randomly chosen institutions—one of which is a true respondent, while the other is a true nonrespondent—where the institution with the higher predicted probability of response is classified as the respondent.

An area of 0.5 under an ROC curve indicates that a correct classification is made 50 percent of the time, with the model providing no predictive benefit. An area of 1.0 indicates that the true respondent always has the higher predicted probability of response as compared to the true nonrespondent, so the model always classifies the two institutions correctly.

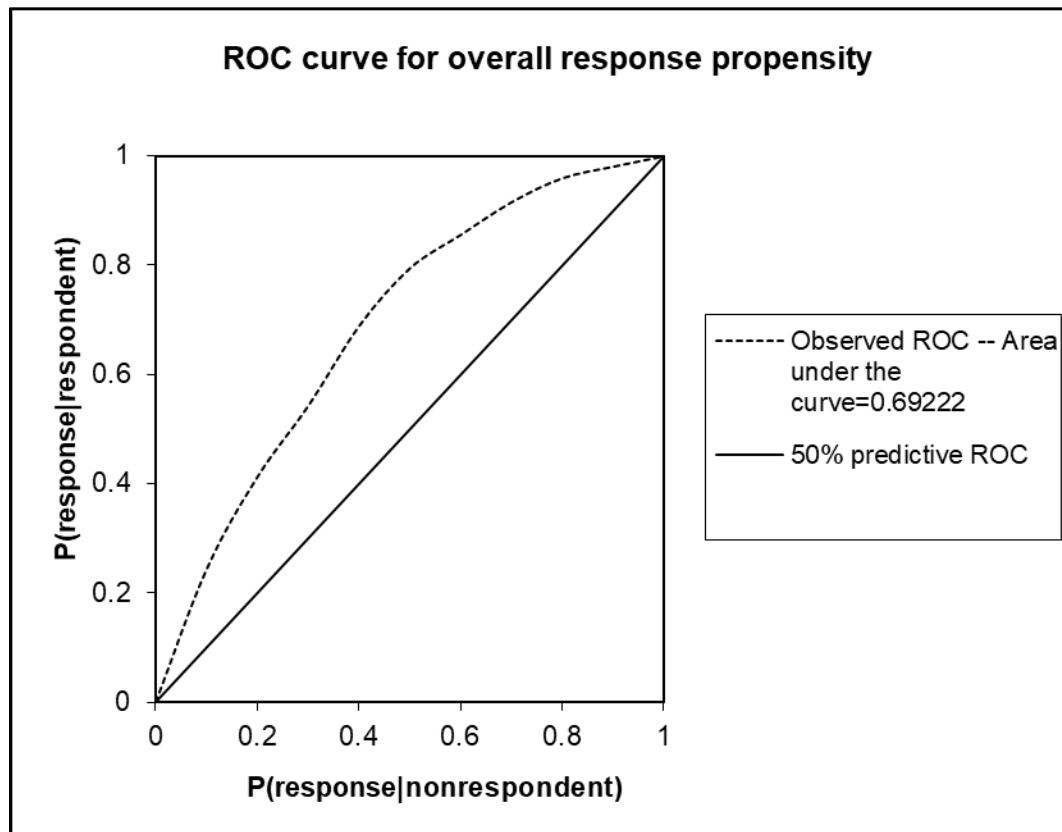
In figure 8, the area under the ROC curve for the undergraduate-enrolling survey institutions is approximately 0.69, so the predicted probabilities give the correct classification about 69 percent of the time. In figure 9, the area under the ROC curve for the graduate-enrolling institutions is approximately 0.63, so the predicted probabilities give the correct classification about 63 percent of the time. In figure 10, the area under the ROC curve for the undergraduate-enrolling study institutions approximately 0.70, so the predicted probabilities give the correct classification about 70 percent of the time.

Researchers can also interpret predictive probabilities from ROC curves in terms of the nonparametric Wilcoxon test statistic, which is used to determine whether the level of a quantitative variable, such as predicted probability of response, is different between two samples (respondents and nonrespondents in this case). The ROC area equals the value of the Wilcoxon test statistic. Viewed in this way, the

⁵⁸ For a more detailed example of the ROC curve used in nonresponse modeling, see Iannacchione (2003).

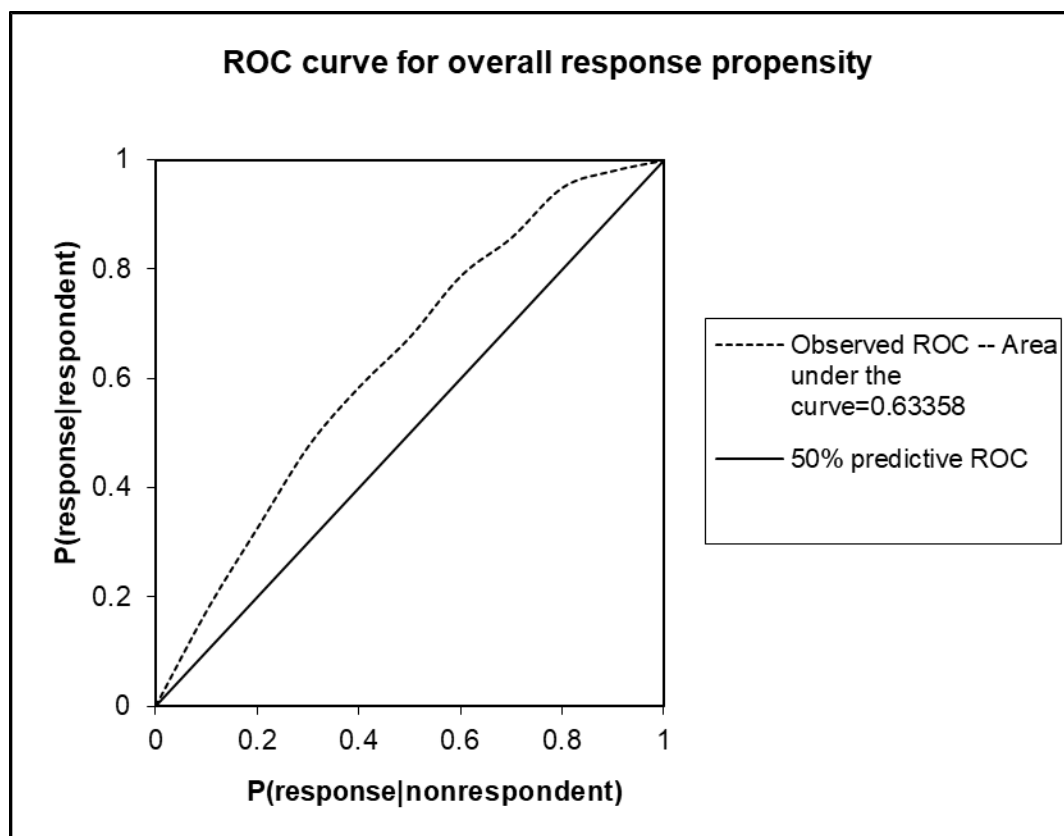
Wilcoxon test rejects the null hypothesis of no predictive ability by showing that the predicted probability of response for the respondents is larger than that for the nonrespondents. Analysts can interpret this result to mean that the variables used in the model are informative predictors of a sample institution's overall response propensity.

Figure 8. Receiver operating characteristic (ROC) curve for overall institution response propensity for undergraduate-enrolling survey institutions: 2019–20



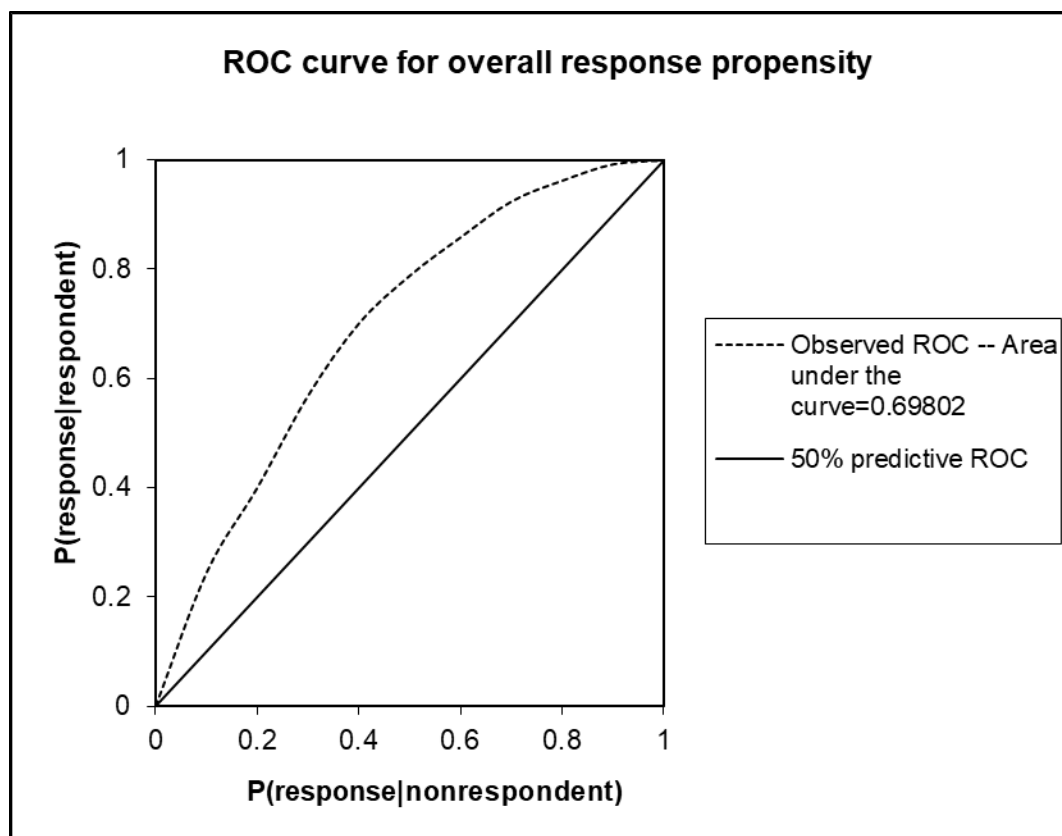
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

Figure 9. Receiver operating characteristic (ROC) curve for overall institution response propensity for graduate-enrolling institutions: 2019–20



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

Figure 10. Receiver operating characteristic (ROC) curve for overall institution response propensity for undergraduate-enrolling study institutions 2019–20



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

7.1.4.2 Student weighting adjustment summary and evaluation

Table 82 and table 83 summarize the student survey weight distributions and the variance inflation caused by the UWE, by control and level of institution for undergraduate and graduate students, respectively. Table 84 summarizes the student study weight distributions and the variance inflation caused by the UWE by control and level of institution.

For survey undergraduates, the UWE is 2.45 overall and ranges from 1.90 for undergraduate survey students in public 4-year, doctorate-granting institutions to 5.09 for undergraduate survey students in private for-profit 4-year institutions. For graduates, the UWE is 2.90 overall and ranges from 2.42 for graduate students in public 4-year, doctorate-granting institutions to 4.35 for graduate students in private for-profit 4-year institutions. For study undergraduates, the UWE is 2.73 overall and ranges from 1.84 for undergraduate students in public 4-year, doctorate-granting institutions to 7.00 for undergraduate survey students in private for-profit 4-year institutions.

Table 82. Undergraduate student survey respondents weight distribution and unequal weighting effects, by control and level of institution: 2019–20

Control and level of institution	Minimum	First quartile	Median	Third quartile	Maximum	Mean	Unequal weighting effect ¹
Total	1.00	54.20	120.21	280.59	4,641.04	211.79	2.45
Public							
Less-than-2-year	2.45	22.13	46.11	101.29	1,079.67	93.82	3.05
2-year	1.01	60.24	121.68	290.48	2,181.84	216.81	2.31
4-year, non-doctorate-granting, primarily subbaccalaureate	1.06	71.23	139.48	388.44	3,064.10	285.77	2.49
4-year, non-doctorate-granting, primarily baccalaureate	2.97	54.20	92.63	231.19	1,396.57	155.11	1.93
4-year, doctorate-granting	1.00	92.15	189.05	414.93	3,402.10	281.74	1.90
Private nonprofit							
2-year or less	1.00	13.63	32.94	97.59	1,386.19	91.95	4.27
4-year, non-doctorate-granting	1.04	67.61	125.40	212.75	2,393.69	170.13	2.03
4-year, doctorate-granting	1.14	65.06	147.85	319.69	4,641.04	234.50	2.51
Private for-profit							
Less-than-2-year	1.58	25.21	62.72	137.27	1,145.27	108.65	2.51
2-year	1.19	14.81	34.88	85.77	1,015.17	68.06	2.83
4-year	1.00	9.65	32.19	124.61	4,044.42	143.45	5.09

¹ Unequal weighting effect was calculated as sample size multiplied by the sum of the squared weights, divided by the sum of the weights squared.

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

Table 83. Graduate student respondents weight distribution and unequal weighting effects, by control and level of institution: 2019–20

Control and level of institution and graduate student type	Minimum	First quartile	Median	Third quartile	Maximum	Mean	Unequal weighting effect ¹
Total	1.00	30.49	88.25	206.16	3,623.67	185.06	2.90
Control and level of institution							
Public							
4-year, non-doctorate-granting, primarily subbaccalaureate	3.39	25.73	45.78	107.32	446.69	97.89	2.49
4-year, non-doctorate-granting, primarily baccalaureate	1.00	21.49	58.90	120.37	1,363.85	91.49	2.45
4-year, doctorate-granting	1.25	40.94	135.78	247.30	1,844.09	222.52	2.42
Private nonprofit							
4-year, non-doctorate-granting	1.11	38.32	75.04	175.08	1,646.08	140.44	2.72
4-year, doctorate-granting	1.12	37.41	91.09	255.01	2,291.02	218.92	2.78
Private for-profit							
4-year	1.00	12.71	45.21	111.95	3,623.67	99.99	4.35
Graduate student type							
Graduate (excluding doctoral–professional practice)	1.00	31.18	85.48	214.74	3,623.67	191.06	2.91
Doctoral–professional practice	1.25	26.70	107.63	182.49	1,565.54	153.44	2.72

¹ Unequal weighting effect was calculated as sample size multiplied by the sum of the squared weights, divided by the sum of the weights squared.

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

Table 84. Undergraduate student study respondents weight distribution and unequal weighting effects, by control and level of institution: 2019–20

Control and level of institution	Minimum	First quartile	Median	Third quartile	Maximum	Mean	Unequal weighting effect ¹
Total	1.00	19.04	40.38	75.16	4,483.44	61.89	2.73
Public							
Less-than-2-year	1.00	7.64	10.19	23.67	565.23	27.78	4.69
2-year	1.00	18.12	37.23	70.72	3,199.82	57.60	2.39
4-year, non-doctorate-granting, primarily subbaccalaureate	1.00	35.16	60.76	116.63	2,890.94	91.18	2.04
4-year, non-doctorate-granting, primarily baccalaureate	1.00	21.02	40.11	67.73	2,323.65	53.48	2.23
4-year, doctorate-granting	1.00	30.00	48.40	80.57	1,381.19	63.95	1.84
Private nonprofit							
2-year or less	1.00	3.61	7.97	33.28	1,559.97	43.36	8.45
4-year, non-doctorate-granting	1.00	12.57	27.49	64.11	1,478.65	55.84	3.06
4-year, doctorate-granting	1.00	15.15	36.14	74.08	1,814.84	66.18	3.46
Private for-profit							
Less-than-2-year	1.00	9.25	30.07	69.60	928.59	55.01	2.94
2-year	1.00	6.58	18.53	47.81	1,338.12	40.62	3.67
4-year	1.00	5.76	21.47	70.55	4,483.44	80.52	7.00

¹ Unequal weighting effect was calculated as sample size multiplied by the sum of the squared weights, divided by the sum of the weights squared.

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

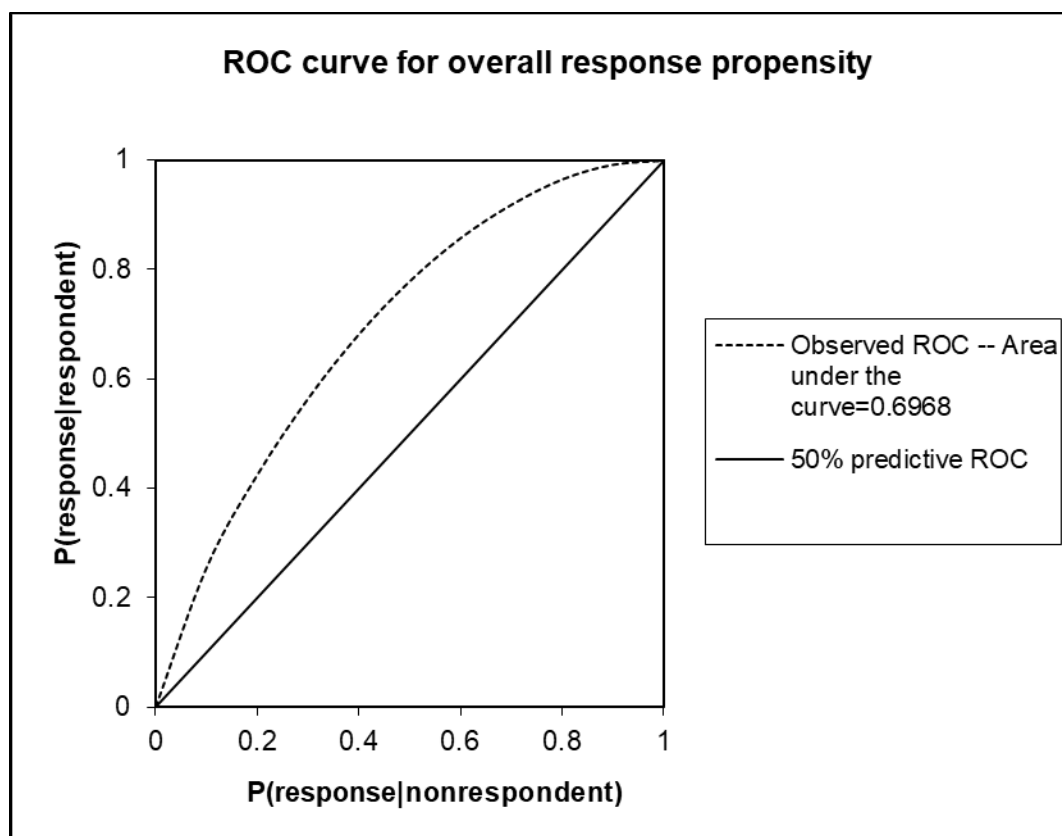
To assess the overall predictive ability of the student nonresponse model, an ROC curve was developed as described in the previous section. However, the predicted probabilities of response (c) for the ROC curve associated with the student nonresponse are the product of the predicted response probabilities obtained at each of the three nonresponse adjustment steps.

Figure 11 shows that the area under the ROC curve for the undergraduate survey students is approximately 0.70, so the predicted probabilities give the correct classification about 70 percent of the time. Figure 12 shows that the area under the ROC curve for the graduate students is approximately 0.69, so the predicted probabilities give the correct classification about 69 percent of the time. Figure 13 shows that the area under the ROC curve for undergraduate study students is approximately 0.85, so the predicted probabilities give the correct classification about 85 percent of the time.

Predictive probabilities from ROC curves can also be interpreted in terms of the nonparametric Wilcoxon test statistic, where the ROC area is equivalent to the value of the Wilcoxon test statistic. Viewed in this way, the Wilcoxon test rejects the null hypothesis of no predictive ability by showing that the predicted probability of response for the respondents is larger than that for the

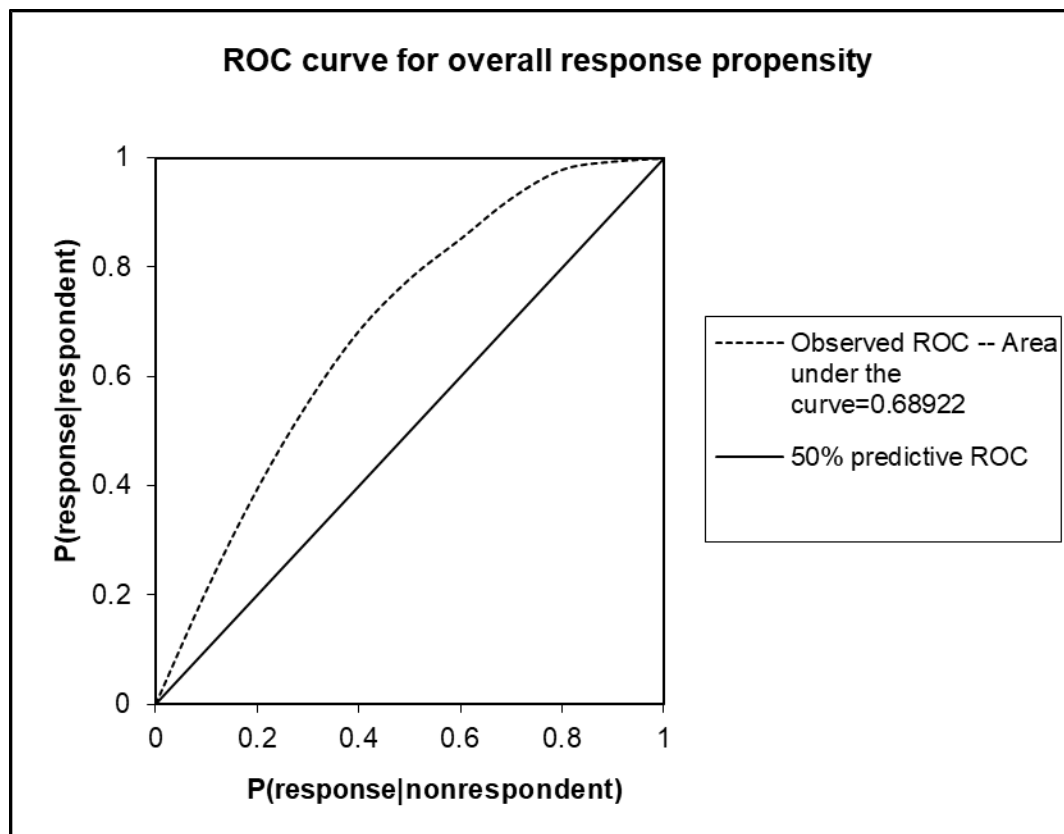
nonrespondents. This result can be interpreted to mean that the variables used in the model are informative predictors of a sample student's overall response propensity.

Figure 11. Receiver operating characteristic (ROC) curve for overall undergraduate survey student response propensity: 2019–20



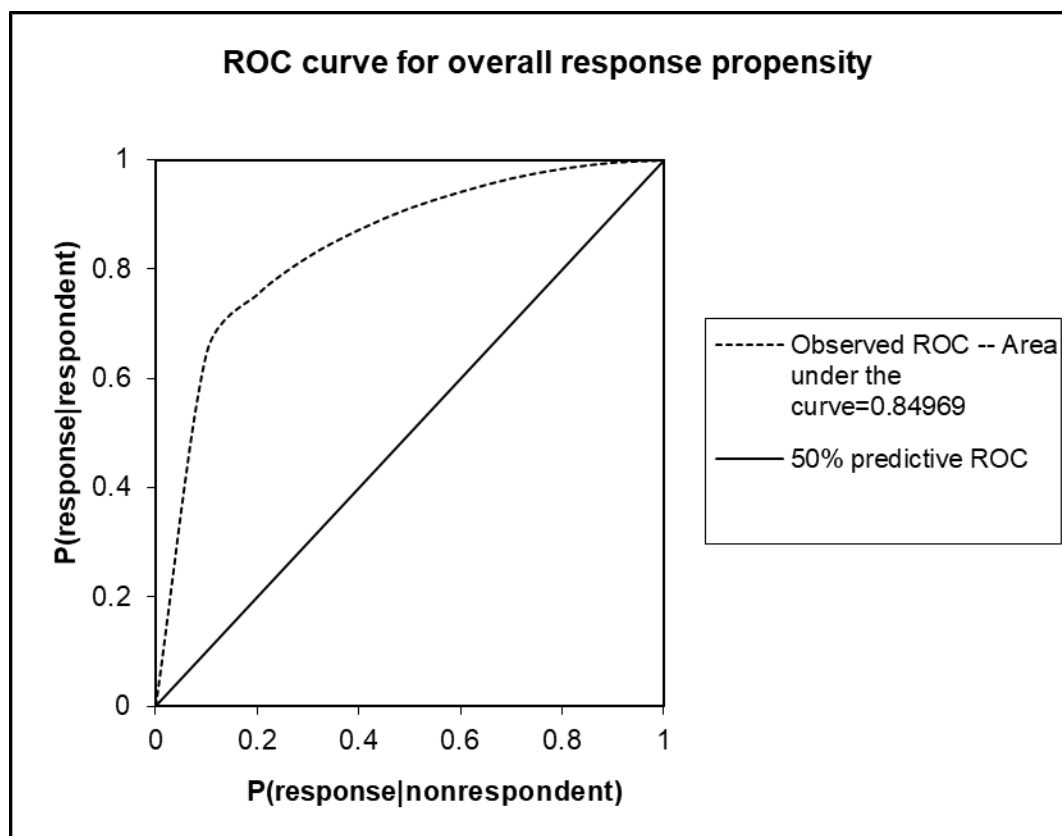
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

Figure 12. Receiver operating characteristic (ROC) curve for overall graduate student response propensity: 2019–20



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

Figure 13. Receiver operating characteristic (ROC) curve for overall undergraduate study student response propensity: 2019–20



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

7.2 Nonresponse Bias Analysis

The accuracy of survey statistics is affected by both random and nonrandom errors. Random errors reduce the precision of survey estimates; 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 (i.e., an 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). Nonrandom errors can result in biased estimates for administrative data as well. The NPSAS:20 data collection procedures were thoroughly developed and tested to minimize nonrandom errors because these errors are difficult to quantify.

Sample studies typically encounter two general types of nonresponse: unit nonresponse and item nonresponse. For NPSAS:20, 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 section 7.1. Item nonresponse occurs when a student or institution respondent is missing values for one or more items. Errors due to nonresponse are often, but not always, nonrandom; 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 that “Any survey stage of data collection with a unit or item response rate less than 85 percent must be evaluated for the potential magnitude of nonresponse bias before the data or any analysis using the data may be released. Estimates of survey characteristics for nonrespondents and respondents are required to assess the potential nonresponse bias” (Seastrom 2014).

Due to many unit response rates below 85 percent (section 7.1.1), institution-level and student-level unit nonresponse bias analyses were conducted for the sample overall, and across institution level and control, regardless of whether the unit response rate was below 85 percent. Item nonresponse bias analyses were performed for all survey and administrative items that had a weighted response rate below 85 percent, using the appropriate analysis weight (undergraduate survey, undergraduate study, or graduate student) and also for students overall and by institution level and control. All analyses were conducted separately for the undergraduate student study, undergraduate student survey, and graduate student survey. The procedures for unit and item nonresponse bias analyses are described in 7.2.1. The institution-level results are summarized in section 7.2.2, while student-level results are summarized in section 7.2.3. Item-level results and response rates are summarized in section 7.2.4. Detailed nonresponse bias tables for all analyses are provided in appendix M.

7.2.1 Nonresponse Bias Analysis Procedures

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: $\hat{B}(\bar{y}_R) = \eta(\bar{y}_R - \bar{y}_{NR})$. As this formulation makes apparent, if

the difference between the respondent and nonrespondent means is small, then biases may be small even with a high nonresponse rate.

Relative bias (*RB*) provides a measure of the magnitude of the bias relative to the sample mean and is estimated as $\widehat{RB}(\bar{y}_R) = \hat{B}(\bar{y}_R)/\hat{\pi}$, often expressed as a percentage. Effect size, as defined by Cohen (1988), is another measure of potential nonresponse bias. For continuous variables, effect size 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. Interpretation of relative bias is subjective, but, generally, relative bias less than 10 percent is good, and analysts should use caution as relative bias increases above 10 percent. 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.

Unit nonresponse bias analysis procedures. For each unit nonresponse bias analysis conducted, nonresponse bias was estimated for each category of selected variables (with nonmissing data for both respondents and nonrespondents) 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 (sections 7.1.2.2 and 7.1.3.6), 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).

Item nonresponse bias analysis procedures. For each item nonresponse bias analysis, nonresponse bias was estimated for each category of selected 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 whether it differed significantly from zero at the 5 percent significance level. Significance tests were complemented by relative bias and effect size calculations.

Imputation procedures (see section 7.4) 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 size-weighted means of category-level differences⁵⁹ were calculated and labeled as significant if any category-level difference was significant. These tests were 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.

7.2.2 Unit Nonresponse 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 IPEDS data files used in institution weight adjustments (section 7.1.2). As defined in section 7.1.1.1, an institution respondent is a sampled institution that provided a student enrollment list from which a sample was selected. The variables for the institution nonresponse bias analysis, listed in table 85, included variables used in the institution nonresponse adjustment (see section 7.1.2.2) and additional variables not used in the weight adjustment model.

⁵⁹ The size-weighted means are weighted using the unweighted count of eligible students in each category for the variable.

Table 85. Variables used in institution-level unit nonresponse bias analysis: 2019–20

Variable	Survey undergraduate	Graduate	Study undergraduate
Control and level of institution	✓	✓	✓
State, control, and level of institution			✓
Carnegie classification 2015: Basic	✓	✓	✓
Degree of urbanization	✓	✓	✓
Region of institution	✓	✓	✓
Historically Black College or University	✓	✓	✓
Hispanic-Serving Institution	✓	✓	✓
Percentage of undergraduate or graduate students enrolled: Black, non-Hispanic	✓	✓	✓
Percentage of undergraduate or graduate students enrolled: Asian or Pacific Islander, non-Hispanic	✓	✓	✓
Percentage of undergraduate or graduate students enrolled: Hispanic	✓	✓	✓
Total undergraduate or graduate enrollment	✓	✓	✓
Total male undergraduate or graduate enrollment	✓	✓	✓
Total female undergraduate or graduate enrollment	✓	✓	✓
Percentage of undergraduate students receiving federal grant aid	✓		✓
Percentage of undergraduate students receiving state/local grant aid	✓		✓
Percentage of undergraduate students receiving institutional 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	✓		✓
Number of full-time, first-time undergraduate students receiving Title IV aid with incomes up to \$30,000	✓		✓
Average amount of grant and scholarship aid received	✓		✓
Percentage of full-time, first-time degree/certificate-seeking undergraduate students graduating within 150 percent of normal time	✓		✓
Number of office and administrative support employees		✓	

NOTE: 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.

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

7.2.2.1 Nonresponse bias analysis for undergraduate-enrolling institutions—Survey

As shown in table 86, for undergraduate-enrolling institutions overall, the institution survey nonresponse weighting adjustment reduced statistically significant bias on the observable characteristics that met reporting requirements. The bias reduction results across institution control and level were mixed as the nonresponse adjustment model (section 7.1.3.6) was conducted with the primary goal of reducing bias in student-level analyses. For most categories of control and level, there was no substantial evidence of bias either before or after weight adjustment; however, a moderate amount of bias remained after weight adjustment for private for-profit institutions.

Before weighting, the median percent relative bias across characteristics was 10.96 for institutions overall and ranged from 3.04 to 59.62 for institutions by control and level (table 86). The statistics summarized in table 86 can be found in appendix tables M-1 through M-12. After weighting, the median percent relative bias was 6.22 for institutions overall and ranged from 2.03 to 43.64 across institution control and level. The percentage of characteristics with significant bias before weighting was 33.61 for institutions overall and ranged from zero to 14.63 across institution control and level. There is a moderate amount of small but statistically significant bias, but characteristics with significant bias are not particularly concerning if the magnitudes are small prior to weight adjustment. After weighting, the percentage was 1.68 for institutions overall and between zero and 21.95 across institution control and level. The median effect size was 0.12 for institutions overall before weighting and ranged from 0.03 to 0.33 across institution control and level. After weighting, the median effect size was 0.10 for institutions overall and ranged from 0.04 to 0.41 across institution control and level.

As seen in table 87, the mean absolute difference between respondent means before and after poststratification was 0.61 for undergraduate-enrolling institutions overall and ranged from 0.00 to 3.11 across institution control and level. The statistics summarized in table 87 can be found in appendix tables M-13 through M-24. The median absolute difference was 0.27 for institutions overall and ranged from 0.00 to 2.08 across institution control and level. The mean absolute difference between the full-sample and respondent means after poststratification was 1.13 for institutions overall and ranged from 0.65 to 8.35 across institution and level. The median absolute difference was 0.81 for institutions overall and ranged from 0.51 to 8.17 across institution and level.

Table 86. Summary of institution-level survey nonresponse bias analysis for undergraduate-enrolling institutions, by control and level of institution: 2019–20

Nonresponse bias statistics ¹	Overall	Public less-than-2-year	Public 2-year	Public 4-year, non-doctorate-granting, primarily subbaccalaureate	Public 4-year, non-doctorate-granting, primarily baccalaureate	Public 4-year, doctorate-granting	Private nonprofit less-than-4-year	Private nonprofit 4-year, non-doctorate-granting	Private nonprofit 4-year, doctorate-granting	Private for-profit less-than-2-year	Private for-profit 2-year	Private for-profit 4-year
Before nonresponse weight adjustments²												
Mean percent relative bias across characteristics	12.91	23.34	6.54	7.80	5.91	3.22	51.11	13.93	7.34	32.34	26.46	22.85
Median percent relative bias across characteristics	10.96	16.13	3.57	5.49	4.25	3.04	59.62	8.91	5.40	33.64	20.11	19.19
Percentage of characteristics with significant bias	33.61	#	12.37	9.72	10.26	7.79	8.00	4.35	1.10	7.06	#	14.63
Median effect size	0.12	‡	0.06	0.10	0.05	0.03	‡	0.15	0.07	0.33	0.27	0.25
After nonresponse weight adjustments³												
Mean percent relative bias across characteristics	10.68	34.10	6.20	7.39	5.83	2.86	45.29	9.02	7.59	36.53	28.34	34.60
Median percent relative bias across characteristics	6.22	20.81	3.75	6.23	4.48	2.03	43.64	6.24	6.57	28.42	19.69	31.85
Percentage of characteristics with significant bias	1.68	#	6.19	4.17	5.13	5.19	4.00	#	1.10	7.06	#	21.95
Median effect size	0.10	‡	0.04	0.08	0.04	0.04	‡	0.08	0.08	0.35	0.28	0.41

Rounds to zero.

‡ Reporting standards not met (fewer than five characteristics had at least five nonrespondents).

¹ Relative bias and effect size are calculated using the weighted differences between respondent and full-sample means. Relative bias is calculated as 100 times the ratio of estimated bias to the weighted full-sample mean. Effect size is calculated as the square root of the sum over categories of the squared differences over full-sample means.

² Respondent and full-sample means are weighted using the institution base weight.

³ Full-sample means are weighted using the institution base weight, and respondent means are 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. "Institution base weight" refers to the institution sampling weight.

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

Table 87. Summary of institution-level survey analysis after poststratification for undergraduate-enrolling institutions, by control and level of institution: 2019–20

Nonresponse bias statistics	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 adjustment¹												
Mean absolute difference across characteristics	0.61	#	0.12	0.18	0.31	0.20	3.11	1.33	0.32	0.64	2.00	1.76
Median absolute difference across characteristics	0.27	#	0.08	0.14	0.18	0.14	2.08	1.09	0.25	0.43	1.65	1.42
Difference between means for full sample and respondents after poststratification adjustment²												
Mean absolute difference across characteristics	1.13	6.33	0.75	1.65	1.17	0.65	8.35	2.22	1.48	5.68	5.12	5.87
Median absolute difference across characteristics	0.81	4.82	0.59	1.58	1.00	0.51	8.17	1.88	1.36	4.54	4.04	5.19

Rounds to zero.

¹ Respondent means before poststratification adjustment are weighted using the institution base weight adjusted for nonresponse. Respondent means after poststratification adjustment are weighted using the institution base weight adjusted for nonresponse and poststratification.² Full-sample means are weighted using the institution base weight, and respondent means are weighted using the institution base weight adjusted for nonresponse and poststratification.

NOTE: Characteristics that did not meet reporting standards were excluded from calculation of summary statistics. "Institution base weight" refers to the institution sampling weight.

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

7.2.2.2 Nonresponse bias analysis for undergraduate-enrolling institutions—Study

As seen in table 88, the nonresponse weight adjustments appear as if they are largely ineffective at reducing institution-level bias; in fact, in many sectors there appears to be an increase in bias after weight adjustment. This is expected because the institution-level nonresponse weight adjustments were conducted with the primary goal of reducing nonresponse bias at the student level. Further, the study nonresponse weight adjustment model did not include all levels of institution control and level; rather, it included public 2-year and public 4-year institutions within states. Nonetheless, for most categories of institution level and control, there is little evidence of large biases either before or after weight adjustment. For private for-profit categories, the median effect sizes after weight adjustment were moderate to large.

Before weighting, the median percent relative bias across characteristics was 14.34 for institutions overall and ranged from 2.84 to 59.62 for institutions across control and level. After weighting, the median percent relative bias was 9.80 for institutions overall and ranged from 2.53 to 44.47 across institution control and level. The percentage of characteristics with significant bias before weighting was 26.80 for institutions overall and ranged from zero to 14.63 across institution control and level. After weighting, the percentage was 5.67 for institutions overall and between zero and 10.42 across institution control and level. The median effect size was 0.12 for institutions overall before weighting and ranged from 0.03 to 0.32 across institution control and level. After weighting, the median effect size was 0.12 for institutions overall and ranged from 0.05 to 0.77 across institution control and level. The statistics summarized in table 88 can be found in appendix tables M-25 through M-36.

As seen in table 89, the mean absolute difference between respondent means before and after poststratification was 0.07 for undergraduate-enrolling institutions overall and ranged from 0.00 to 0.52 across institution control and level. The median absolute difference was 0.04 for institutions overall and ranged from 0.00 to 0.39 across institution control and level. The mean absolute difference between the full-sample and respondent means after poststratification was 0.96 for institutions overall and ranged from 0.76 to 10.29 across institution and level. The median absolute difference was 0.41 for institutions overall and ranged from 0.56 to 10.39 across institution and level. The statistics summarized in table 87 can be found in appendix tables M-37 through M-48.

Table 88. Summary of institution-level study nonresponse bias analysis for undergraduate-enrolling institutions, by control and level of institution: 2019–20

Nonresponse bias statistics ¹	Overall	Public less-than-2-year	Public 2-year	Public 4-year, non-doctorate-granting, primarily subbaccalaureate	Public 4-year, non-doctorate-granting, primarily baccalaureate	Public 4-year, doctorate-granting	Private nonprofit less-than-4-year	Private nonprofit 4-year, non-doctorate-granting	Private nonprofit 4-year, doctorate-granting	Private for-profit less-than-2-year	Private for-profit 2-year	Private for-profit 4-year
Before nonresponse weight adjustments²												
Mean percent relative bias across characteristics	16.99	22.83	5.70	7.76	5.91	3.27	51.11	12.29	6.17	32.11	26.46	22.85
Median percent relative bias across characteristics	14.34	13.05	3.09	6.03	4.25	2.84	59.62	7.24	4.61	33.34	20.11	19.19
Percentage of characteristics with significant bias	26.80	#	11.46	11.11	10.26	2.60	8.00	2.20	#	7.06	#	14.63
Median effect size	0.12	‡	0.05	0.11	0.05	0.03	‡	0.15	0.04	0.32	0.27	0.25
After nonresponse weight adjustments³												
Mean percent relative bias across characteristics	19.95	35.77	9.42	11.80	7.97	3.22	44.47	16.16	7.63	30.79	57.60	30.67
Median percent relative bias across characteristics	9.80	25.59	6.66	10.13	7.27	2.53	44.47	13.03	6.96	24.88	37.89	25.81
Percentage of characteristics with significant bias	5.67	#	10.42	6.94	7.69	5.19	#	2.20	#	5.88	9.20	7.32
Median effect size	0.12	‡	0.08	0.15	0.07	0.05	‡	0.19	0.07	0.34	0.77	0.34

Rounds to zero.

‡ Reporting standards not met (fewer than five characteristics had at least five nonrespondents).

¹ Relative bias and effect size are calculated using the weighted differences between respondent and full-sample means. Relative bias is calculated as 100 times the ratio of estimated bias to the weighted full-sample mean. Effect size is calculated as the square root of the sum over categories of the squared differences over full-sample means.² Respondent and full-sample means are weighted using the institution base weight.³ Full-sample means are weighted using the institution base weight, and respondent means are 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. "Institution base weight" refers to the institution sampling weight.

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

Table 89. Summary of institution-level study analysis after poststratification for undergraduate-enrolling institutions, by control and level of institution: 2019–20

Nonresponse bias statistics	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 adjustment¹												
Mean absolute difference across characteristics	0.07	0.50	0.11	#	#	#	0.52	0.30	0.14	0.20	0.33	0.24
Median absolute difference across characteristics	0.04	0.29	0.07	#	#	#	0.39	0.23	0.14	0.16	0.20	0.17
Difference between means for full sample and respondents after poststratification adjustment²												
Mean absolute difference across characteristics	0.96	7.10	1.26	2.77	1.63	0.76	10.29	2.80	1.45	6.08	8.73	5.96
Median absolute difference across characteristics	0.41	3.18	0.95	2.52	1.77	0.56	10.39	2.21	1.25	5.08	6.17	5.16

Rounds to zero.

¹ Respondent means before poststratification adjustment are weighted using the institution base weight adjusted for nonresponse. Respondent means after poststratification adjustment are weighted using the institution base weight adjusted for nonresponse and poststratification.

² Full-sample means are weighted using the institution base weight, and respondent means are weighted using the institution base weight adjusted for nonresponse and poststratification.

NOTE: Characteristics that did not meet reporting standards were excluded from calculation of summary statistics. "Institution base weight" refers to the institution sampling weight.

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

7.2.2.3 *Nonresponse bias analysis for graduate-enrolling institutions*

As shown in table 90, for graduate-enrolling institutions overall, the institution nonresponse weighting adjustment resulted in a slight bias reduction on the observable characteristics that met reporting requirements. Across categories of control and level, there were mixed results, with some evidence of bias increasing; however, there was no evidence of substantive bias either before or after weight adjustment. As was done for the undergraduate-enrolling institution weights, graduate-enrolling institution weight adjustments were conducted with the goal of reducing student-level bias rather than institution-level bias.

Before weighting, the median percent relative bias across characteristics was 6.63 for institutions overall and ranged from 5.58 to 23.26 for institutions by control and level. After weighting, the median percent relative bias was 4.36 for institutions overall and ranged from 2.37 to 16.75 across institution control and level. The percentage of characteristics with significant bias before weighting was 19.70 for institutions overall and ranged from 2.33 to 34.21 across institution control and level. After weighting, the percentage was 12.12 for institutions overall and between 4.26 and 15.79 across institution control and level. The median effect size was 0.06 for institutions overall before weighting and ranged from 0.02 to 0.20 across institution control and level. After weighting, the median effect size was 0.05 for institutions overall and ranged from 0.01 to 0.23 across institution control and level. The statistics summarized in table 90 can be found in appendix tables M-49 through M-55.

Table 90. Summary of institution-level nonresponse bias analysis for graduate-enrolling institutions, by control and level of institution: 2019–20

Nonresponse bias statistics 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 adjustments²							
Mean percent relative bias across characteristics	8.10	‡	5.96	3.20	15.06	9.70	23.80
Median percent relative bias across characteristics	6.63	‡	5.58	1.93	13.04	9.77	23.26
Percentage of characteristics with significant bias	19.70	‡	5.88	2.33	4.00	4.26	34.21
Median effect size	0.06	‡	0.04	0.02	0.13	0.09	0.20
After nonresponse weight adjustments³							
Mean percent relative bias across characteristics	9.78	‡	11.31	3.46	16.86	10.40	26.88
Median percent relative bias across characteristics	4.36	‡	8.24	2.37	13.41	8.34	16.75
Percentage of characteristics with significant bias	12.12	‡	5.88	11.63	8.00	4.26	15.79
Median effect size	0.05	‡	0.12	0.01	0.16	0.10	0.23

‡ Reporting standards not met (fewer than five characteristics had at least five nonrespondents).

¹ Relative bias and effect size are calculated using the weighted differences between respondent and full-sample means. Relative bias is calculated as 100 times the ratio of estimated bias to the weighted full-sample mean. Effect size is calculated as the square root of the sum over categories of the squared differences over full-sample means.

² Respondent and full-sample means are weighted using the institution base weight.

³ Full-sample means are weighted using the institution base weight, and respondent means are 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. "Institution base weight" refers to the institution sampling weight.

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

As shown in table 91, the mean absolute difference between respondent means before and after poststratification was 0.18 for graduate-enrolling institutions overall and ranged from 0.01 to 1.49 across control and level of institution. The median absolute difference was 0.10 for institutions overall and ranged from zero to 1.35 across control and level of institution. The mean absolute difference between the full-sample and respondent means after poststratification was 1.05 for institutions overall and ranged from 0.73 to 4.64 across control and level of institution. The median absolute difference was 0.66 for institutions overall and ranged from 0.50 to 4.10 across control and level of institution. The statistics summarized in table 91 can be found in appendix tables M-56 through M-62.

Table 91. Summary of institution-level analysis after poststratification for graduate-enrolling institutions, by control and level of institution: 2019–20

Nonresponse bias statistics	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 adjustment¹							
Mean absolute difference across characteristics	0.18	‡	1.49	0.03	0.21	0.01	1.11
Median absolute difference across characteristics	0.10	‡	1.35	0.02	0.14	#	0.90
Difference between means for full sample and respondents after poststratification adjustment²							
Mean absolute difference across characteristics	1.05	‡	2.23	0.73	2.92	2.20	4.64
Median absolute difference across characteristics	0.66	‡	1.81	0.50	2.41	1.80	4.10

Rounds to zero.

‡ Reporting standards not met (fewer than five characteristics had at least five nonrespondents).

¹ Respondent means before poststratification adjustment are weighted using the institution base weight adjusted for nonresponse. Respondent means after poststratification adjustment are weighted using the institution base weight adjusted for nonresponse and poststratification.

² Full-sample means are weighted using the institution base weight, and respondent means are weighted using the institution base weight adjusted for nonresponse and poststratification.

NOTE: Characteristics that did not meet reporting standards were excluded from calculation of summary statistics. "Institution base weight" refers to the institution sampling weight.

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

7.2.3 Unit Nonresponse Bias Analysis: Student Level

Student-level nonresponse bias analyses were conducted for the student survey and study, overall and within each control and level of institution, separately for undergraduate students and graduate students. In addition, the study analyses were conducted by state and by public 2-year and public 4-year institution sectors within each state for undergraduate students. The results are summarized below. Respondent definitions are given in section 7.1.1.2.

The variables for the student nonresponse bias analysis, listed in table 92, included variables used in the student nonresponse adjustment (section 7.1.3.6) and additional variables not used in the weight adjustment model.

Table 92. Variables used in student-level nonresponse bias analyses: 2019–20

Variable	Survey undergraduate	Graduate	Study undergraduate
Control and level of institution	✓	✓	✓
State, control, and level of institution			✓
Region of institution	✓	✓	✓
Federal aid status	✓	✓	✓
Direct Loan status	✓	✓	✓
Total Direct Loan amount received	✓	✓	✓
Age as of December 31, 2019	✓	✓	✓
Institution aid status	✓	✓	✓
State aid status	✓	✓	✓
Veteran status	✓	✓	✓
Race/ethnicity	✓	✓	✓
Gender	✓	✓	✓
Institution total undergraduate or graduate enrollment	✓	✓	✓
Field of study or major	✓	✓	✓
Degree program	✓	✓	✓
Pell Grant status	✓		✓
Total Pell amount received	✓		✓
Received any aid excluding private loans	✓	✓	✓
First-time beginner status (sampled)	✓		✓
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	✓		✓
Average amount of grant and scholarship aid received	✓		✓
Student type (Doctor's degree–professional practice, other graduate)		✓	
Number of office and administrative support employees		✓	
Total income from CPS:20	✓	✓	
Parents' highest education from CPS:20	✓	✓	
Has dependents from CPS:20	✓	✓	
Student's marital status from CPS:20	✓	✓	

NOTE: CPS = Central Processing System. Continuous variables were converted to categorical variables using quartiles. Variables related to student population reflect the population of undergraduate students for analyses of undergraduate students and of graduate students for analyses of graduate students.

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

7.2.3.1 Nonresponse bias analysis for undergraduate students—Survey

As seen in table 93, there were many small but statistically significant biases on observable characteristics before the nonresponse weight adjustment, and these were largely diminished after weight adjustment in undergraduate students overall and across institution control and level. Before weight adjustment, the median percent relative bias across characteristics was 9.56 for undergraduate students overall, ranging from 4.32 to 15.66 across institution control and level. After weight adjustment, the median was 0.00 for undergraduate students overall and ranged from 1.65 to 9.97 across institution control and level.

The percentage of characteristics with significant bias before weight adjustment was 76.61 for undergraduate students overall and between 20.73 and 65.14 across institution control and level. After weight adjustment the percentage was 18.55 for undergraduate students overall and between 6.67 and 29.13 across institution control and level. The median effect size for undergraduate students before weight adjustment was 0.09 for students overall and between 0.04 and 0.15 across institution control and level, while after weight adjustment the median effect size was 0.00 for undergraduate students overall and 0.01 to 0.12 across institution control and level. The statistics summarized in table 93 are provided in appendix tables M-63 through M-74.

As seen in table 94, the mean absolute difference between respondent means before and after poststratification was 0.94 for undergraduate students overall and ranged from 0.71 to 8.78 across institution control and level. The median absolute difference was 0.47 for undergraduate students overall and ranged from 0.32 to 4.33 across institution control and level. The mean absolute difference between the full-sample and respondent means after poststratification was 1.12 for undergraduate students overall and ranged from 0.88 to 9.43 across institution and level. The median absolute difference was 0.64 for undergraduate students overall and ranged from 0.59 to 4.81 across institution and level. The statistics summarized in table 94 are provided in appendix tables M-75 through M-86.

Table 93. Summary of undergraduate student-level survey nonresponse bias analysis, by control and level of institution: 2019–20

Nonresponse bias statistics ¹	Overall	Public less- than- 2-year	Public 2-year	Public 4-year, non- doctorate- granting, primarily subbaccalaureate	Public 4-year, non- doctorate- granting, primarily baccalaureate	Public 4-year, doctorate- granting	Private nonprofit less-than- 4-year	Private nonprofit 4-year, non- doctorate- granting	Private nonprofit 4-year, doctorate- granting	Private for- profit less- than- 2-year	Private for- profit 2-year	Private for- profit 4-year
Before nonresponse weight adjustments²												
Mean percent relative bias across characteristics	11.25	13.90	13.58	13.95	8.18	9.36	19.92	7.49	10.10	13.99	17.09	8.88
Median percent relative bias across characteristics	9.56	10.11	11.53	11.69	4.58	5.33	15.66	4.32	7.74	9.32	10.17	6.35
Percentage of characteristics with significant bias	76.61	24.00	65.14	56.31	49.53	57.27	20.73	46.08	53.85	37.89	40.78	34.91
Median effect size	0.09	0.09	0.14	0.10	0.06	0.07	0.15	0.04	0.05	0.10	0.12	0.05
After nonresponse weight adjustments³												
Mean percent relative bias across characteristics	5.28	10.37	9.51	9.07	6.80	5.97	12.70	7.16	6.82	8.89	18.27	6.43
Median percent relative bias across characteristics	#	4.83	2.04	2.67	2.22	1.65	5.45	2.51	2.21	4.20	9.97	2.83
Percentage of characteristics with significant bias	18.55	6.67	22.94	14.56	24.30	16.36	7.32	13.73	22.12	10.53	29.13	15.09
Median effect size	#	0.05	0.01	0.03	0.02	0.01	0.07	0.02	0.02	0.04	0.12	0.02

Rounds to zero.

¹ Relative bias and effect size are calculated using the weighted differences between respondent and full-sample means. Relative bias is calculated as 100 times the ratio of estimated bias to the weighted full-sample mean. Effect size is calculated as the square root of the sum over categories of the squared differences over full-sample means.

² Respondent and full-sample means are weighted using the student base weight.

³ Full-sample means are weighted using the institution base weight, and respondent means are weighted using the student base weight adjusted for nonresponse.

NOTE: Variables and characteristics that did not meet reporting standards were excluded from calculation of summary statistics. "Student base weight" refers to the student sampling weight (final institution weight times student sampling adjustment) adjusted for student multiplicity and unknown eligibility.

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

Table 94. Summary of undergraduate student-level survey analysis after poststratification, by control and level of institution: 2019–20

Nonresponse bias statistics	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 adjustment¹												
Mean absolute difference across characteristics	0.94	2.44	0.71	0.78	0.83	1.29	8.78	1.45	1.41	3.63	8.16	3.91
Median absolute difference across characteristics	0.47	1.95	0.32	0.50	0.33	0.43	4.33	0.82	0.94	1.77	3.33	2.74
Difference between means for full sample and respondents after poststratification adjustment²												
Mean absolute difference across characteristics	1.12	2.86	1.22	0.99	0.88	1.51	9.43	1.36	1.58	3.93	6.78	4.12
Median absolute difference across characteristics	0.64	2.11	0.68	0.59	0.65	0.61	4.81	0.85	0.98	2.61	3.65	2.88

¹ Respondent means before poststratification adjustment are weighted using the student base weight adjusted for nonresponse. Respondent means after poststratification adjustment are weighted using the student base weight adjusted for nonresponse and poststratification.

² Full-sample means are weighted using the student base weight, and respondent means are weighted using the student base weight adjusted for nonresponse and poststratification.

NOTE: Characteristics that did not meet reporting standards were excluded from calculation of summary statistics. "Student base weight" refers to the student sampling weight (final institution weight times student sampling adjustment) adjusted for student multiplicity and unknown eligibility.

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

7.2.3.2 *Nonresponse bias analysis for undergraduate students—Study*

As seen in table 95, there were a moderate amount of small but statistically significant biases in observable characteristics before the nonresponse weight adjustment, and these were generally reduced after weight adjustment in undergraduate students overall and across institution control and level. The bias reduction results across institution and control were mixed because the nonresponse adjustment model did not include all levels of institution control. As described in section 7.1.3.6.3, the adjustment for undergraduate study nonresponse was conducted separately within each state and included, as a model variable, institution and control categorized by public 2-year, public 4-year, and “all other sectors.”

Before weight adjustment, the median percent relative bias across characteristics was 6.63 for undergraduate students overall, ranging from 1.91 to 12.69 across institution control and level. After weight adjustment, the median was 0.00 for undergraduate students overall and ranged from 1.20 to 9.56 across institution control and level. The percentage of characteristics with significant bias before weight adjustment was 36.64 for undergraduate students overall and between 0.00 and 50.55 across institution control and level. After weight adjustment the percentage was 14.66 for undergraduate students overall and between 2.50 and 38.46 across institution control and level. The median effect size for undergraduate students before weight adjustment was 0.05 for students overall and between 0.03 and 0.13 across institution control and level, while after weight adjustment the median effect size was 0.01 for undergraduate students overall and 0.02 to 0.13 across institution control and level. The statistics summarized in table 95 are provided in appendix tables M-87 through M-98.

As seen in table 96, the mean absolute difference between respondent means before and after poststratification was 0.46 for undergraduate students overall and ranged from 0.54 to 9.70 across institution control and level. The median absolute difference was 0.07 for undergraduate students overall and ranged from 0.23 to 4.48 across institution control and level. The mean absolute difference between the full-sample and respondent means after poststratification was 0.69 for undergraduate students overall and ranged from 1.40 to 12.30 across institution and level. The median absolute difference was 0.08 for undergraduate students overall and ranged from 0.65 to 7.69 across institution and level. The statistics summarized in table 96 are provided in appendix tables M-99 through M-110. Nonresponse bias analyses were also conducted by state and for 2-year public institutions and 4-year public institutions within each state, including the District of Columbia and Puerto Rico Summary statistics, similar to those provided in tables table 95 and table 96, are provided in appendix table M-111.

Table 95. Summary of undergraduate student-level study nonresponse bias analysis, by control and level of institution: 2019–20

Nonresponse bias statistics ¹	Overall	Public less- than- 2-year	Public 2-year	Public 4-year, non- doctorate- granting, primarily subbaccalaureate	Public 4-year, non- doctorate- granting, primarily baccalaureate	Public 4-year, doctorate- granting	Private nonprofit less-than- 4-year	Private nonprofit 4-year, non- doctorate- granting	Private nonprofit 4-year, doctorate- granting	Private for- profit less- than- 2-year	Private for- profit 2-year	Private for- profit 4-year
Before nonresponse weight adjustments²												
Mean percent relative bias across characteristics	12.89	12.51	15.81	12.27	10.39	8.87	22.85	10.84	8.83	20.04	14.95	9.14
Median percent relative bias across characteristics	6.63	1.91	7.61	5.08	4.76	3.05	12.58	5.72	2.85	12.69	8.74	3.99
Percentage of characteristics with significant bias	36.64	#	50.55	36.78	32.97	47.31	2.27	19.54	29.89	30.99	30.26	3.75
Median effect size	0.05	0.13	0.05	0.08	0.06	0.04	0.10	0.03	0.03	0.11	0.11	0.04
After nonresponse weight adjustments³												
Mean percent relative bias across characteristics	3.91	14.33	13.41	10.07	8.45	6.87	19.57	7.14	7.75	16.20	13.32	7.31
Median percent relative bias across characteristics	#	3.80	2.54	3.27	2.80	1.23	8.74	1.20	1.73	9.56	9.02	3.20
Percentage of characteristics with significant bias	14.66	15.56	38.46	19.54	21.98	19.35	15.91	21.84	29.89	32.39	25.00	2.50
Median effect size	0.01	0.13	0.05	0.07	0.03	0.02	0.08	0.02	0.03	0.08	0.08	0.02

Rounds to zero.

¹ Relative bias and effect size are calculated using the weighted differences between respondent and full-sample means. Relative bias is calculated as 100 times the ratio of estimated bias to the weighted full-sample mean. Effect size is calculated as the square root of the sum over categories of the squared differences over full-sample means.² Respondent and full-sample means are weighted using the student base weight.³ Full-sample means are weighted using the institution base weight, and respondent means are weighted using the student base weight adjusted for nonresponse.

NOTE: Variables and characteristics that did not meet reporting standards were excluded from calculation of summary statistics. "Student base weight" refers to the student sampling weight (final institution weight times student sampling adjustment) adjusted for student multiplicity and unknown eligibility.

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

Table 96. Summary of undergraduate student-level study analysis after poststratification, by control and level of institution: 2019–20

Nonresponse bias statistics	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 adjustment¹												
Mean absolute difference across characteristics	0.46	3.07	0.54	0.69	0.84	1.38	9.70	1.20	2.02	4.39	6.48	3.13
Median absolute difference across characteristics	0.07	2.32	0.23	0.46	0.31	0.43	4.48	0.52	1.25	2.48	3.36	1.48
Difference between means for full sample and respondents after poststratification adjustment²												
Mean absolute difference across characteristics	0.69	2.97	2.28	1.40	1.40	1.46	12.30	1.67	2.06	6.24	7.31	3.39
Median absolute difference across characteristics	0.08	2.16	0.74	0.85	0.71	0.65	7.69	0.82	1.31	4.81	4.60	1.88

¹ Respondent means before poststratification adjustment are weighted using the student base weight adjusted for nonresponse. Respondent means after poststratification adjustment are weighted using the student base weight adjusted for nonresponse and poststratification.

² Full-sample means are weighted using the student base weight, and respondent means are weighted using the student base weight adjusted for nonresponse and poststratification.

NOTE: Characteristics that did not meet reporting standards were excluded from calculation of summary statistics. "Student base weight" refers to the student sampling weight (final institution weight times student sampling adjustment) adjusted for student multiplicity and unknown eligibility.

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

7.2.3.3 *Nonresponse bias analysis for graduate students*

As seen in table 97, there was a moderate amount of small but statistically 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 institution control and level. Before weight adjustment, the median percent relative bias across characteristics was 6.46 for graduate students overall, ranging from 4.72 to 7.81 across institution control and level. After weight adjustment, the median was 0.00 and ranged from 2.42 to 6.20 across institution control and level. The percentage of characteristics with significant bias before weight adjustment was 50.00 for graduate students overall and between 18.18 and 46.03 across institution control and level. After weight adjustment, the percentage was 16.30 for graduate students overall and between 9.23 and 28.57 across institution control and level. The median effect size before weight adjustment for graduate students overall was 0.05 and between 0.04 and 0.08 across institution control and level. After weight adjustment, the effect size was 0.00 for graduate students overall and 0.02 to 0.05 across institution control and level. The statistics summarized in table 97 are provided in appendix tables M-112 to M-118.

Table 97. Summary of graduate student-level nonresponse bias analysis, by control and level of institution: 2019–20

Nonresponse bias statistics 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 adjustments²							
Mean percent relative bias across characteristics	10.78	‡	12.56	10.97	13.39	11.72	14.13
Median percent relative bias across characteristics	6.46	‡	6.57	5.98	7.81	7.06	4.72
Percentage of characteristics with significant bias	50.00	‡	30.77	42.68	46.03	45.00	18.18
Median effect size	0.05	‡	0.07	0.04	0.08	0.06	0.07
After nonresponse weight adjustments³							
Mean percent relative bias across characteristics	8.86	‡	10.58	9.94	12.06	9.92	11.92
Median percent relative bias across characteristics	#	‡	4.68	2.42	6.20	2.69	5.64
Percentage of characteristics with significant bias	16.30	‡	9.23	15.85	28.57	12.50	23.38
Median effect size	#	‡	0.03	0.02	0.05	0.02	0.05

Rounds to zero.

‡ Reporting standards not met (fewer than five characteristics had at least 30 nonrespondents).

¹ Relative bias and effect size are calculated using the weighted differences between respondent and full-sample means. Relative bias is calculated as 100 times the ratio of estimated bias to the weighted full-sample mean. Effect size is calculated as the square root of the sum over categories of the squared differences over full-sample means.

² Respondent and full-sample means are weighted using the student base weight.

³ Full-sample means are weighted using the institution base weight, and respondent means are weighted using the student base weight adjusted for nonresponse.

NOTE: Variables and characteristics that did not meet reporting standards were excluded from calculation of summary statistics. "Student base weight" refers to the student sampling weight (final institution weight times student sampling adjustment) adjusted for student multiplicity and unknown eligibility.

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

As seen in table 98, the mean absolute difference between respondent means before and after poststratification was 0.96 for graduate students overall and ranged from 1.25 to 3.07 across institution control and level. The median absolute difference was 0.44 for graduate students overall and ranged from 0.52 to 2.83 across institution control and level. The mean absolute difference between the full-sample and respondent means after poststratification was 1.50 for graduate students overall and ranged from 1.81 to 2.69 across institution and level. The median absolute difference was 0.64 for graduate students overall and ranged from 0.92 to 2.18 across institution and level. The statistics summarized in table 98 can be found in appendix tables M-119 through M-125.

Table 98. Summary of graduate student-level analysis after poststratification, by control and level of institution: 2019–20

Nonresponse bias statistics	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 adjustment¹							
Mean absolute difference across characteristics	0.96	‡	1.44	1.25	2.91	1.30	3.07
Median absolute difference across characteristics	0.44	‡	1.08	0.52	1.92	1.01	2.83
Difference between means for full sample and respondents after poststratification adjustment²							
Mean absolute difference across characteristics	1.50	‡	2.39	1.94	2.68	1.81	2.69
Median absolute difference across characteristics	0.64	‡	1.33	0.92	1.89	1.01	2.18

‡ Reporting standards not met (fewer than five characteristics had at least 30 nonrespondents).

¹ Respondent means before poststratification adjustment are weighted using the student base weight adjusted for nonresponse. Respondent means after poststratification adjustment are weighted using the student base weight adjusted for nonresponse and poststratification.

² Full-sample means are weighted using the student base weight, and respondent means are weighted using the student base weight adjusted for nonresponse and poststratification.

NOTE: Characteristics that did not meet reporting standards were excluded from calculation of summary statistics. “Student base weight” refers to the student sampling weight (final institution weight times student sampling adjustment) adjusted for student multiplicity and unknown eligibility.

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

7.2.4 Item Nonresponse Bias Analysis

Item nonresponse bias analyses were conducted for each item with a weighted response rate below 85 percent, overall and within each control and level of institution. The variables used are the same as those used in the student-level unit analyses (section 7.2.3). The results are summarized below.

Weighted item response rates were calculated for all items in the derived variable files, using the analysis weight (WTA000 or WTB000), and appear in appendix tables N-4 to N-7. In accordance with NCES statistical standard 1-3-5 (Seastrom 2014), item response rates (RRIs) were calculated as the ratio of the number of unit 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 treated as an eligible item nonrespondent. Similarly, respondents to the abbreviated questionnaire are considered item nonrespondents to all items not on the abbreviated questionnaire, so items on the abbreviated questionnaire tend to have higher response rates than items not on the questionnaire. To convert refusals, some institutions were allowed to provide only 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.⁶⁰

7.2.4.1 Item nonresponse bias analysis for undergraduate students—Survey

Appendix table O-4 contains weighted item response rates for 501 variables in the derived file for the undergraduate student survey. Response rates for students overall ranged from 22.0 percent for *Length of time studied abroad* (STABTIME) to 100 percent for 25 variables from administrative sources such as NSLDS. Another 151 variables had response rates over 95 percent. Across institution control and level, the lowest response rates observed were 1.5 percent for *Reason for not applying for financial aid* (REANOAPA-REANOAPF) for students in private for-profit 2-year institutions. These study abroad variables and financial aid variables are examples of items with high nonresponse due to both the abbreviated questionnaire and unknown eligibility. The questions were not on the abbreviated questionnaire, and the eligibility rate for these questions is relatively low (e.g., only students who studied abroad are eligible for STABTIME). Therefore, a large portion of nonrespondents were imputed to be ineligible for the questions.

There were 237 variables with overall response rates below 85 percent, and an additional 67 variables that had a response rate below 85 percent in one or more categories by institution and level, for a total of 304 variables for which nonresponse bias analyses were conducted. The results of these analyses are summarized in appendix table M-126. The items with the highest potential for bias due to nonresponse before imputation are those that apply to a small portion of the sample, in particular *Work-study job: located primarily on or off campus* (SJONOFF) and *Work-study job: related to major or coursework* (SJMAJOR). Although they had modest response rates of 72 percent, the median effect size was 0.78 and the percentage of estimated biases that were statistically significant

⁶⁰ 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.

was around 86 percent. Thirteen other variables had between 86 and 89 percent of characteristics with statistically significant bias and medium effect sizes. The differences between the preimputation and postimputation means were statistically significant for these variables, suggesting that the bias was reduced by imputation.

7.2.4.2 Item nonresponse bias analysis for undergraduate students— Study

Appendix table O-5 contains weighted item response rates for 47 variables on the derived file for the undergraduate student study. Response rates for students overall ranged from 53.9 percent for *Student budget minus EFC minus total aid* (SNEED2_AC) to 100.0 percent for *NPSAS institution state* (INSTSTAT). Across institution control and level, the lowest response rate observed was 36.9 percent for *Income percentile rank for full-time, full-year students* (PCTALLFTFY_AC) for students in public less-than-2-year institutions. Item response rates were also computed for the undergraduate study by state and by control and level of institution. Summary statistics for these response rates are provided in appendix table O-6.

There are 36 variables with overall response rates below 85 percent, and an additional 9 variables with response rates below 85 percent in at least one category of control and level of institution, for a total of 45 variables for which item nonresponse bias analysis was conducted. The results of these analyses are summarized in appendix table M-127. The items with the highest potential for nonresponse bias were income percentile variables such as *Income percentile rank for full-time full-year students* (PCTALLFTFY), *Income percentile for independent students* (PCTINDEP_AC), and *Income percentile for dependent students* (PCTDEP_AC). The largest median effect size observed for these variables was 0.37, and the largest percentage of characteristics with statistically significant bias was 47.6. The differences between the preimputation and postimputation means were statistically significant for these variables, suggesting that the bias was reduced by imputation.

7.2.4.3 Item nonresponse bias analysis for graduate students

Appendix table O-7 contains response rates for 434 variables on the derived file for the graduate student survey. Response rates for students overall ranged from 24.9 percent for *Grant amount exceeding federal need* (GRTSNEED) to 100 percent for 23 variables from administrative sources such as NSLDS. Another 147 variables had response rates over 95 percent. Across institution control and level,

the lowest response rate observed was 7.8 percent for MEALPLANFQ for students in private for profit 4-year institutions.

There are 121 variables with overall response rates below 85 percent, and there are an additional 119 variables with a response rate below 85 percent in at least one category of institution control and level, for a total of 240 variables for which item nonresponse bias was conducted. The results of these analyses are summarized in appendix table M-128. The items with the highest potential for bias due to nonresponse before imputation are those that apply to a small portion of the sample, in particular *Likelihood of using income-driven student loan repayment plans* (USEIDR), which had 54.3 percent of characteristics with significant bias and a median effect size of 0.47. Similar results were seen for *Likelihood of using loan forgiveness program* (USELFP): There were an additional 42 variables with higher percentages of characteristics with significant bias, up to 73.2 percent for *Tuition as percent of income* (INCPCT2), and median effect sizes between 0.11 and 0.42. The differences between the preimputation and postimputation means were statistically significant for 29 of these variables, suggesting that the bias was reduced by imputation.

7.3 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:20, 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 differ from those that would be expected if the sample was an SRS and the observations were independent and identically distributed random variables.

When testing hypotheses (conducting t tests, regression analyses, etc.) 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 the 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 7.3.1, and section 7.3.2 contains a discussion of the replicate weights created for the bootstrap procedure. Use of software packages for proper variance estimation is discussed in section 7.3.3.

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 an SRS 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 7.3.4 and appendix P.

7.3.1 Taylor Series

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

For stratified multistage 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:20, 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. In addition, separate analysis strata and PSUs were defined for undergraduate survey and study students to reflect the differences in the responding institutions. These sampling strata and PSUs 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. The PSUs were defined as the 2,160 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 2,150 undergraduate-enrolling study institutions, 2,120 undergraduate-enrolling survey institutions, and 1,410 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 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. Analysis PSUs were created separately for undergraduate survey students, undergraduate study students, and graduate students. Specifically, each analysis PSU for the undergraduate study students, undergraduate survey students, and graduates was created to contain at least four responding undergraduate study students, undergraduate survey students, and graduate students, respectively. Containing at least four respondents ensures stable variance estimates so that analyses could be conducted correctly on the separate undergraduate study, undergraduate survey, and graduate student analysis files. After the PSUs were combined, the resulting analysis PSUs were paired to form analysis strata. This process resulted in 1,002 analysis strata for undergraduate survey students, 1,052 analysis strata for undergraduate study students, and 427 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 (WTA_ANALSTR and WTB_ANALSTR) and analysis PSU (WTA_ANALPSU and WTB_ANALPSU). The set of variables that were used when assuming the first-stage units were sampled without replacement and that account for the FPC include the analysis stratum (WTA_FANALSTR and WTB_FANALSTR), analysis PSU (WTA_FANALPSU and WTB_FANALPSU), analysis SSU (WTA_FANALSSU and WTB_FANALSSU), and the weighted count of PSUs in an analysis stratum (WTA_PSUCOUNT and WTB_PSUCOUNT).

Ultimately,

- WTA_FANALSTR and WTB_FANALSTR equal the institution variance estimation stratum WTA_ANALSTR and WTB_ANALSTR, respectively; and
- WTA_FANALPSU and WTB_FANALPSU equal WTA_ANALPSU and WTB_ANALPSU, respectively.

WTA_FANALSSU and WTB_FANALSSU were created by randomly dividing the NPSAS:20 analysis PSUs into two parts.

These variables are by-products of the bootstrap variance estimation weights (described in section 7.3.2), 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.

7.3.2 Bootstrap Replicate Weights

A separate set of 200 bootstrap weights was constructed as was done for NPSAS:04, NPSAS:08, NPSAS:12, NPSAS:16, and NPSAS:18-AC. The final student weights (WTA000 and WTB000) described in section 7.1 were 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:

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

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

The replication method used was bootstrap with 200 replicates and with an FPC applied. The NPSAS:20 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 institution sampling strata (public 2-year, public 4-year, and “all other sectors”) 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 above; 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 ;
- \hat{N}_h = the frame count of institutions in stratum h ;
- m_{hi} = the number of SSUs or students selected from institution i in stratum h ;
- n_h^* = the bootstrap sample size of PSUs in stratum h when bootstrap sampling is at the PSU level in stratum h ;
- n_{hi}^* = the number of times PSU hi is selected in the bootstrap sample when bootstrap sampling is at the PSU level;
- m_{hi}^* = the bootstrap sample size of SSUs in PSU hi when bootstrap sampling is at the SSU level in stratum h ;
- m_{hij}^* = the number of times SSU 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),

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

2. Generate a uniform (0, 1) random number R_h for each stratum h .
3. If $R_h \leq FPC_h$, 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. Repeat steps 1 through 5 two hundred times for the undergraduate survey and study samples separately and 200 times for the graduate sample to form 200 replicate samples for each type of student (undergraduate survey students, undergraduate study students, and graduate students).

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, institutions were treated as being sampled with replacement or, equivalently, assumed to be sampled from a

large population of institutions. This assumption allowed the variance stratum finite population factor, FPC_h , to be treated as 1.00. The variance estimate was then based on institution variances only, and the within-institution component was fully represented in expectation. The exception was that institutions sampled with certainty were treated as strata; this set the finite population factor at the institution level to 0, and the variance estimate was based entirely on the within-institution variance. This did not account for nonresponse among certainty institutions.

A finite population factor less than 1.00 was applied at the institution stage of sampling and appropriately reduced the contribution of the institution component of variance but also reduced the contribution represented in expectation for the within-institution components of variance. It also allowed certainty institutions to be treated as a sample when there were nonrespondents. Taylor-series formulas added back a partial contribution from the within-institution 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.⁶¹ 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, the 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, institutions 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

⁶¹ 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.

adjustments described in section 7.1.4.2⁶² were applied to each replicate to create the 200 replicate weights included on each analysis file (WTA001–WTA200 and WTB001–WTB200) so that the variances would be estimated to account for the weight adjustments. For 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:20 to achieve model convergence for all replicates and to minimize the effect of different models on the variance estimates.⁶³

7.3.3 Software Use for Variance Estimation

Table 99 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 N. This example code, along with the code in table 99, can be helpful in writing code in other software packages.

⁶² The institution weight adjustments cannot be replicated due to the bootstrap methodology used.

⁶³ The sums of the replicate weights vary slightly due to the model adjustments, but replicate weights are only used for variance estimates and not for point estimates.

Table 99. 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:20 in selected survey data analysis software: 2019–20

Analysis weight for estimates	WTA000	WTB000
Taylor-series variance estimation (with replacement)		
Variance estimation stratum and PSU variables	WTA_ANALSTR and WTA_ANALPSU	WTB_ANALSTR and WTB_ANALPSU
Software: statements, parameters, and keywords for Taylor-series variance estimation (with replacement)		
SUDAAN	DESIGN = WR WEIGHT WTA000; NEST WTA_ANALSTR WTA_ANALPSU;	DESIGN = WR WEIGHT WTB000; NEST WTB_ANALSTR WTB_ANALPSU;
Stata	svyset WTA_ANALPSU [pweight = WTA000], strata (WTA_ANALSTR) vce(LINEARIZED)	svyset WTB_ANALPSU [pweight = WTB000], strata (WTB_ANALSTR) vce(LINEARIZED)
SAS survey data analysis procedures	VARMETHOD = TAYLOR WEIGHT WTA000; STRATA WTA_ANALSTR; CLUSTER WTA_ANALPSU;	VARMETHOD = TAYLOR WEIGHT WTB000; STRATA WTB_ANALSTR; CLUSTER WTB_ANALPSU;
IBM SPSS complex samples ¹	CSPLAN ANALYSIS /PLAN FILE='myfile.csaplan' /PLANVARS ANALYSISWEIGHT=WTA000 /DESIGN STRATA=WTA_ANALSTR CLUSTER=WTA_ANALPSU /ESTIMATOR TYPE=WR	CSPLAN ANALYSIS /PLAN FILE='myfile.csaplan' /PLANVARS ANALYSISWEIGHT=WTB000 /DESIGN STRATA=WTB_ANALSTR CLUSTER=WTB_ANALPSU /ESTIMATOR TYPE=WR
R survey package ²	mydesign<-svydesign(id=~WTA_ANALPSU, strata=~WTA_ANALSTR, weights=~WTA000, data=mydata)	mydesign<-svydesign(id=~WTB_ANALPSU, strata=~WTB_ANALSTR, weights=~WTB000, data=mydata)
Taylor-series variance estimation (without replacement)		
Variance estimation stratum, PSU, SSU, and count variables	WTA_FANALSTR, WTA_FANALPSU, WTA_FANALSSU, and WTA_PSUCOUNT	WTB_FANALSTR, WTB_FANALPSU, WTB_FANALSSU, and WTB_PSUCOUNT
Software: statements, parameters, and keywords for Taylor-series variance estimation (without replacement)		
SUDAAN	DESIGN = WOR WEIGHT WTA000; NEST WTA_FANALSTR WTA_FANALPSU WTA_FANALSSU; TOTCNT WTA_PSUCOUNT _minus1__zero_;	DESIGN = WOR WEIGHT WTB000; NEST WTB_FANALSTR WTB_FANALPSU WTB_FANALSSU; TOTCNT WTB_PSUCOUNT _minus1__zero_;
Stata	svyset WTA_FANALPSU [pw=WTA000], strata(WTA_FANALSTR) fpc(WTA_PSUCOUNT) WTA_FANALSSU, vce(LINEARIZED)	svyset WTB_FANALPSU [pw=WTB000], strata(WTB_FANALSTR) fpc(WTB_PSUCOUNT) WTB_FANALSSU, vce(LINEARIZED)
R survey package ²	mydesign<-svydesign(id=~WTA_FANALPSU, strata=~WTA_FANALSTR, weights=~WTA000, fpc=~WTA_PSUCOUNT, data=mydata)	mydesign<-svydesign(id=~WTB_FANALPSU, strata=~WTB_FANALSTR, weights=~WTB000, fpc=~WTB_PSUCOUNT, data=mydata)

See notes at end of table.

Table 99. 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:20 in selected survey data analysis software: 2019–20—Continued

Analysis weight for estimates	WTA000	WTB000
Bootstrap variance estimation		
Replicate weight variables	WTA001 – WTA200	WTB001 – WTB200
Software: statements, parameters, and keywords for BRR variance estimation		
SUDAAN	DESIGN = BRR WEIGHT WTA000; REPWGT WTA001 – WTA200;	DESIGN = BRR WEIGHT WTB000; REPWGT WTB001 – WTB200;
Stata	svyset [pweight=WTA000], brrweight(WTA001 - WTA200) vce(BRR) mse	svyset [pweight=WTB000], brrweight(WTB001 - WTB200) vce(BRR) mse
SAS survey data analysis procedures	VARMETHOD = BRR WEIGHT WTA000; REPWEIGHTS WTA001 – WTA200;	VARMETHOD = BRR WEIGHT WTB000; REPWEIGHTS WTB001 – WTB200;
WesVar	Method: BRR Full sample weight: WTA000 Replicates: WTA001 – WTA200	Method: BRR Full sample weight: WTB000 Replicates: WTB001 – WTB200
R survey package ²	mydesign<-svrepdesign(type="BRR", weights=~WTA000, repweights= "WTA00[1-200]", combined.weights=FALSE, data=mydata)	mydesign<-svrepdesign(type="BRR", weights=~WTB000, repweights= "WTB00[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: BRR = balanced repeated replication. 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, 2019–20 National Postsecondary Student Aid Study (NPSAS:20).

7.3.4 Variance Approximation

The design effect (DEFF) for a given estimate, $\hat{\theta}$, is defined as

$$\text{DEFF}(\hat{\theta}) = \frac{\text{Var}_{\text{design}}(\hat{\theta})}{\text{Var}_{\text{srs}}(\hat{\theta})},$$

where $\text{Var}_{\text{design}}(\hat{\theta})$ = the variance for a given estimate accounting for the study design and

$\text{Var}_{\text{srs}}(\hat{\theta})$ = the unadjusted variance for a given estimate for an SRS.

The square root of the design effect (DEFT) is another measure, which analysts can express as the ratio of the standard errors, or

$$\text{DEFT}(\hat{\theta}) = \frac{\text{SE}_{\text{design}}(\hat{\theta})}{\text{SE}_{\text{srs}}(\hat{\theta})},$$

where

$\text{SE}_{\text{design}}(\hat{\theta})$ = the standard error for a given estimate accounting for the study design and

$\text{SE}_{\text{srs}}(\hat{\theta})$ = the standard error for a given estimate for SRS.

Most complex multistage sampling designs like NPSAS:20 result in design effects greater than 1.0; that is, the design-based variance is larger than the SRS variance. Appendix P 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 above and in appendix N.

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 and to the different sampling rates between student strata.

As discussed above, Taylor-series linearization and replication techniques can be used to compute more precise standard errors for data from complex 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:20 data without using one of the software packages for analysis of complex sample study data can use the design effect tables in appendix P 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 99 in the previous section and appendix M.)

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

$$\text{normalized weight} = \text{weight} * n/N,$$

where n is the sample size (i.e., the number of cases with a valid main sampling weight) and N is the sum of weights.

As the second step in the approximation, the standard errors produced by the statistical software, the test statistics, or the sample weight used in analysis can be adjusted to reflect the actual complex design of the study. To adjust the standard error of an estimate, the analyst should multiply the standard error produced by the statistical software by the square root of the *DEFF*, or the *DEFT*. The *DEFF* and *DEFT* used to make adjustments can be calculated for specific estimates, can be the median *DEFF* and *DEFT* across a number of variables, or can be the median *DEFF* and *DEFT* for a specific subgroup in the population. Adjusted standard errors can then be used in hypothesis testing, for example, when calculating t and F statistics.

A second option is to adjust the t and F statistics produced by statistical software packages using unadjusted (i.e., SRS) standard errors. To do this, first conduct the desired analysis weighted by the normalized weight and then divide a t statistic by the *DEFT* or divide an F statistic by the *DEFF*. A third alternative is to create a

new analytic weight variable in the data file by dividing the normalized analytic weight by the *DEFF* and using the adjusted weight in the analyses.

7.4 Imputation

Missing data, with the exception of missing code -3 (see section 6.2), 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),⁶⁴ 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

⁶⁴ 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.

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:20, there were five iterations, which improved quality without significantly slowing down the imputation process.

Similar to NPSAS:18-AC, because dependent and independent undergraduate financial aid may differ according to state law, state financial aid variables were imputed separately by state and dependency. Graduate students were imputed as a single group.

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 to 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, the distributions within imputation classes were verified to be similar for the observed and imputed data and the completed (observed and imputed) distribution for income was concluded to be reasonable. Item response rates are shown in section 7.2.4, and the observed and imputed distributions for eight key variables are provided in appendix O, tables O-1 through O-3.

7.5 Disclosure Risk Analysis and Avoidance

In preparing data files for release, NCES takes steps, including a formal disclosure risk analysis, to minimize the likelihood that individual students participating in the study can be identified. Every effort is made to protect the confidentiality of information about specific individuals, including performing data-swapping procedures on NPSAS:20 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:20 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

- American Association for Public Opinion Research. (2016). *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys* (9th ed.) Alexandria, VA: Author. Retrieved June 7, 2023, from <https://aapor.org/wp-content/uploads/2022/11/Standard-Definitions20169theditionfinal.pdf>.
- Box, G.E.P., and Cox, D.R. (1964). An Analysis of Transformations. *Journal of the Royal Statistics Society, Series B*, 26(2): 211–234.
- 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.
- Cominole, M., Siegel, P., Dudley, K., Roe, D., and Gilligan, T. (2006). *2004 National Postsecondary Student Aid Study (NPSAS:04) Full Scale Methodology Report* (NCES 2006–180). U.S. Department of Education. Washington, DC: National Center for Education Statistics.
- 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, 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 Survey Research Methods Section, American Statistical Association* (pp. 792–796). Alexandria, VA: American Statistical Association.

- Hanley, J.A., and McNeil, B.J. (1982). The Meaning and Use of the Area Under a Receiver Operating Characteristic (ROC) Curve. *Radiology*, 143(1): 29–36. Retrieved June 7, 2023, from <https://doi.org/10.1148/radiology.143.1.7063747>.
- Iannacchione, V. (1982, February). *Weighted Sequential Hot Deck Imputation Macros*. Paper presented at the Seventh Annual SAS Users Group International Conference, San Francisco.
- Iannacchione, V. (2003). Sequential Weight Adjustments for Location and Cooperation Propensity for the 1995 National Survey of Family Growth. *Journal of Official Statistics*, 16: 31–43.
- Judkins, D.R. (1997). Imputing for Swiss Cheese Pattern of Missing Data. In *Proceedings of the Statistics Canada Symposium 97, New Directions in Surveys and Censuses* (pp. 143–148). 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: 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* (pp. 329–341). New York: John Wiley & Sons, Inc.
- Office of Postsecondary Education. (2020). *Guidance for Interruptions of Study Related to Coronavirus (COVID-19)*. Washington, DC: Federal Student Aid. Retrieved February 5, 2021, from <https://fsapartners.ed.gov/knowledge-center/library/electronic-announcements/2020-03-05/guidance-interruptions-study-related-coronavirus-covid-19-updated-june-16-2020>.

- 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.
- Satterthwaite, F.E. (1946). An Approximate Distribution of Estimates of Variance Components. *Biometrics Bulletin*, 2(6): 110–114.
- Seastrom, M. (2014). *NCES Statistical Standards* (NCES 2014-097). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved June 7, 2023, from <https://nces.ed.gov/statprog/2012/>.
- Tukey, J.W. (1977). *Exploratory Data Analysis*. Reading, MA: Addison-Wesley.
- Wine, J., Bryan, M., and Siegel, P. (2014). *2011–12 National Postsecondary Student Aid Study (NPSAS:12) Data File Documentation* (NCES 2014-182). U.S. Department of Education. Washington, DC: National Center for Education Statistics.
- Woodruff, R.S. (1971). A Simple Method for Approximating the Variance of a Complicated Estimate. *Journal of the American Statistical Association*, 66(344): 411–414.