NATIONAL CENTER FOR EDUCATION STATISTICS
Institute of Education Sciences

2007–08 National Postsecondary Student Aid Study (NPSAS:08) Full-scale Methodology Report

Technical Report

NCES 2011-188

U.S. DEPARTMENT OF EDUCATION

This page intentionally left blank.

2007–08 National Postsecondary Student Aid Study (NPSAS:08) Full-scale Methodology Report

Technical Report

December 2010

Melissa Cominole
John Riccobono
Peter Siegel
Lesa Caves
RTI International

Tracy Hunt-White
Project Officer
National Center for Education Statistics

NCES 2011-188

U.S. DEPARTMENT OF EDUCATION

U.S. Department of Education
Arne Duncan
Secretary

Institute of Education Sciences
John Easton
Director

National Center for Education Statistics
Stuart Kerachsky
Acting 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, and other education policymakers, practitioners, data users, and the general public.

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
1990 K Street NW
Washington, DC 20006-5651

November 2010

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

This publication is only available online. To download, view, and print the report as a PDF file, go to the NCES World Wide Web 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
Cominole, M., Riccobono, J., Siegel, P., and Caves, L. (2010). *2007–08 National Postsecondary Student Aid Study (NPSAS:08) Full-scale Methodology Report* (NCES 2011-188). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved [date] from http://nces.ed.gov/pubsearch.

Content Contact
Aurora D'Amico
(202) 502–7334
aurora.d'amico@ed.gov
(202) 502-7334

Executive Summary

Introduction

The 2007–08 National Postsecondary Student Aid Study (NPSAS:08), conducted for the U.S. Department of Education's National Center for Education Statistics (NCES), has collected comprehensive data on how students and their families pay for postsecondary education. The first NPSAS study was conducted in 1986–87 to meet the need for national-level data about significant financial aid issues. Since 1987, NPSAS has been conducted every 3 to 4 years, the last cycle, NPSAS:04, having been conducted during the 2003–04 academic year. The primary objective of NPSAS:08 is to produce reliable national estimates of characteristics related to financial aid for postsecondary students at both the undergraduate and graduate levels.

NPSAS:08 served as the base year of data collection for the Baccalaureate and Beyond Longitudinal Study (B&B). The B&B series provides a longitudinal study of the education, work, financial, and personal experiences of students who have completed a bachelor's degree at a given point in time. To be eligible for the B&B:08 cohort, a student must have met the requirements to complete a bachelor's degree between July 1, 2007, and June 30, 2008. A cohort of baccalaureate recipients was identified as part of NPSAS:08, and the B&B study will collect further data from them in 1 year (2009) and again in 4 years (2012).

A field test study was conducted for the purpose of testing the procedures planned for the full-scale NPSAS:08 study. NPSAS:08 included a new set of interview items to obtain baseline measures of the awareness of two new federal grants—the Academic Competitiveness Grant (ACG) and the National Science and Mathematics Access to Retain Talent (SMART) Grant—introduced in 2006. In addition, NPSAS:08 introduced the use of the Virtual Call Center. This supplement to the call center allowed interviewers to conduct telephone interviews from their homes, while it increased the telephone interviewing capacity for NPSAS:08.

The NPSAS:08 sample was designed to provide state-level representative estimates for undergraduate students within four degree-granting institutional sectors in six states: (1) public 2-year institutions; (2) public 4-year institutions; (3) private, not-for-profit 4-year institutions; and (4) private, for-profit institutions. The six states included in this sample were California, Georgia, Illinois, New York, Minnesota, and Texas.

This report details the methodological procedures and results from NPSAS:08. The full-scale study was conducted by RTI International (RTI), with assistance from MPR Associates, Inc. (MPR), for NCES of the U.S. Department of Education, Washington, DC, as authorized by the Higher Education Act of 2008, 20 U.S.C. § 1015 (2008).

Sample Design

The full-scale NPSAS:08 was based on a sample of all students enrolled in eligible postsecondary institutions throughout the United States and Puerto Rico during the 2007–08

academic year. [1] The institutional sample drew from all sectors—public, private not-for-profit, and private for-profit institutions at the 4-year, 2-year, and less-than-2-year levels. Additionally, the sample included state-representative samples for six states. To prevent an institution from being selected into both the field test and full-scale samples, both samples were selected at the same time, before the field test study. For the institutional sample, 1,960 institutions were selected.[2] Of these institutions, 1,940 were found to be eligible for NPSAS:08. Institutions were asked to provide enrollment lists for the purpose of selecting the student sample. Enrollment lists were obtained from 1,730 of the eligible institutions (90 percent, weighted).

The student sample was selected on a flow basis as the lists were received from institutions. Of the 137,800 students initially sampled, 132,800 were determined to be eligible for participation in NPSAS:08. The eligible sample included 84,240 (96 percent) sample members from public institutions; 16,610 (95 percent) from private, for-profit institutions; and 31,950 (97 percent) from private, not-for-profit institutions. On the basis of information from responding institutions provided by their enrollment lists, approximately 89 percent (118,610) of the eligible sample were classified as undergraduates; 11 percent (14,190) were classified as graduate or first-professional students (on the basis of enrollment list classification).[3]

Instrumentation

Computer-Assisted Data Entry Instrument

A web-based computer-assisted data entry (CADE) software system was used for the abstraction, or collection, of student records from institutions. The CADE system consisted of eight sections: contact information, tuition, need analysis, enrollment, undergraduate admissions, demographics, financial aid awarded, and institutional student information records. Before institutions were allowed to enter administrative record data for study respondents, records for all students sampled from that institution were requested from the U.S. Department of Education's Central Processing System (CPS), which maintains financial aid application data. This information was preloaded into the CADE system to provide edit checks for the data entered by the institution.

Student Interview

The NPSAS:08 full-scale student interview was designed as a web-based instrument to be used for both web and telephone respondents. This design required that several important features be embedded in the instrument, such as extensive help text to assist respondents and warnings to alert them when a response fell outside a predetermined range of likely responses.

The interview consisted of seven sections grouped by topic. The first section determined student eligibility for the NPSAS:08 study, gathered information needed to identify respondents who might be eligible for further follow-up in B&B, and obtained enrollment information for the 2007–

[1] To be considered eligible to participate in NPSAS:08, an institutions had to have met the following qualifications: offered an educational program designed for persons who had completed secondary education; offered at least one academic, occupational, or vocational program of study lasting at least 3 months or 300 clock hours; offered courses open to persons other than the employees or members of the company or group (e.g., union) that administered the institution; been located in the 50 states, the District of Columbia, or Puerto Rico; been other than a U.S. Service Academy; and signed a Title IV participation agreement with the U.S. Department of Education.

[2] The numbers appearing in the tables and text of this report have been rounded to the nearest tens to maintain the confidentiality of study respondents; however, percentages are based on unrounded numbers.

[3] A first-professional student is a student who is enrolled in one of the following degree programs: chiropractic, dentistry, law, medicine, optometry, osteopathic medicine, pharmacy, podiatry, ministry or divinity, or veterinary medicine.

08 school year. The second section elicited information about educational experiences, such as languages studied, distance-education courses, and transfer credit history. The third section asked about financial aid and student expenses. Included in this section were items about aid-based employment at the NPSAS institution, such as work-study jobs, assistantships, and fellowships, as well as the new ACG and SMART Grant items. The fourth section focused on non-aid-based employment and personal finances. The fifth section gathered background and demographic information about sample members and their families. The sixth section, administered only to potentially eligible B&B students, addressed future plans in terms of postgraduation employment, education, and teaching. The final section requested contact information from potentially eligible B&B sample members for the follow-up longitudinal studies.

Data Collection Design and Outcomes

Training

Training programs were developed for different types of project staff: institutional coordinators, help desk agents, telephone interviewers, tracers, and supervisors. All interviewing staff received a general training on telephone interviewing techniques. Help desk agents received specific training on frequently asked questions about the study, instrument, and technical issues related to completion of the instrument on the Web. Help desk agents were also trained how to conduct the student interview when requested to do so by sample members. Programs on successfully locating and interviewing sample members were developed for all telephone interviewers. Topics covered in telephone interviewer training included administrative procedures required for case management; quality control of interactions with sample members, parents, and other contacts; the purpose of NPSAS:08 and the uses of the data to be collected; and the organization and operation of the web instrument to be used in data collection. Tracers received specialty training on how to use specific locating databases and tools to find hard-to-reach sample members. Supervisors received the same training that help desk agents and telephone interviewers received, together with additional sessions related to project administration and supervision.

Institutional Contacting

Once institutions were sampled, attempts were made to contact the chief administrator of each institution (e.g., the president or the chancellor) to verify institutional eligibility, solicit participation, and request the appointment of an institutional coordinator to oversee data collection within the institution. Institutional coordinators were asked to provide lists of all eligible students enrolled at any time between July 1, 2007, and June 30, 2008.[4] Several quality checks of the lists were implemented before students from each institution were sampled. Of the 1,960 eligible institutions sampled, 1,730 provided lists, which resulted in an overall institutional participation rate of 90 percent (weighted).

[4] In order to be considered eligible to participate in NPSAS:08, a student had to have attended a NPSAS-eligible institution during the 2007–08 academic year and to have met these further requirements: enrolled in either (1) an academic program or at least one course for credit that could be applied toward fulfilling the requirements for an academic degree, or (2) an occupational or vocational program that required at least 3 months or 300 clock hours of instruction to receive a degree, certificate, or other formal award; not been concurrently enrolled in high school; and not been enrolled solely in a General Educational Development (GED) or other high school completion program.

Institutional Record Abstraction

Institutions could choose from among three modes for student record abstraction: (1) *self-CADE,* by which institutional staff entered data directly into the web-based CADE system; (2) *data-CADE,* by which institutional staff provided student record information in data file uploads according to specifications; and (3) *field-CADE,* by which trained field data collectors used laptops to abstract student record data into the CADE system. Nearly two-thirds of institutions (63 percent) chose to submit student data by self-CADE. About 36 percent of institutions submitted student records by data-CADE. Approximately 1 percent of institutions submitted their data by field-CADE. Of the 1,940 eligible and participating institutions, about 1,670 institutions provided institutional record information for about 130,410 students. Approximately 126,620 of the CADE student records received were considered complete student records. The institution- and student-level weighted response rates for record abstraction were both 96 percent.

Student Locating and Interviewing

The NPSAS:08 data collection design involved initial locating of sample members, providing an opportunity for the student to complete the web interview, following up with web nonrespondents after 3 weeks, and attempting to conduct a telephone interview with them, if necessary. Upon receipt of student enrollment lists, batch-locating activities were implemented to update address and telephone information. Sources for this task included the CPS, the U.S. Postal Service National Change of Address system, and Telematch (after batch locating). Students were sent a notification mailing containing an initial letter, informational brochure, and username and password for completing the interview via the Web. Telephone contacting began for web nonrespondents 3 weeks after the initial mailing.

Telephone interviewers' locating and tracing activities coincided with efforts to gain cooperation from sample members. Any case for which the interviewer had exhausted all tracing options and was unable to gain any contact with the sample member was sent to RTI's Tracing Operations Unit (TOPS). Telephone interviewers followed up with cases for which TOPS had obtained further contact information. Sample members for whom no additional information could be obtained were determined to be final unlocatable sample members.

Student data collection consisted of three phases. The first, the early response phase, paid a $30 incentive to sample members who completed the web interview or called in to complete a telephone interview during the first 3 weeks of data collection. Approximately 48 percent of completed interviews were obtained during this phase. In the second phase of data collection, production interviewing, interviewers made outbound calls to complete telephone interviews. No incentives were offered during production interviewing, which accounted for 11 percent of all completed interviews. The third phase of data collection, the nonresponse-conversion phase, focused on obtaining interviews from sample members who had refused to participate, were hard to reach, or had insisted on receiving the incentive before completing the interview. Sample members who completed an interview during this phase of data collection received an incentive of $30. About 41 percent of completed interviews were completed during the nonresponse-conversion phase.

Of the approximately 132,800 eligible sample members, 95,360 (71 percent, weighted) completed the student interview. Of these sample members, 18,010 were identified as baccalaureate recipients during the NPSAS year and therefore eligible for inclusion in the B&B:08 cohort. On average, students spent about 26 minutes completing the interview. Web respondents were able to

complete the interview in approximately 25 minutes, while telephone respondents completed the interview in about 29 minutes.

Study Respondents

Key variables were identified across the various data sources to determine the minimum requirements to support the analytic needs of the study. Sample members who met these minimum requirements were classified as *study respondents*. To be a study respondent, a sample member had to have data for three critical variables—student type (undergraduate, graduate, or first-professional student), gender, and date of birth. In addition, sample members had to have data—from any source—for at least 8 variables out of a set of 15 additional prespecified variables. These variables include dependency status, income, expected family contribution (EFC), class level, and race. About 127,700 of 132,800 eligible sample members (96 percent) were classified as NPSAS:08 study respondents. For most study respondents (approximately 88 percent), data were available from at least two of the major data sources (e.g., student interview, institutional student records, or the CPS).

Evaluation of Operations and Data Quality

Evaluations of operations and procedures focused on the timeline for data collection from both institutions (CADE) and students (web and telephone interviews), tracing and locating procedures, refusal-conversion efforts, the effectiveness of incentives for increasing early response via the Web and for refusal conversion, and the length of the student interview. Evaluations of data quality included an examination of items with high rates of missing data, the accuracy of data collected in coding systems, use of online help text, item-level nonresponse-conversion efforts, and question delivery and data-entry quality control procedures.

Analysis Weights

Cross-sectional weights were developed for student-level analyses. Analysis strata and PSUs were created to allow for variance estimation using the Taylor series procedure, and bootstrap replicate weights were computed to allow for variance estimation using the bootstrap technique. Weighted response rates, nonresponse bias analyses, and survey design effect tables are reported.

Data Files

The final data files are available as a set of restricted research files, which are fully documented by an electronic codebook (ECB), and as a public release Data Analysis System (DAS), which also contains full documentation.[5] The first NPSAS:08 DAS was adjudicated and approved for public release in April 2009. The primary analysis file contains data for 127,700 study respondents and more than 600 variables developed from multiple sources.

[5] The ECB and DAS are both fully documented software products available from NCES. The DAS is available online at http://nces.ed.gov/das/.

Products

- *The Profile of Undergraduates in U.S. Postsecondary Education Institutions: 2007–08* will describe the demographic and enrollment characteristics of undergraduate students.
- *Student Financing of Undergraduate Education: 2007–08* will address undergraduate tuition, total price of attendance, types and sources of financial aid received, net price, financial aid need, and unmet need.
- *Student Financing of Graduate and First-Professional Education: 2007–08* will describe the demographic and enrollment characteristics of graduate and first-professional students, as well as the types and sources of financial aid received.
- *2007–08 National Postsecondary Student Aid Study (NPSAS:08): Student Financial Aid Estimates for 2007–08* report is a First Look publication based on the 2007–08 data; it gives the percentages of students receiving various types of financial aid and average amounts received, by type of institution attended, attendance pattern, dependency status, and income level.
- *2007–08 National Postsecondary Student Aid Study (NPSAS:08):Undergraduate Financial Aid Estimates for 2007–08 by Type of Institution* report is the second publication based on the 2007–08 data. This report focuses only on undergraduates and includes separate tables for those who attended public 4-year institutions; those who attended private, not-for-profit 4-year institutions; those who attended public 2-year institutions; and those who attended private, for-profit postsecondary institutions during the 2007–08 academic year. It gives average tuition and fees, average total price of attendance, and the percentages of undergraduates receiving various types and combinations of financial aid, together with average amounts received and with a particular focus on grants and loans.

Foreword

This report describes and evaluates the methods and procedures used in the 2007–08 National Postsecondary Student Aid Study (NPSAS:08). NPSAS:08 included important changes from previous NPSAS cycles. The student sample was augmented to oversample students who had received Academic Competitiveness Grants and SMART Grants. This oversampling permitted inclusion of several new interview items, allowing policymakers to gauge student knowledge about these new grant programs. As did NPSAS:04, NPSAS:08 oversampled institutions in selected states to provide state-level representative data on undergraduate students in four institutional sectors.

We hope that the information provided in this report will be useful to a wide range of interested readers. Additional information about NPSAS:08 is available on the Web at http://www.nces.ed.gov/surveys/npsas/.

Tom Weko
Associate Commissioner
Postsecondary Studies Division

This page intentionally left blank.

Acknowledgments

The authors would like to acknowledge the many individuals who participated in the 2007–08 National Postsecondary Student Aid Study (NPSAS:08). Thanks to all the institutions that participated and, especially, the students who were interviewed by phone and via the Web. This study would have been impossible without their invaluable contributions.

This page intentionally left blank.

Contents

Executive Summary iii

Foreword ix

Acknowledgments xi

List of Tables xvii

List of Figures xxi

Chapter 1. Overview 1

1.1 Background and Purpose of NPSAS 1
1.2 Special Features of NPSAS:08 2
1.2.1 State Representative Sample 2
1.2.2 The Academic Competitiveness Grant and the Science and Mathematics Access to Retain Talent Grant 2
1.2.3 Virtual Call Center 2
1.3 Schedule and Products of NPSAS:08 3

Chapter 2. Design and Methodology 5

2.1 Sampling 5
2.1.1 Target Population and Sampling Overview 5
2.1.2 Institutional Sample and Eligibility 14
2.1.3 Student Sample and Eligibility 18
2.2 Sources of Data 22
2.3 Instrument Development 22
2.3.1 Computer-Assisted Data Entry 22
2.3.2 Student Interview 24
2.3.3 Evaluation of Data Quality 27
2.4 Data Collection 28
2.4.1 Contacting Institutions 28
2.4.2 Student Enrollment List Acquisition 32
2.4.3 Administrative Data for Sample Members 34
2.4.4 Data Abstraction From Student Records 35
2.4.5 Student Interviewing 38
2.4.6 Phases of Student Interviewing 42
2.4.7 Locating and Contacting Students 43
2.5 Systems 46
2.5.1 Hatteras and the Instrument Development and Documentation System 46
2.5.2 Integrated Management System 47
2.5.3 Variable Tracking System 48

Chapter 3. Data Collection Outcomes 49

3.1 Institutional Participation 49
3.1.1 National Institutional Sample 49
3.1.2 State Institutional Samples 50
3.2 Study Respondents 51

3.2.1 National Student Sample 52
3.2.2 State Student Samples 53
3.3 Data Collection Results 55
3.3.1 Student Record Matching 55
3.3.2 Student Record Abstraction 61
3.3.3 Student Locating and Response Rates 64
3.4 Student Locating 65
3.4.1 Database Batch Tracing Before Data Collection 66
3.4.2 Intensive Tracing During Data Collection 67
3.5 Interviewing Outcomes 69
3.5.1 Interview Completion by Data Collection Phase 69
3.5.2 Interview Completion by Mode of Administration 72
3.6 Completeness of Data Records Among Study Respondents 75

Chapter 4. Evaluation of Field Operations and Data Quality 77

4.1 Enrollment List Collection 77
4.1.1 Early Contacting Activities 77
4.1.2 Institutional Participation 78
4.1.3 Quality of Enrollment Lists 80
4.2 Student Record Abstraction 82
4.2.1 Preloading Data Into the Computer-Assisted Data Entry System 82
4.2.2 Timeliness of Record Abstraction 83
4.2.3 Computer-Assisted Data Entry Data Completeness 85
4.3 Student Interviewing 86
4.3.1 Timing Analysis 86
4.3.2 Call Counts 90
4.3.3 Online Coding 92
4.3.4 Help Text 94
4.3.5 Item Nonresponse 96
4.3.6 Conversion Text 98
4.3.7 Verification Checks 99
4.3.8 Help Desk 99
4.3.9 Telephone Interviews and Quality Assurance Monitoring 100
4.3.10 Virtual Call Center 103
4.3.11 Respondent Profiles 104
4.4 Identification of the Baccalaureate and Beyond Cohort 108

Chapter 5. Variable Construction and File Development 111

5.1 Overview of the NPSAS:08 Data Files 111
5.2 Online Coding and Verification 112
5.2.1 Online Coding 112
5.2.2 Range and Consistency Checks 113
5.3 Data Editing 113
5.4 Data Perturbation 115
5.5 Statistical Imputations 115
5.6 Composite and Derived Variable Construction 116

Chapter 6. Unit Nonresponse Bias Analyses, Weighting, and Variance Estimation.....119

6.1 Institution Nonresponse Bias Analysis and Weighting..... 121
6.1.1 Initial Institution Weight Components.....121
6.1.2 Assessing Institution Nonresponse Bias.....122
6.1.3 Adjusting Institution Weights.....124
6.1.4 Institution Weighting Adjustment Performance 131
6.2 Student Nonresponse Bias Analysis and Weighting..... 132
6.2.1 Initial Student Weight Components.....132
6.2.2 Assessing Student Nonresponse Bias.....136
6.2.3 Adjusting Student Weights 137
6.2.4 Student Weighting Adjustment Performance 158
6.3 Item Nonresponse Bias Analysis..... 159
6.4 Variance Estimation 174
6.4.1 Taylor Series.....175
6.4.2 Bootstrap Replicate Weights 175
6.4.3 Variance Approximation.....178

References..... 179

Appendixes

A Technical Review Panel..... A-1
B Sampling Details.....B-1
C Facsimiles.....C-1
D NPSAS:08 Endorsements D-1
E Data Collection Materials.....E-1
F DataCADE Specifications G-1
G Training Materials.....F-1
H Imputations H-1
I Analysis Variables.....I-1
J Generalized Exponential Model Adjustment Procedure.....J-1
K Nonresponse Bias Analysis.....K-1
L Weighted-Analysis Considerations L-1
M Design Effects.....M-1

This page intentionally left blank.

List of Tables

1. Start and end dates for major NPSAS:08 activities 3
2. Target numbers of sample students, by institutional stratum and student type: 2008 12
3. Target numbers of sample students in the six state-representative samples, by institutional stratum and type of student: 2008 13
4. Institutional sampling rates and number of certainty and noncertainty institutions sampled, by institutional stratum: 2008 17
5. Planned and achieved NPSAS:08 student samples, by student stratum and institution level: 2008 20
6. Initial classification of NPSAS:08 student sample, by institutional characteristics and student type: 2008 21
7. Summary of NPSAS:08 evaluations: 2008 28
8. NPSAS:08 training sessions for interviewing and tracing staff: 2008 38
9. Numbers of sampled, eligible, and participating institutions and enrollment list participation rates, by institutional characteristic: National sample: 2008 50
10. Numbers of sampled, eligible, and participating institutions and enrollment list participation rates, by institutional characteristic: State samples: 2008 51
11. Numbers of sampled and eligible students and response rates, by institutional characteristic: National sample: 2008 53
12. Numbers of sampled, eligible undergraduate students and response rates, by institutional characteristic: State samples: 2008 54
13. Results of Central Processing System matching for 2007–08 and 2008–09, by institutional characteristic and student type: 2008 57
14. Results of National Student Loan Data System matching, by institutional characteristic and student type: 2008 58
15. Results of National Student Clearinghouse matching, by institutional characteristic and student type: 2008 59
16. Results of ACT and SAT Reasoning Test matching, by institutional characteristic and student type: 2008 60
17. Student record abstraction method: 2008 61
18. Institution-level computer-assisted data entry completion rates, by institutional characteristic and abstraction method: 2008 62
19. Student-level computer-assisted data entry completion rates, by institutional characteristic and student type: 2008 63
20. Student locating, by institutional characteristic and student type: 2008 66
21. Batch processing record match rates, by tracing source: 2008 67

22. Students requiring intensive tracing, by institutional characteristic and student type: 2008 68
23. Locate and interview rates, by intensive tracing efforts: 2008 68
24. Student interview completion results, by institutional characteristic and student type: 2008 69
25. Student interview response rates, by data collection phase, and by institution characteristic and student type: 2008 70
26. Results of prompting during the early response phase of data collection: 2008 72
27. Student interview completion rates: 2008 73
28. Student interview completion rates by mode of completion, and by institutional characteristic and student type: 2008 74
29. Data completeness for NPSAS:08 study respondents, by data source: 2008 75
30. Cumulative flow of enrollment list receipt: 2004 and 2008 77
31. NPSAS:08 institutional enrollment list participation, by prior NPSAS participation: 2008 78
32. Distribution of participating NPSAS:08 institutions, by 2005 Carnegie institutional classification: 2008 79
33. NPSAS:08 participation of Historically Black Colleges and Universities: 2008 80
34. Types of discrepancies encountered with student enrollment lists, by highest level of offering: 2008 81
35. Source of data elements preloaded into the computer-assisted data entry system: 2008 83
36. Computer-assisted data entry item-level completion rates, by method of abstraction: 2008 86
37. Average interview time, by mode of administration: 2008 87
38. Average time to complete the NPSAS student interview, by interview section and completion mode: 2008 88
39. Average time to complete the education experiences section, by student type: 2008 88
40. Average time to complete the current employment section, by employment status: 2008 89
41. Average time to complete the web student interview, by web connection type: 2008 90
42. Call counts, by interview status and mode of completion: 2008 91
43. Call counts, by institutional characteristic and student type: 2008 92
44. Interview items with high rates of help text access, by mode of administration: 2008 95
45. Item nonresponse for interview items with more than 10 percent of data missing, by section: 2008 96

46. Conversion rates for critical items, by type of response: 2008 98
47. Help desk incident type: 2008 100
48. Characteristics of interview respondents, by data collection phase: 2008 105
49. Characteristics of interview respondent, by completion type: 2008 106
50. Characteristics of interview respondent, by completion mode: 2008 107
51. Characteristics of interview respondent, by language of interview administration 2008 108
52. Baccalaureate determination, by student type: 2008 109
53. Description of codes for missing data: 2008 113
54. Summary of institution nonresponse bias analysis, by type of institution: 2008 124
55. Weight adjustment factors for institution poststratification: 2008 125
56. Weight adjustment factors for institution nonresponse adjustment: 2008 127
57. Institution weight distribution and unequal weighting effects, by type of institution: 2008 131
58. Weight adjustment factors for unknown student eligibility status: 2008 134
59. Weight adjustment factors for student location nonresponse adjustment: 2008 139
60. Weight adjustment factors for student refusal and other nonresponse adjustment: 2008 143
61. Weight adjustment factors for student poststratification: 2008 150
62. Student weight distribution and unequal weighting effects: 2008 158
63. Summary of weighted item response rates for all students and by type of institution: 2008 161

This page intentionally left blank.

List of Figures

1. Sequence of NPSAS:08 sampling operations: 2008 8
2. NPSAS:08 student sampling strata: 2008 11
3. Structure and content of computer-assisted data entry student record abstraction instrument: 2008 24
4. Structure and flow of the NPSAS:08 student interview: 2008 25
5. The NPSAS:08 institutional website home page: 2008 30
6. NPSAS:08 web-CADE main menu: 2008 37
7. NPSAS:08 student website: 2008 40
8. Data collection phases for NPSAS:08: 2008 43
9. NPSAS:08 data collection plan: 2008 44
10. NPSAS:08 student locating results: 2008 64
11. Percentage distribution of interview completions, by data collection phase: 2008 71
12. Cumulative flow of computer-assisted data entry completions: 2004 and 2008 84
13. Cumulative flow of computer-assisted data entry completions, by abstraction mode: 2008 84
14. Recode results for major field of study, by completion mode: 2008 93
15. Upcode results, by coding system and completion mode: 2008 94
16. Pop-up message displayed after three consecutive blank screens: 2008 99
17. Question delivery error rates, by week: 2008 101
18. Data entry error rates, by week: 2008 102
19. Receiver Operating Characteristic curve for overall institution response propensity: 2008 132
20. Receiver Operating Characteristic curve for overall student response propensity: 2008 159

This page intentionally left blank.

Chapter 1. Overview

This report describes the methods and results for the 2007–08 National Postsecondary Student Aid Study (NPSAS:08). The 2008 study was conducted for the National Center for Education Statistics (NCES), U.S. Department of Education, Washington, DC, as authorized by the Higher Education Act of 2008, 20 U.S.C. § 1015 (2007). Previous cycles of NPSAS, as well as two longitudinal studies deriving from it—the Beginning Postsecondary Students Longitudinal Study (BPS) and the Baccalaureate and Beyond Longitudinal Study (B&B)—are authorized by the following legislation:

- the Higher Education Act of 1965, as amended by the Higher Education Opportunity Act of 2008, 20 U.S.C. § 1015 (2008);
- the General Education Provisions Act, as amended, 20 U.S.C. §§ 9541 to 9548 (2007);
- the Higher Education Act of 1965, as amended by the Higher Education Amendments of 1986, 20 U.S.C. § 1070 et seq. (2007); and
- the National Education Statistics Act of 1994, as amended, 20 U.S.C. §§ 9541 to 9547 and 9573 (2007).

1.1 Background and Purpose of NPSAS

NPSAS is a comprehensive, nationwide study to determine how students and their families pay for postsecondary education; it features a nationally representative sample of U.S. students (aided and nonaided) in postsecondary education institutions. The sample comprises undergraduate, graduate, and first-professional students. These students attend all kinds of institutions, including public and private institutions, for-profit and not-for-profit institutions, and less-than-2-year institutions to 4-year colleges and universities.

The first NPSAS was conducted during the 1986–87 school year to meet the need for nationally representative data about significant postsecondary financial aid issues. NPSAS has been conducted every 3 to 4 years since 1987. Prior to the current cycle, the most recent NPSAS was conducted during the 2003-04 academic year. Beginning in 1990, each NPSAS data collection has provided base-year data for either the BPS or the B&B longitudinal study. NPSAS:08 serves as the base-year study for B&B. B&B eligible students were surveyed again in 2009, and it is anticipated they will be surveyed again in 2012.

The main objective of NPSAS is to produce reliable national estimates of characteristics related to financial aid for postsecondary students. No other single national database contains student-level records for students receiving financial aid from the many varied programs funded by the federal government, the states, postsecondary institutions, employers, and private organizations. The study addresses two questions with important policy implications for financial aid programs:

- How do students and their families finance postsecondary education?
- Who applies for and who receives aid?

This report describes the data collection plan, outcomes, and analysis for NPSAS:08. All procedures and methods were developed in consultation with a Technical Review Panel composed

of experts in higher education (appendix A). Chapter 2 describes the study design and data collection methods. Chapter 3 presents the overall outcomes of the several stages of data collection. Chapter 4 presents the results of data-quality analysis, with procedural and methodological evaluations. Chapter 5 delineates procedures used in variable construction and data file preparation. Chapter 6 presents the nonresponse bias analyses, weighting procedures, and variance estimation. Materials used during the study are provided as appendixes to the report and cited in the text where appropriate.

All analyses conducted to evaluate the effectiveness of the NPSAS:08 procedures are discussed. Unless otherwise indicated, a probability level of .05 was the criterion in all tests of significance. Throughout the report, numbers of sample institutions and students have been rounded to ensure the confidentiality of individual student data; therefore, row and column entries in tables may not sum to their respective totals, and reported percentages (which were calculated from unrounded numbers) may differ somewhat from those that would result from these rounded numbers.

1.2 Special Features of NPSAS:08

1.2.1 State Representative Sample

As with NPSAS:04, NPSAS:08 oversampled institutions in selected states to provide state-level representative data on undergraduate students in four institutional sectors (public 4-year, public 2-year, private not-for-profit 4-year, and private for-profit 4-year). The states included in the NPSAS:08 augmentation were California, Georgia, Illinois, Minnesota, New York, and Texas. The decision to include a six-state representative sample was made after the initial NPSAS:08 field test and full-scale institutional samples had been selected; the full-scale institutional sample was therefore augmented to allow the data to be representative at the state level.

1.2.2 The Academic Competitiveness Grant and the Science and Mathematics Access to Retain Talent Grant

A special set of items was included in the NPSAS:08 student interview to ascertain students' knowledge of the Academic Competitiveness Grant (ACG) and the Science and Mathematics Access to Retain Talent (SMART) Grant. These two grant programs were introduced by the federal government in 2006. Respondents who were eligible to receive these grants were asked whether they were aware of them. Those respondents who knew of them were then asked a series of questions related to the grants. To ensure that a sufficient number of students would provide responses to this set of questions, the NPSAS:08 student sample was augmented to include additional students who each received either an ACG or a SMART Grant. These data will provide researchers with baseline measures about students' knowledge of these new grant programs.

1.2.3 Virtual Call Center

Another important change in NPSAS:08 was the introduction of the Virtual Call Center (VCC) to supplement the on-site call center. Although all previous studies conducted telephone interviewing solely in an on-site call center, the introduction of new technology allowed telephone interviewers to accomplish securely and remotely, from their homes, the same tasks that were accomplished solely in the call center (e.g., help desk, prompting, locating, and interviewing). A subset of telephone interviewers conducted outbound telephone interviewing, remotely from their

homes. This innovation made data collection more efficient because it increased interviewing capacity without requiring call center expansion.

1.3 Schedule and Products of NPSAS:08

This section presents the schedule for the major activities for the full-scale study (table 1) and describes the publications that will be available.

Table 1. Start and end dates for major NPSAS:08 activities

Activity	Start date[1]	End date[2]
Select institutional sample[3]	10/3/2005	9/21/2007
Mail information and make phone contact with chief administrator	9/24/2007	11/16/2007
Mail information and make phone contact with institutional coordinator	10/2/2007	12/20/2007
Obtain enrollment lists for student sampling	1/14/2008	7/15/2008
Select student samples	1/15/2008	7/15/2008
Request/obtain CPS data	1/16/2008	8/29/2008
Receive list of SMART Grant/ACG recipients	1/10/2008	1/10/2008
Preload CPS data into CADE records	2/1/2008	7/22/2008
Conduct CADE record abstraction	2/1/2008	9/19/2008
Conduct student interviews		
Web	3/10/2008	9/29/2008
Telephone	3/31/2008	9/29/2008

[1] Date the activity was initiated for the first applicable school or its associated students.
[2] Date the activity was completed for the last applicable school or its associated students.
[3] The NPSAS:08 field test and full-scale institutional samples were selected at the same time to prevent institutions being selected for both samples. Prior to the start of full-scale data collection in spring 2007, the sample was freshened with 2005–06 IPEDS data to ensure that all potentially eligible institutions were included. Finally, in fall 2007, the sample was augmented to include additional institutions for the six-state augmentation; consequently, approximately 20 institutions were included in both samples.
NOTE: ACG = Academic Competitiveness Grant; CADE = computer-assisted data entry; CPS = Central Processing System; IPEDS = Integrated Postsecondary Education Data System SMART = Science and Mathematics Access to Retain Talent.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

Reports based on NPSAS:08 will be published as follows:

- *The Profile of Undergraduates in U.S. Postsecondary Education Institutions: 2007–08* will describe the demographic and enrollment characteristics of undergraduate students.
- *Student Financing of Undergraduate Education: 2007–08* will address undergraduate tuition, total price of attendance, types and sources of financial aid received, net price, financial aid need, and unmet need.
- *Student Financing of Graduate and First-Professional Education: 2007–08* will describe the demographic and enrollment characteristics of graduate and first-professional students, as well as the types and sources of financial aid received.

The following, together with the Data Analysis System (DAS), are available on the NCES website at http://nces.ed.gov/:

- *2007–08 National Postsecondary Student Aid Study (NPSAS:08): Student Financial Aid Estimates for 2007–08* report is a First Look publication based on the 2007–08 data; it gives the

percentages of students receiving various types of financial aid and average amounts received, by type of institution attended, attendance pattern, dependency status, and income level.

- *2007–08 National Postsecondary Student Aid Study (NPSAS:08): Undergraduate Financial Aid Estimates for 2007–08 by Type of Institution* report is the second publication based on the 2007–08 data. This report focuses only on undergraduates and includes separate tables for those who attended public 4-year, those who attended private not-for-profit 4-year, those who attended public 2-year, and those who attended private for-profit postsecondary institutions during the 2007–08 academic year. It gives average tuition and fees, average total price of attendance, and the percentages of undergraduates receiving various types and combinations of financial aid, together with average amounts received and with a particular focus on grants and loans.
- *2007–08 National Postsecondary Student Aid Study (NPSAS:08): Undergraduate Data Analysis System* (NPSAS:08 Undergraduate DAS) contains the data on a sample of about 114,000 undergraduates who were enrolled at any time between July 1, 2007, and June 30, 2008, in about 1,730 postsecondary institutions. The data represent all undergraduate students enrolled in postsecondary institutions in the 50 states, the District of Columbia, and Puerto Rico that were eligible to participate in the federal financial aid programs under Title IV of the Higher Education Act.
- *2007–08 National Postsecondary Student Aid Study (NPSAS:08): Graduate Data Analysis System* (NPSAS:08 Graduate DAS) contains the data on a sample of about 14,000 graduate students who were enrolled at any time between July 1, 2007, and June 30, 2008, in about 1,730 postsecondary institutions. The data represent all graduate students enrolled in postsecondary institutions in the 50 states, the District of Columbia, and Puerto Rico that were eligible to participate in the federal financial aid programs under Title IV of the Higher Education Act.

To access the NPSAS:08 DAS application or one of the NPSAS:08 reports, contact Aurora D'Amico, or visit the website at http://nces.ed.gov/das/:

Aurora D'Amico
Postsecondary Studies Division
National Center for Education Statistics
1990 K Street NW
Washington, DC 20006-5652
(202) 502-7334
aurora.d'amico@ed.gov

Analysts can also use the QuickStats system to generate simple tables and graphs quickly using the NPSAS:08 data. To access this system, visit the website at http://nces.ed.gov/datalab/quickstats/. The survey data files used in developing variables in the DASs, as well as the associated electronic codebooks and file documentation, are available to researchers who have obtained a restricted data license from NCES. Information on the NCES Statistical Standards Program, including Restricted Use Data Licenses Procedures, is available from the NCES website: http://nces.ed.gov/statprog/. Additional information on obtaining a restricted data license may be found in the NCES *Restricted-Use Data Procedures Manual* (U.S. Department of Education 1999) at http://nces.ed.gov/statprog/rudman/.

Chapter 2. Design and Methodology

This chapter describes in detail the design and methods implemented for the 2007–08 National Postsecondary Student Aid Study (NPSAS:08). First, the two-stage sampling design is described. The content and design of the computer-assisted data entry (CADE) and student instruments are then explained, followed by the data collection process. The steps involved in contacting, locating, and interviewing sample members are next delineated, and the chapter concludes with a description of the data collection systems.

2.1 Sampling

2.1.1 Target Population and Sampling Overview

The NPSAS:08 target population consisted of all eligible students enrolled at any time between July 1, 2007, and June 30, 2008, in eligible postsecondary institutions in the United States or Puerto Rico. For institutions to be eligible for the study, they had to have signed Title IV participation agreements with the U.S. Department of Education, making them eligible for federal student aid programs. For students to be eligible for the study, they had to have been

- enrolled in an academic program or taken at least one course for credit applicable toward an academic degree, or
- enrolled in an occupational or vocational program that required at least 3 months or 300 clock hours of instruction for receipt of a degree, certificate, or other formal award.

Students were ineligible if they were concurrently enrolled in high school or enrolled solely in a General Educational Development (GED) program or other high school completion program.

The institutional sampling frame for NPSAS:08 was constructed from the 2004–05 Integrated Postsecondary Education Data System (IPEDS:2004–05) Institutional Characteristics, Fall Enrollment, and Completions files. The IPEDS data used for the initial sampling frame were collected in 2005; the IPEDS data used for sample freshening, in 2007.[6] Consequently, any institutions that came into existence or became eligible between the IPEDS data collections in 2006 and June 30, 2008, were not included on the sampling frame used for freshening. Ineligible institutions (e.g., institutions located outside the United States and Puerto Rico, central offices, military academies) were deleted from the sampling frame.

The remaining institutions in the sampling frame were partitioned into 46 institutional strata based on institutional level and control, highest level of offering, and proportion of bachelor's degrees awarded in education.[7] NPSAS:08 also included state-representative undergraduate student samples for four degree-granting institutional sectors in six states: (1) public 2-year institutions; (2) public 4-year institutions; (3) private, not-for-profit 4-year institutions; and (4) private, for-profit

[6] The process of freshening the sample involved adding a sample of all institutions that became eligible between 2005 and 2006 to the original institutional sample.

[7] The proportion of bachelor's degrees awarded in education is used to ensure sufficient numbers of sample students receiving a bachelor's degree in education. Such students are an important analysis domain for the Baccalaureate and Beyond Longitudinal Study (B&B).

institutions.[8] The 46 institutional strata—22 nationally representative and 24 state representative—were as follows:

- public less-than-2-year;
- public 2-year;
- public 4-year non-doctorate-granting bachelor's high education;
- public 4-year non-doctorate-granting bachelor's low education;
- public 4-year non-doctorate-granting master's high education;
- public 4-year non-doctorate-granting master's low education;
- public 4-year doctoral high education;
- public 4-year doctoral low education;
- public 4-year first-professional-granting high education;
- public 4-year first-professional-granting low education;
- private not-for-profit less-than-2-year;
- private not-for-profit 2-year;
- private not-for-profit 4-year non-doctorate-granting bachelor's high education;
- private not-for-profit 4-year non-doctorate-granting bachelor's low education;
- private not-for-profit 4-year non-doctorate-granting master's high education;
- private not-for-profit 4-year non-doctorate-granting master's low education;
- private not-for-profit 4-year doctoral high education;
- private not-for-profit 4-year doctoral low education;
- private not-for-profit 4-year first-professional-granting high education;
- private not-for-profit 4-year first-professional-granting low education;
- private for-profit less-than-2-year;
- private for-profit 2-year-or-more;
- California public 2-year;
- California public 4-year;
- California private not-for-profit 4-year;
- California private for-profit degree-granting;

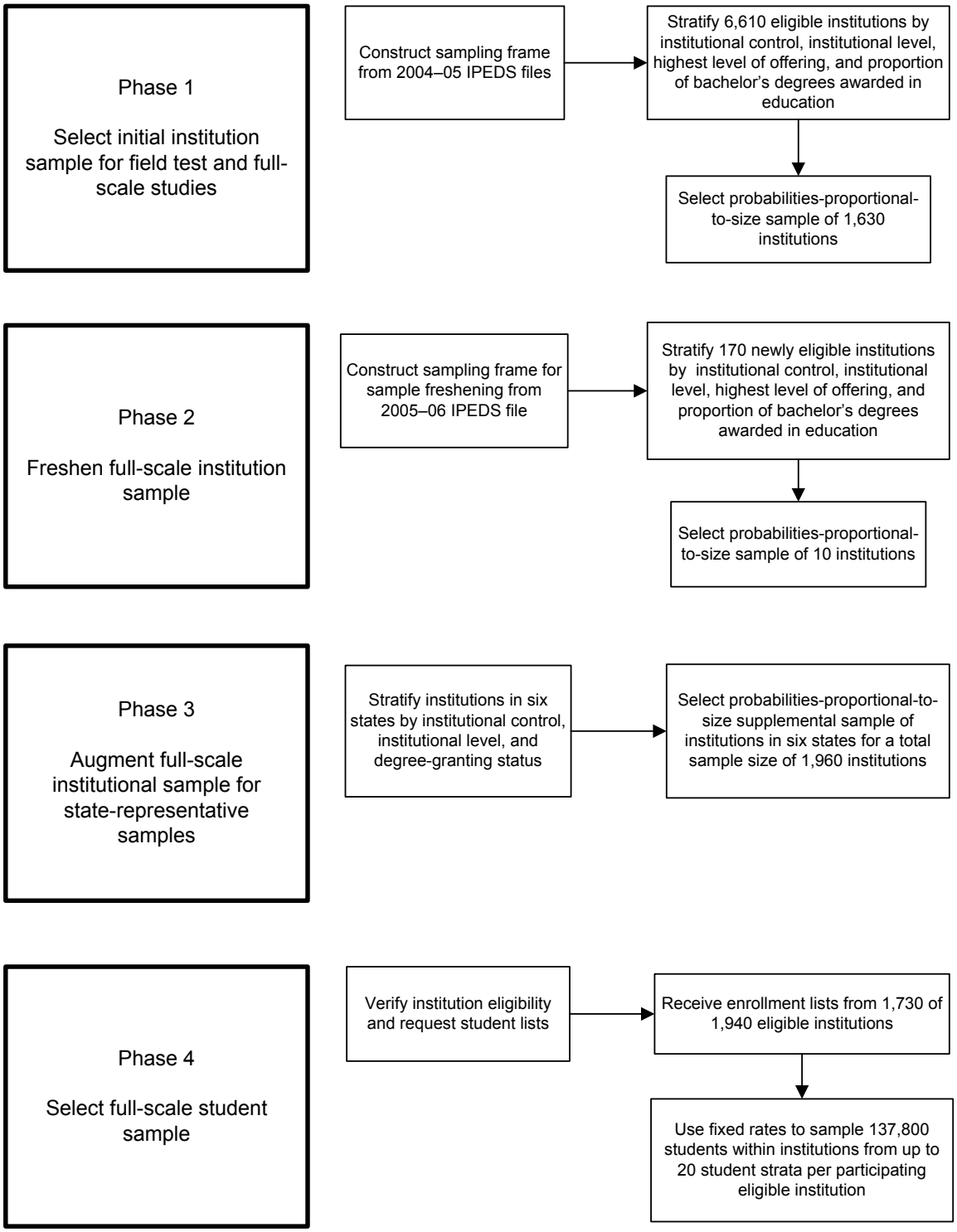
[8] These six states were selected on the basis of undergraduate enrollment numbers in the four sectors; prior inclusion in the NPSAS:04 12-state sample, with high levels of cooperation and participation in that survey; and interstate variation or statewide recent changes in tuition and state grant policies, which enable comparative research and analysis. The six states were California, Georgia, Illinois, New York, Minnesota, and Texas.

- Texas public 2-year;
- Texas public 4-year;
- Texas private not-for-profit 4-year;
- Texas private for-profit degree-granting;
- New York public 2-year;
- New York public 4-year;
- New York private not-for-profit 4-year;
- New York private for-profit degree-granting;
- Illinois public 2-year;
- Illinois public 4-year;
- Illinois private not-for-profit 4-year;
- Illinois private for-profit degree-granting;
- Georgia public 2-year;
- Georgia public 4-year;
- Georgia private not-for-profit 4-year;
- Georgia private for-profit degree-granting;
- Minnesota public 2-year;
- Minnesota public 4-year;
- Minnesota private not-for-profit 4-year; and
- Minnesota private for-profit degree-granting.

High education refers to the 20 percent of institutions with the highest proportions of their baccalaureate degrees awarded in education (based on the most recent IPEDS Completions file). The remaining 80 percent of institutions are classified as *low education*, meaning they have a lower proportion of baccalaureate degrees awarded in education. The statistical sampling process for NPSAS:08 comprised four phases (figure 1).

Institutions were selected by means of Chromy's (1979) sequential probability minimum replacement sampling algorithm, which resembles systematic sampling, to select institutions with probabilities proportional to a composite measure of size based on expected enrollment. Initially a sample of 1,630 institutions was selected in August 2006 so that these institutions could be notified of their selection early and a separate field test sample could be selected from the remaining institutions on the sampling frame. To keep the sample representative of the population, in July 2007 the sample was freshened to include about 10 newly eligible institutions. In September 2007, after the decision was made to include state-representative undergraduate samples in the six states, a supplemental sample of institutions was selected in these states and added to the existing sample.

Figure 1. Sequence of NPSAS:08 sampling operations: 2008



NOTE: IPEDS = Integrated Postsecondary Education Data System.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

The next step in the institutional sampling process was to determine eligibility. Of the sampled institutions (*n* = 1,960), about 1,940 (99 percent) were found to be NPSAS-eligible institutions.[9]

The sampling frames for selecting students were enrollment lists provided by the eligible sample institutions of students meeting eligibility criteria and enrolled during the 2007–08 academic year. Student enrollment lists were sampled as they were received, using equal probability stratified systematic sampling. Student sampling strata were as follows:

- in-state potential baccalaureate recipients who were business majors;
- out-of-state potential baccalaureate recipients who were business majors;
- in-state potential baccalaureate recipients who were science, technology, engineering, or mathematics (STEM) majors and Science and Mathematics Access to Retain Talent (SMART) Grant recipients;
- out-of-state potential baccalaureate recipients who were STEM majors and SMART Grant recipients;
- in-state potential baccalaureate recipients who were STEM majors and not SMART Grant recipients;
- out-of-state potential baccalaureate recipients who were STEM majors and not SMART Grant recipients;
- in-state potential baccalaureate recipients in all other majors who were SMART Grant recipients;
- out-of-state potential baccalaureate recipients in all other majors who were SMART Grant recipients;
- in-state potential baccalaureate recipients in all other majors who were not SMART Grant recipients;
- out-of-state potential baccalaureate recipients in all other majors who were not SMART Grant recipients;
- in-state other undergraduate students who were SMART Grant recipients;
- out-of-state other undergraduate students who were SMART Grant recipients;
- in-state other undergraduate students who were Academic Competitiveness Grant (ACG) recipients;
- out-of-state other undergraduate students who were ACG recipients;
- in-state other undergraduate students who were not SMART Grant or ACG recipients;
- out-of-state other undergraduate students who were not SMART Grant or ACG recipients;
- master's students;

[9] Institutional eligibility requirements for NPSAS:08 are discussed in section 2.1.2.

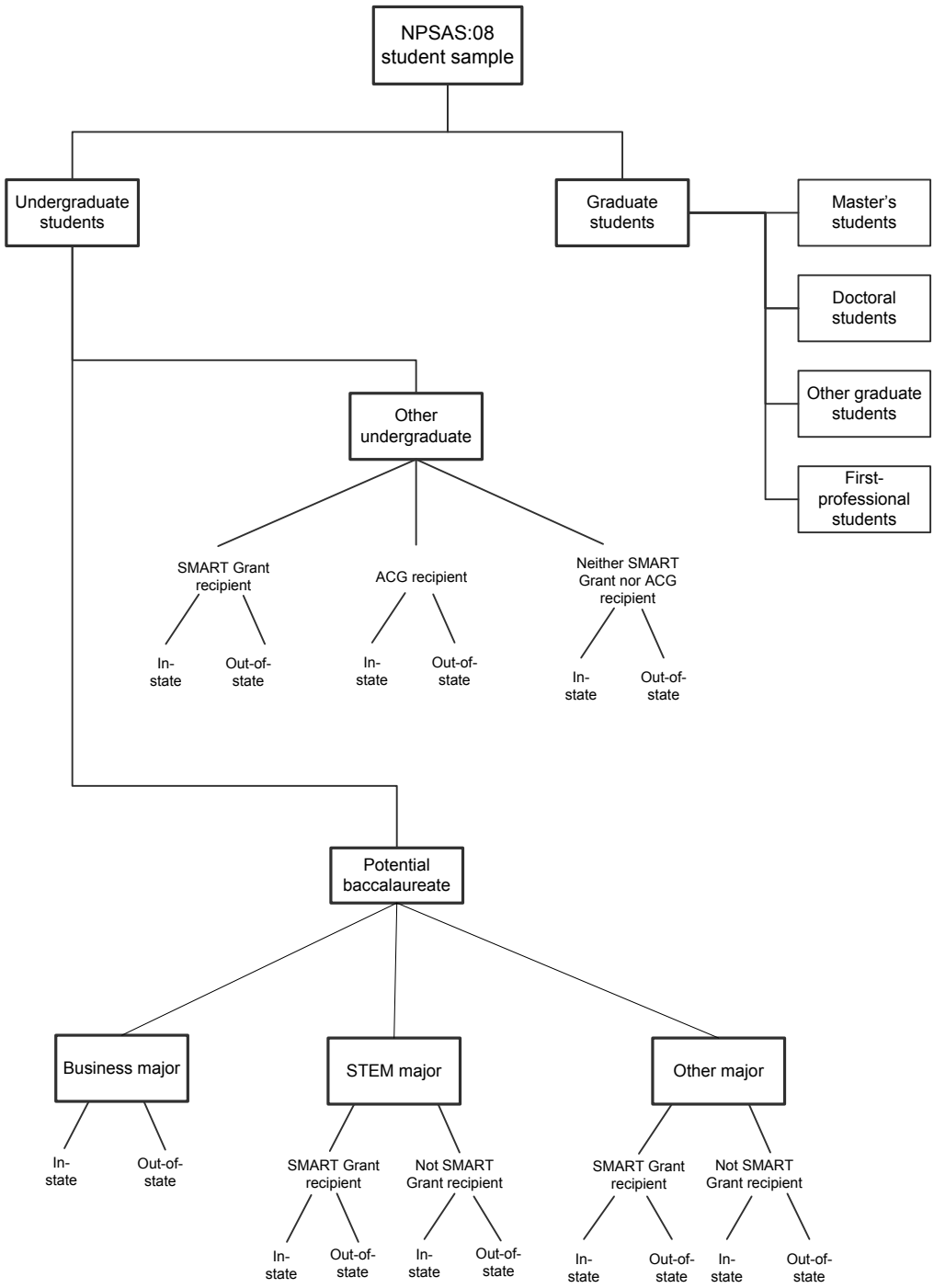
- doctoral students;
- other graduate students; and
- first-professional students.

These 20 strata are most easily conceptualized in groups presented graphically (figure 2).

Potential baccalaureate recipients were students identified by the institution as having received or as expecting to receive a baccalaureate degree between July 1, 2007, and June 30, 2008. These students were stratified separately from other undergraduate students because they were oversampled to allow for sufficient numbers to be surveyed in the 2009 and 2012 follow-up Baccalaureate and Beyond Longitudinal Studies (B&B:08/09/12). *Potential baccalaureate recipients* and *other undergraduate students* were each divided into in-state and out-of-state strata because undergraduate in-state students were oversampled in the six states with state-representative samples. These in-state and out-of-state strata were used for all institutions to allow for sampling ease and consistency; however, in states without state-representative samples, in-state students were sampled at the same rate as out-of-state students. Additionally, STEM majors, SMART Grant recipients, and ACG recipients were oversampled to allow for sufficient numbers for analysis. A large percentage of baccalaureate recipients major in business, so within the set of potential baccalaureate recipients, business majors were undersampled to ensure that the baccalaureate sample did not consist primarily of business majors.

For each student stratum, the enrollment list was sampled at a rate designed to provide approximately equal student-level probabilities. To more accurately estimate the overall sample yield, student sampling rates were revised after sufficient lists had been received. The sampling rates were set to meet the sample sizes for the national sample (table 2) and to meet the sample size for the state sample (table 3). The final sample included 137,800 students; the target was 138,050 (see table 2).

Figure 2. NPSAS:08 student sampling strata: 2008



NOTE: ACG = Academic Competitiveness Grant; SMART = Science and Mathematics Access to Retain Talent; STEM = science, technology, engineering, or mathematics.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

Table 2. Target numbers of sample students, by institutional stratum and student type: 2008

	All students			Undergraduates									Graduate students			First professionals		
				All			Bachelor's			Other								
Type of institution	Number expected	Number achieved	Percent achieved	Number expected	Number achieved	Percent achieved	Number expected	Number achieved	Percent achieved	Number expected	Number achieved	Percent achieved	Number expected	Number achieved	Percent achieved	Number expected	Number achieved	Percent achieved
All institutions	138,050	137,800	99.8	124,980	125,110	100.1	30,890	29,470	95.4	94,090	95,650	101.7	11,020	10,760	97.7	2,050	1,920	93.9
Public less-than-2-year	2,200	1,730	78.7	2,200	1,730	78.7	†	†	†	2,200	1,730	78.7	†	†	†	†	†	†
Public 2-year	39,500	39,340	99.6	39,500	39,330	99.6	†	100	†	39,500	39,230	99.3	†	10	†	†	†	†
Public 4-year non-doctorate-granting	13,920	16,120	115.8	12,660	15,440	122.0	4,750	5,890	123.9	7,910	9,550	120.8	1,270	680	53.7	†	†	†
Public 4-year doctorate-granting	32,900	30,280	92.0	28,620	25,490	89.0	13,740	11,610	84.5	14,880	13,880	93.2	3,640	4,210	115.7	640	590	92.0
Private not-for-profit less-than-4-year	2,340	2,080	88.5	2,340	2,080	88.5	†	10	†	2,340	2,070	88.2	†	†	†	†	†	†
Private not-for-profit 4-year non-doctorate-granting	15,320	14,200	92.7	12,770	12,840	100.5	6,340	5,550	87.5	6,430	7,290	113.3	2,540	1,360	53.4	†	10	†
Private not-for-profit 4-year doctorate-granting	14,690	16,480	112.2	9,840	10,860	110.4	4,840	4,660	96.4	5,000	6,200	123.9	3,440	4,290	124.7	1,410	1,330	94.0
Private for-profit less-than-2-year	6,620	6,610	99.9	6,620	6,610	99.9	†	†	†	6,620	6,610	99.9	†	†	†	†	†	†
Private for-profit 2-year-or-more	10,540	10,960	103.9	10,430	10,740	103.0	1,220	1,640	134.9	9,210	9,100	98.8	130	210	169.8	†	†	†

† Not applicable.
NOTE: Detail may not sum to totals because of rounding. Baccalaureate and other undergraduate counts are based on having received or expecting to receive a bachelor's degree at the time of sampling.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

Table 3. Target numbers of sample students in the six state-representative samples, by institutional stratum and type of student: 2008

	All undergraduates			Baccalaureates			Other undergraduates		
	Number expected	Number achieved	Percent achieved	Number expected	Number achieved	Percent achieved	Number expected	Number achieved	Percent achieved
California	12,140	11,150	91.8	2,400	2,210	92.0	9,740	8,940	91.8
Public 2-year	5,920	5,370	90.8	†	†	†	5,920	5,370	90.8
Public 4-year	3,510	3,470	98.8	1,570	1,490	94.5	1,940	1,980	102.3
Private not-for-profit 4-year	1,360	1,260	92.3	670	560	83.2	690	700	101.3
Private for-profit degree-granting	1,350	1,060	78.1	160	170	105.1	1,190	890	74.5
Georgia	8,340	9,120	109.3	1,450	2,230	153.6	6,890	6,890	100.0
Public 2-year	4,240	4,170	98.4	†	†	†	4,240	4,170	98.4
Public 4-year	1,460	1,930	132.0	660	1,410	215.1	810	520	64.6
Private not-for-profit 4-year	1,290	1,360	105.4	640	640	100.8	650	720	109.8
Private for-profit degree-granting	1,350	1,660	122.8	160	180	112.0	1,190	1,480	124.0
Illinois	7,000	7,500	107.1	1,520	1,680	110.5	5,480	5,820	106.1
Public 2-year	2,740	2,760	100.5	†	†	†	2,740	2,760	100.5
Public 4-year	1,610	1,640	101.9	720	710	98.8	890	930	104.4
Private not-for-profit 4-year	1,300	1,680	128.6	640	720	112.1	660	950	144.6
Private for-profit degree-granting	1,350	1,430	106.1	160	250	157.6	1,190	1,180	99.2
Minnesota	7,770	9,220	118.6	1,470	1,850	126.1	6,310	7,370	116.8
Public 2-year	3,630	4,640	127.7	†	†	†	3,630	4,640	127.7
Public 4-year	1,540	2,190	142.2	690	1,210	176.2	850	970	114.5
Private not-for-profit 4-year	1,250	1,140	90.8	620	470	76.3	630	670	105.1
Private for-profit degree-granting	1,350	1,260	93.1	160	160	103.2	1,190	1,090	91.6
New York	10,470	10,880	103.9	2,480	2,590	104.4	7,990	8,310	104.0
Public 2-year	4,160	4,180	100.6	†	†	†	4,160	4,180	100.6
Public 4-year	2,890	2,780	96.2	1,300	1,080	83.4	1,600	1,700	106.6
Private not-for-profit 4-year	2,080	2,380	114.7	1,030	1,190	116.2	1,050	1,190	113.2
Private for-profit degree-granting	1,350	1,540	114.1	160	300	188.6	1,190	1,240	104.0
Texas	8,220	7,560	92.1	2,150	1,730	80.7	6,070	5,830	96.0
Public 2-year	2,560	2,430	94.9	†	†	†	2,560	2,430	94.9
Public 4-year	3,050	2,710	88.7	1,370	1,060	77.1	1,680	1,650	98.2
Private not-for-profit 4-year	1,250	1,100	88.0	620	460	75.0	630	640	100.8
Private for-profit degree-granting	1,350	1,320	97.9	160	210	133.5	1,190	1,110	93.0

† Not applicable.
NOTE: Detail may not sum to totals because of rounding. Baccalaureate and other undergraduate counts are based on having received or expected to receive a bachelor's degree at the time of sampling.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

The sample size for NPSAS:08 was larger than that of past NPSAS cycles primarily to ensure sufficient yield for analytic purposes. The sample size was designed so that respondent yield would be sufficient for analyses even if actual response rates were lower than the targeted rates.

2.1.2 Institution Sample and Eligibility

The target population for NPSAS:08 included nearly all Title IV participating postsecondary institutions in the 50 states, the District of Columbia, and Puerto Rico. To be eligible for NPSAS:08, an institution, during the 2007–08 academic year, had to have

- offered an educational program designed for persons who had completed secondary education;
- offered at least one academic, occupational, or vocational program of study lasting at least 3 months or 300 clock hours;
- offered courses open to persons other than the employees or members of the company or group (e.g., union) that administered the institution;
- been located in the 50 states, the District of Columbia, or Puerto Rico;
- been other than a U.S. Service Academy;[10] and
- signed a Title IV participation agreement with the U.S. Department of Education.

Institutions providing only avocational, recreational, or remedial courses or only in-house courses for their own employees were excluded. The listed eligibility requirements match those used in previous NPSAS rounds, with two exceptions: (1) the Title IV participation requirement began with NPSAS:2000, and (2) a requirement that more than one correspondence course be offered ended with NPSAS:04.

The institutional and student samples were allocated to the strata in accordance with the cost/variance optimization procedure (appendix B). Student sampling rates, which were used to compute institution-level composite measures of size, were based on the IPEDS:2004–05 Fall Enrollment Survey counts and the required sample sizes (for details, see appendix B).

With use of Chromy's (1979) sequential probability minimum replacement sampling algorithm, which resembles systematic sampling, an independent sample of institutions was selected for each institutional stratum. This algorithm led to the selection of institutions with probabilities proportional to a composite measure of size based on expected enrollment. To avoid multiple selections of sample institutions, those with expected frequencies of selection greater than unity (1.00) were selected with certainty. The remainder of the institution sample was selected from the remaining institutions within each stratum. To ensure the positive pairwise probabilities of selection required for proper variance estimation (Chromy 1981), the sampling algorithm was implemented with a random start for each institutional stratum.

The sample of institutions was initially selected in August 2006 to allow the field test sample institutions to be selected from the complement of the full-scale sample. This plan eliminated the possibility of an institution's being burdened with participation in both the field test and full-scale studies. Because additional institutions were later selected for the state augmentation, however, 20 institutions that participated in the field test study were included in the full-scale study.

[10] These academies were not eligible for this financial aid study because of their unique funding/tuition base.

In July 2007, the sample was freshened with use of the IPEDS:2005–06 Institutional Characteristics, Fall Enrollment, and Completions files; any newly eligible institutions in the sampling frame were in this way included to ensure representativeness of the current population. The measures of size for the supplemental sampling frame from which the freshened sample was selected were based on the IPEDS:2005–06 Fall Enrollment Survey counts. The sample size for freshened institutions was set so that they would have similar probabilities of selection as if they were selected in the original sample. This minimized the chance of unequal student weights and variance estimates due to different probabilities of selection. This process added about 10 institutions to the sample.

In summer 2007, augmentation of the NPSAS:08 full-scale sample was undertaken to include state-representative samples of undergraduate students in four sectors from six states. This augmentation enabled state-level analyses and comparisons of many of the most pertinent issues in postsecondary financial aid and costs. The choice of states for the sample augmentation was based on several considerations:

- Undergraduate enrollment numbers were considered in four sectors: (1) public 2-year institutions; (2) public 4-year institutions; (3) private, not-for-profit 4-year institutions; and (4) private, for-profit institutions. Tuition and grant policies in the states with the largest enrollments had the greatest effect on national patterns and trends; the undergraduate representation in a national sample was already so large that the cost of sample augmentation would be relatively low.
- Prior inclusion in the 12-state sample of NPSAS:04 and high levels of cooperation and participation in that survey were considered. States' participation in NPSAS is not mandatory, so institutional cooperation is necessary to achieve the response rates and yields required for reliable estimates. Smaller states that were willing and helpful in NPSAS:04 and achieved high yields and response rates were considered more likely than others to cooperate currently.
- States were considered when they had different tuition and state grant policies from other states or had recent changes in these policies; these states would provide opportunities for comparative research and analysis.

With these considerations as criteria, the sample was augmented for six states: California, Georgia, Illinois, Minnesota, New York, and Texas.

Sample selection for the state-representative institutional samples adhered to guidelines regarding the number of schools in a given sector. For instance, if a sector within a state contained 36 or fewer institutions then all institutions were selected. If a sector within a state contained more than 36 institutions then 30 were selected. It was sometimes necessary to select all institutions in a sector within a state to adequately represent state-level systems. In order to augment the original institutional sample to allow for state representation, three scenarios were possible:

1. The original institutional sample contained a sufficient number of institutions within each sector, so no additional institutions were needed.
2. If a sector within a state had more than 36 institutions but the original sample did not contain a sufficient number of institutions to be representative, the original sample was discarded and a new sample was selected. Statistically, this method was the cleanest, in terms of computing probabilities of selection and not introducing additional variance.

3. If a sector within a state had 36 or fewer institutions and the original sample did not contain all 36, the institutions not included in the original sample were added to the augmented sample. In some cases, a sector within a state may have had more than 36 institutions but it was necessary to select all of the institutions in the sector to adequately represent state-level systems. In these cases, all institutions not included in the original sample were added to the augmented sample.

The institutional sampling rates and the numbers of certainty and noncertainty institutions selected for each of the 22 national strata and the 24 state strata are reported (table 4). The institutions included in the national sample were selected from all 46 strata, whereas institutions included in the state samples were selected only from the 24 state strata. Within each institutional stratum, additional *implicit stratification* was accomplished by sorting the stratum sampling frame by the following classifications:[11] (1) level of for-profit 2-year-or-more institutions; (2) Historically Black Colleges and Universities indicator; (3) Hispanic-serving-institution indicator; (4) Carnegie classifications of postsecondary institutions; (5) the Office of Business Economics region (Bureau of Economic Analysis of the U.S. Department of Commerce region); and (6) the measure of institution size.[12] Further implicit stratification within region and by state and system was completed for some state systems, the objective being to approximate proportional representation of institutions on these measures.[13]

[11] *Implicit stratification* is the process in which substrata are created by sorting the data within stratum when using the probability minimum replacement sampling algorithm.

[12] For sorting purposes, Alaska and Hawaii were combined with Puerto Rico in the Outlying Areas region rather than in the Far West region.

[13] The implicit stratification within region and by state and system was completed on the initial institution sample before the decision was made to have state-representative samples.

Table 4. Institutional sampling rates and number of certainty and noncertainty institutions sampled, by institutional stratum: 2008

Institutional stratum[1,2]	Size of universe[3]	Sampling rate	Number of sample institutions		
			Total	Certainty	Noncertainty
Total	6,777	0.29	1,960	1,030	930
Public					
Less-than-2-year	247	0.09	20	#	20
2-year	835	0.23	190	30	160
4-year non-doctorate-granting bachelor's high education	16	0.63	10	#	10
4-year non-doctorate-granting bachelor's low education	48	0.52	30	10	20
4-year non-doctorate-granting master's high education	36	0.50	20	#	20
4-year non-doctorate-granting master's low education	152	0.41	60	10	50
4-year doctoral high education	23	1.00	20	20	#
4-year doctoral low education	73	1.00	70	70	#
4-year first-professional-granting high education	28	1.00	30	30	#
4-year first-professional-granting low education	96	1.00	100	100	#
Private not-for-profit					
Less-than-2-year	102	0.07	10	#	10
2-year	224	0.06	10	#	10
4-year non-doctorate-granting bachelor's high education	74	0.19	10	#	10
4-year non-doctorate-granting bachelor's low education	278	0.10	30	#	30
4-year non-doctorate-granting master's high education	88	0.18	20	#	20
4-year non-doctorate-granting master's low education	323	0.58	190	80	110
4-year doctoral high education	29	0.31	10	#	10
4-year doctoral low education	95	0.33	30	#	30
4-year first-professional-granting high education	58	1.00	60	60	#
4-year first-professional-granting low education	260	0.30	80	40	40
Private for-profit					
Less-than-2-year	1,476	0.07	100	#	100
2-year-or-more	1,029	0.08	80	10	80
California					
Public 2-year	114	0.65	70	30	50
Public 4-year	33	1.00	30	30	#
Private not-for-profit 4-year	96	0.36	40	20	10
Private for-profit degree-granting	106	0.28	30	10	20

See notes at end of table.

Table 4. Institutional sampling rates and number of certainty and noncertainty institutions sampled, by institutional stratum: 2008—Continued

Institutional stratum[1,2]	Size of universe[3]	Sampling rate	Number of sample institutions		
			Total	Certainty	Noncertainty
Georgia					
Public 2-year	54	1.00	50	50	#
Public 4-year	21	1.00	20	20	#
Private not-for-profit 4-year	29	1.00	30	30	#
Private for-profit degree-granting	24	1.00	20	20	#
Illinois					
Public 2-year	48	0.63	30	10	20
Public 4-year	12	1.00	10	10	#
Private not-for-profit 4-year	64	0.47	30	10	20
Private for-profit degree-granting	28	1.00	30	30	#
Minnesota					
Public 2-year	31	1.00	30	30	#
Public 4-year	11	1.00	10	10	#
Private not-for-profit 4-year	27	1.00	30	30	#
Private for-profit degree-granting	31	1.00	30	30	#
New York					
Public 2-year	38	1.00	40	40	#
Public 4-year	38	1.00	40	40	#
Private not-for-profit 4-year	139	0.41	60	30	30
Private for-profit degree-granting	42	0.71	30	20	10
Texas					
Public 2-year	64	0.47	30	10	20
Public 4-year	44	0.86	40	30	10
Private not-for-profit 4-year	48	0.63	30	20	10
Private for-profit degree-granting	45	0.67	30	20	10

Rounds to zero.
[1]Each institutional stratum is mutually exclusive
[2] Institutional stratum reflects institutional categorization as determined from the 2004–05 Integrated Postsecondary Education Data System (IPEDS) file; some errors in this classification were uncovered when institutions were contacted.
[3] Based on the 2004–05 IPEDS file.
NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

2.1.3 Student Sample and Eligibility

The postsecondary students eligible for NPSAS:08 were those who attended a NPSAS-eligible institution during the 2007–08 academic year and who were

- enrolled in *either* (1) an academic program or at least one course for credit that could be applied toward fulfilling the requirements for an academic degree, *or* (2) an occupational or vocational program that required at least 3 months or 300 clock hours of instruction to receive a degree, certificate, or other formal award;
- not concurrently enrolled in high school; and
- not enrolled *solely* in a GED or other high school completion program.

Each sampled institution verified as NPSAS-eligible was asked to provide a list of all students who satisfied all the NPSAS eligibility conditions, preferably an "unduplicated" electronic list (i.e., one in which each student's name appeared only once), together with identifying, classifying, and locating information.

The NPSAS:08 student sample of 138,000 included an augmentation of about 5,000 additional recipients of SMART Grants and ACGs to ensure sufficient numbers of these students for analysis relating to these new sources of student financial aid. The U.S. Department of Education provided a file of SMART Grant recipients so that students on the enrollment lists could be matched to this file and identified as recipients. Some schools did not have fall 2007 SMART Grant or ACG recipients on the file, but data were available on SMART Grant or ACG recipients from the previous school year.

As student lists were received from institutions, students were sampled by means of stratified systematic sampling with predetermined sampling rates that varied by student stratum. To eliminate cross-institution duplication, the Social Security numbers of those selected from an institution were compared with Social Security numbers of students who had already been selected from other institutions. Multiplicity adjustments in the sample weighting (described in more detail in section 6.2.1) accounted for the fact that any students who attended more than one institution during the NPSAS year had more than one chance of selection.

Some institutional systems sent in lists for multiple institutions or campuses. If the lists were separate for each institution or campus, then the samples were selected separately and independently. If the lists were combined into one list with no identifier mapping students to institution or campus, then one student sample was selected that represented all of the institutions or campuses included on the list. In such cases, sampling rates were adjusted, and a weight adjustment was made (see section 6.1.1).

Initial student sampling rates were calculated for each sample institution, using sampling rates designed to generate approximately equal probabilities of selection within the ultimate institution-by-student sampling strata (see appendix B). Sometimes these rates were modified, however, as follows:

- The student sampling rates were increased so that the sample size for each sample institution was at least 10 students (if possible) to ensure sufficient yield for variance estimation.
- Student sampling rates were decreased if the sample size was greater than 300.
- To ensure that the desired student sample sizes were achieved, sample yield was monitored throughout enrollment list collection, and student sampling rates were adjusted periodically for institutions for which sample selection had not yet been performed.

These adjustments to the initial sampling rates resulted in some additional variability in the student sampling rates and therefore in some increase in survey design effects (variance inflation; see section 6.4.3).

The planned and achieved sample sizes are reported by student stratum and level of offering (table 5). Also reported is the initial classification of the student sample overall and by institutional type and student stratum (table 6). The achieved sample yield of 137,800 was very close to the targeted 138,050 because institutional participation rates were correctly estimated and the sampling rates throughout the sampling process were persistently monitored and adjusted. Overall, more

baccalaureates, other undergraduate students,[14] and doctoral students were selected into the sample than planned, with fewer ACG, master's, and other graduate students than planned (for details about sample allocation, see appendix B, section B.4).[15]

Table 5. Planned and achieved NPSAS:08 student samples, by student stratum and institutional level: 2008

Student stratum[1]	Institutional level[2]	Students sampled: Number expected[3]	Number achieved[4]	Percent[5]
Total	All institutions	138,050	137,800	99.8
Baccalaureate business	4-year	2,500	3,270	131.0
Baccalaureate STEM and SMART Grant	4-year	3,270	4,660	142.7
Baccalaureate STEM and no SMART Grant	4-year	2,840	3,100	109.1
Baccalaureate neither STEM nor SMART Grant	4-year	170	280	164.5
Baccalaureate other	4-year	22,120	18,000	81.4
Baccalaureate all	2-year	#	150	#
Other undergraduate SMART Grant	Subtotal	580	1,310	226.1
	Less-than-2-year	#	#	#
	2- to 3-year	10	190	1,745.5
	4-year	570	1,110	196.3
Other undergraduate ACG	Subtotal	3,590	2,520	70.2
	Less-than-2-year	#	20	#
	2- to 3-year	1,780	600	33.4
	4-year	1,800	1,900	105.3
Other undergraduate	Subtotal	89,920	91,820	102.1
	Less-than-2-year	9,310	8,790	94.4
	2- to 3-year	40,390	42,510	105.3
	4-year	40,230	40,520	100.7
Master's	4-year	7,880	6,530	82.8
Doctoral	4-year	3,100	3,760	121.5
Other graduate	4-year	40	470	1,128.6
First-professional	4-year	2,050	1,920	93.9

Rounds to zero.
[1] As expected, the sampling frames misclassified some individual students with respect to baccalaureate, other undergraduate, graduate, and first-professional status; numbers presented in this table are based on the sampling frame classification. The in-state and out-of-state strata have been combined.
[2] Institutional level is based on level confirmed by the school during institutional contacting.
[3] Based on sample allocation and 2004–05 Integrated Postsecondary Education Data System (IPEDS) file enrollment counts.
[4] The student sample was drawn from 1,730 eligible institutions determined to be eligible and providing enrollment lists.
[5] Percentage reported reflects the ratio of "achieved" to "expected."
NOTE: Detail may not sum to totals because of rounding. ACG = Academic Competitiveness Grant; SMART = Science and Mathematics Access to Retain Talent; STEM = science, technology, engineering, or mathematics.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

[14] *Other undergraduate students* are defined as any undergraduate student not classified as a potential baccalaureate student.
[15] *Other graduate students* are defined as any graduate student not classified as a doctoral student.

Table 6. Initial classification of NPSAS:08 student sample, by institutional characteristics and student type: 2008

	Total sample[1]		Student sampling type[2, 3]							
			Bachelor's		Other undergraduate		Graduate		First-professional	
Institutional characteristic	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
All institutions	137,800	100.0	29,470	100.0	95,650	100.0	10,760	100.0	1,920	100.0
Institutional level										
Less-than-2-year	8,820	6.4	#	#	8,810	9.2	#	#	#	0.1
2-year	43,460	31.5	150	0.5	43,300	45.3	10	#	#	0.1
4-year non-doctorate-granting	37,930	27.5	12,950	44.0	22,880	23.9	2,080	19.3	20	1.1
4-year doctorate-granting	47,590	34.5	16,360	55.5	20,650	21.6	8,680	80.7	1,900	98.8
Institutional control										
Public	87,470	63.5	17,600	59.7	64,390	67.3	4,900	45.5	590	30.6
Private not-for-profit	32,760	23.8	10,220	34.7	15,560	16.3	5,650	52.5	1,330	69.3
Private for-profit	17,570	12.7	1,640	5.6	15,710	16.4	210	2.0	#	0.1
Type of institution										
Public less-than-2-year	1,730	1.3	#	#	1,730	1.8	#	#	#	#
Public 2-year	39,340	28.5	100	0.3	39,230	41.0	10	#	#	0.1
Public 4-year non-doctorate-granting	16,120	11.7	5,890	20.0	9,550	10.0	680	6.3	#	0.1
Public 4-year doctorate-granting	30,280	22.0	11,610	39.4	13,880	14.5	4,210	39.2	590	30.4
Private not-for-profit less-than-4-year	2,080	1.5	10	#	2,070	2.2	#	#	#	#
Private not-for-profit 4-year non-doctorate-granting	14,200	10.3	5,550	18.8	7,290	7.6	1,360	12.6	10	0.4
Private not-for-profit 4-year doctorate-granting	16,480	12.0	4,660	15.8	6,200	6.5	4,290	39.9	1,330	69.0
Private for-profit less-than-2-year	6,610	4.8	#	#	6,610	6.9	#	#	#	0.1
Private for-profit 2-year-or-more	10,960	8.0	1,640	5.6	9,100	9.5	210	2.0	#	0.1

Rounds to zero.
[1] The student sample was drawn from 1,730 eligible institutions that provided enrollment lists.
[2] As expected, the sampling frames misclassified some individual students as to baccalaureate, other undergraduate, graduate, and first-professional status; numbers presented in this table are based on the sampling frame classification.
[3] The 10 baccalaureate strata have been combined, the 6 other undergraduate strata have been combined, and the master's, doctorate, and other graduate strata have been combined.
NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

2.2 Sources of Data

Information for the NPSAS:08 study was obtained from several sources:

- *Student records*. Data came from institutional financial aid and registrar records on the student at the sampled institution. Using the CADE instrument (web-CADE), institutional personnel or field data collectors entered the data at the institution in 2008 (self-CADE or field-CADE). Alternatively, institutions could directly download the data to a data file (data-CADE).
- *Student interviews*. Data came directly from web or telephone interviews of sampled students.
- *Central Processing System (CPS)*. Data came from the U.S. Department of Education database of federal financial aid applications for the 2007–08 and 2008–09 academic years.
- *National Student Loan Data System*. Data came from the U.S. Department of Education database of Title IV federal loans, Pell Grants, ACGs, and SMART Grants for records dated before September 2008.
- *National Student Clearinghouse*. Data came from a commercial database for the collection of postsecondary enrollment, degree, and certificate records on behalf of participating postsecondary institutions for the 2007–08 academic year.
- *ACT File*. Data came from the student's most recent ACT score and survey data record for the 2001–02 through 2006–07 academic years.
- *SAT File*. Data came from the student's SAT Reasoning Test (SAT) score and survey data record, as obtained from the College Board for high school graduation years 2002–07.
- *IPEDS*. Institution-level data were obtained from the U.S. Department of Education's NCES 2005–06 database of descriptive information about individual postsecondary institutions.[16]

2.3 Instrument Development

Two data collection instruments were developed for NPSAS:08. The first was designed to collect student-level data from institutional records; the second was a student interview to collect data directly from the sampled student.

2.3.1 Computer-Assisted Data Entry

NPSAS:08 used three modes for student record abstraction: (1) institutions entered data directly into the web-based CADE system (referred to as *self-CADE*); (2) institutions provided student record information in data files according to specifications (*data-CADE*); and (3) trained field data collectors traveled to the institution, abstracted the student record data, and entered them into the web-based CADE system (*field-CADE*).

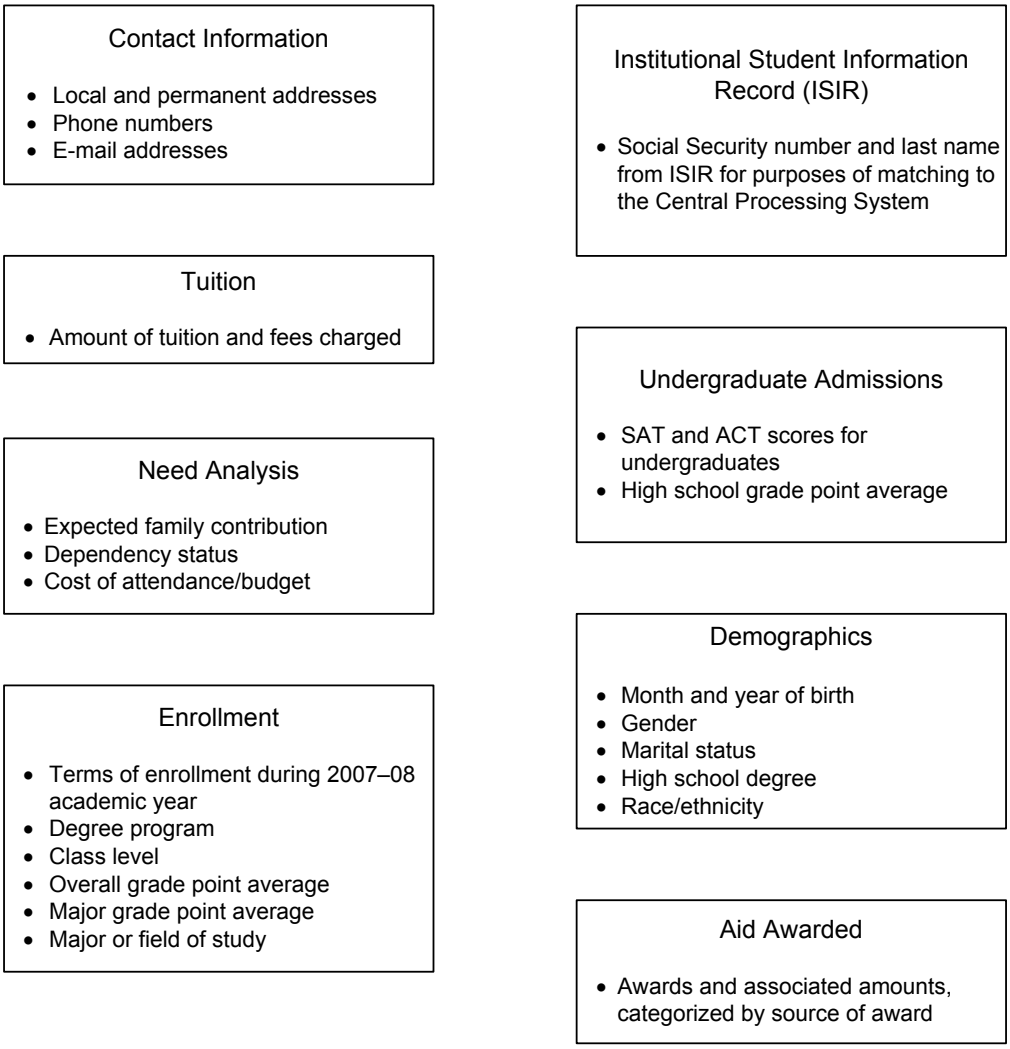
[16] The institutions in the freshened sample include 2006–07 IPEDS data.

The web-based CADE instrument was created with ASP.NET technology combined with a structured query language (SQL) server database. The overall content of the NPSAS:08 CADE instrument resembled that of the instruments used in previous NPSAS cycles (since NPSAS:96) and effectively derived the desired data elements from the institutions. The items from NPSAS:04 that were specific to the Beginning Postsecondary Students (BPS) cohort were removed from the instrument, while items that were necessary for identifying the B&B cohort were included. In addition, a small number of items were added (e.g., major grade point average and Classification of Instructional Programs [CIP] code for major) and deleted.

To increase efficiency, improvements were made to the appearance and navigation of the CADE instrument in NPSAS:08. The NPSAS:04 instrument included multiple screens per section; the NPSAS:08 instrument condensed all the items for each section onto one screen, which decreased the time required to complete each section of the instrument, which in turn decreased burden on institutions. In addition, the student selection and section selection screens were combined. Once a student was selected, the user clicked on links to navigate to other sections to enter data for that student, without returning to a section selection screen. Users had to return to the selection screen only when they began to enter data for other students.

The CADE instrument consisted of eight sections: Aid Awarded, Need Analysis, Institutional Student Information Record, Contact Information, Demographics, Undergraduate Admissions, Enrollment, and Tuition (figure 3; for a facsimile of the instrument, see appendix C). All eight instrument sections were available from the selection page and from every data-entry screen once the user had selected a specific student.

Figure 3. Structure and content of computer-assisted data entry student record abstraction instrument: 2008



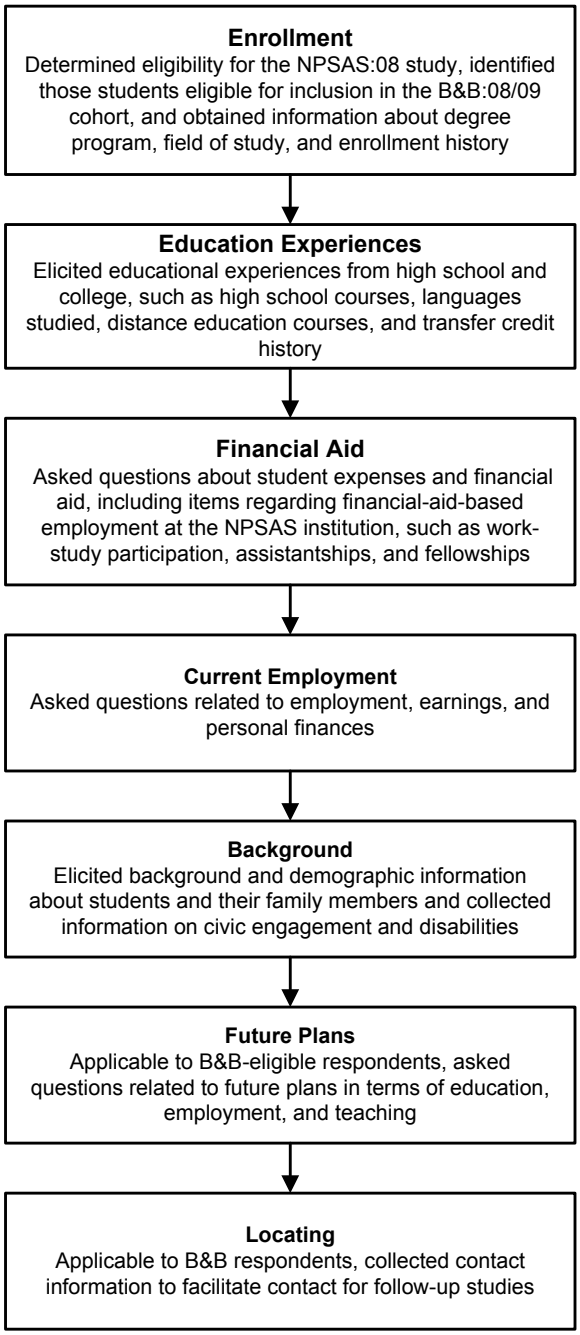
NOTE: Computer-assisted data entry sections were not sequential; users could enter student record information in any order.
SAT = SAT Reasoning Test.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

2.3.2 Student Interview

The NPSAS:08 student interview was developed as a mixed-mode, web-based instrument that respondents accessed either through the Web or by a telephone interview. To determine what items to include, items from past NPSAS cycles were evaluated. Items that were pertinent only to the BPS cohort were excluded, while items relevant to the B&B cohort were included. Also included were new items reflecting policy issues and current topics of interest to researchers. These items were specific to both the B&B cohort and the entire NPSAS:08 sample and included the ACG and SMART Grant items.

The student instrument consisted of seven sections that are displayed in figure 4. A facsimile of the student instrument is presented in appendix C.

Figure 4. Structure and flow of the NPSAS:08 student interview: 2008



NOTE: B&B = Baccalaureate and Beyond; NPSAS = National Postsecondary Student Aid Study.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

A single web-based instrument was used for both web and telephone interviews. Mixed-mode surveys introduce challenges, as well as benefits, not experienced with single-mode surveys. Web interviews enable sample members to complete the survey at their convenience; however, the presence of an interviewer affords the respondent the opportunity to clarify question meanings and affords interviewers the opportunity to probe unclear responses. The mixed-mode, web-based instrument was developed to replicate the benefits of a telephone interviewer in a web-based

interview; the goal was to maintain data quality and still make the process as efficient as possible for respondents. The NPSAS student interview included the following features to accommodate the mixed-mode nature of the survey:

- Questions were worded so that they could, while preserving question integrity, be read either directly by a web respondent or aloud by an interviewer to a telephone respondent.
- Help text was included on every form to provide definitions and to clarify question meaning.
- Verification boxes were displayed whenever respondents entered out-of-range values for items.
- Explicit "don't know" responses were allowed as legitimate responses on only some forms (e.g., parents' income, high school grade point average). For the remaining items, when respondents did not know the answer to an item or wished not to answer an item, they simply left the screen blank and proceeded with the interview.
- To minimize item nonresponse, conversion text was added to selected critical items when a respondent left one of these items unanswered.[17]
- After three consecutive screens with no response, a pop-up box was displayed to encourage participation. The message reiterated the importance of the study and completeness of data, reminded the respondent of the confidentiality of his or her responses, and requested that the respondent complete the items left blank.

Respondents were guided through each section of the interview according to skip logic, which took into account the respondent's characteristics and answers to previous answers (e.g., routing them past questions inapplicable to them). Efficient skip logic was crucial to minimizing respondent burden. The interview also incorporated previous answers (e.g., school name, major or field of study) into question wording to further customize the respondent's interview and increase clarity.

The NPSAS:08 student interview incorporated the use of coding systems to standardize the collection of some information, such as postsecondary institutions attended and major or field of study. For the postsecondary-institution coding system, a database was constructed from the set of institutions in the 2005–06 IPEDS, which was developed by NCES (http://nces.ed.gov/ipeds/). Respondents entered the city, state, and name of the institution into a text box. Once the information was submitted, the coder searched the database and returned a list of possible matches. The respondent or interviewer then selected the correct institution. If the respondent did not find a matching institution, the text string entered was retained, but the respondent was asked to supply the level and control of the institution.

The coding system for major or field of study worked much the way the institution-coding system worked. This system used a database constructed from the 2000 CIP taxonomy, also developed by NCES (http://nces.ed.gov/pubs2002/cip2000/). Respondents entered a major or field of study into a text box. The coder then conducted a search of the database, using the keywords entered, and provided a list of possible matches. The respondent selected the correct

[17] *Conversion text* is text that appears on the screen after the respondent left the form blank initially. After selecting next, the item reappears with the additional text attempting to convince the respondent to answer the item. For more information regarding conversion text, see section 4.3.6.

major or field of study or, if no areas matched, manually coded the major with the use of two dropdown boxes. Respondents selected a general major category from the first dropdown box and then selected a specific major category from a second dropdown box.

Once programmed, the student instrument was tested multiple times to ensure that each component was working properly. Both project staff and NCES staff used mock cases to evaluate the skip logic, question wording, screen layout, and efficiency of the instrument for the various student profiles expected to occur in the sample. Furthermore, testing was conducted at various times of the day, using different internet connections and browsers, to ensure that respondents would not encounter problems in accessing and completing the NPSAS:08 student instrument.

An abbreviated interview was developed for two primary purposes: (1) to convert nonrespondents and (2) to provide a Spanish-language translation. The abbreviated interview consisted of the first section, which included items related to eligibility and enrollment. This interview was used as a refusal-conversion tool at the end of the data collection period. It was translated into Spanish so that bilingual telephone interviewers could administer it to Spanish-speaking respondents. All components of the abbreviated interview, including help text and conversion text, were translated into Spanish to maintain consistency across languages.

2.3.3 Evaluation of Data Quality

Throughout the entire NPSAS:08 data collection process, all components were carefully monitored for quality assurance and continual improvement. Debriefings were held, and many evaluations were conducted during each stage of data collection to ensure continued excellence and efficiency (table 7).

Table 7. Summary of NPSAS:08 evaluations: 2008

Major area of evaluation	Evaluation approaches
Training for data collection	Debrief field abstractors.[1] Debrief CATI staff.[1]
Enrollment list acquisition	Analyze overall response rate, accuracy, and time to produce lists.
Student record abstraction	Analyze overall outcomes, including institutional participation, nonresponse, and refusal. Analyze data quality (missing data) under conditions of web-CADE, field-CADE, and data file production approaches. Debrief institutional coordinators.[1] Debrief field staff.[1]
Student tracing and locating activities	Debrief tracing staff and supervisors.[1] Analyze all sources and levels of tracing results and costs.
Student interviewing	Analyze quality control monitoring data. Analyze CATI operational parameters (e.g., numbers of calls per case, total interviewer hours per completed interview). Analyze interview response burden, overall and by section. Debrief interviewers, monitors, and supervisors.[1] Analyze response rates and patterns of interview nonresponse. Analyze impact of financial incentive on response rate. Analyze effectiveness of various strategies for handling answering machines.
Nonresponse bias analysis	Analyze nonresponse bias at the following levels: institutional, student, and item.

[1]Informal debriefings of staff involved in different data collection tasks were conducted throughout the study. Information gathered through these debriefings was used to enhance understanding of the outcomes of more formal evaluations and is therefore not described separately in this report.
NOTE: CADE = computer-assisted data entry; CATI = computer-assisted telephone interviewing.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

2.4 Data Collection

2.4.1 Contacting Institutions

The institution sample for NPSAS:08 consisted of 1,960 institutions, 1,940 of which were eligible to participate in the study. In addition, in six states—California, Georgia, Illinois, Minnesota, New York, and Texas—institutions were oversampled to provide state-level representative data. Each of the oversampled states in NPSAS:08 was part of the 12-state oversample selected for NPSAS:04.

Before institutions were contacted, endorsements were obtained from 26 major professional associations (for the complete list of endorsements, see appendix D). This list included organizations that had previously endorsed NPSAS, as well as new organizations, such as the State Higher Education Executive Officers and the Western Interstate Commission for Higher Education. Regional organizations in the oversampled states and organizations representing schools in the for-profit and less-than-2-year sectors were specifically targeted for inclusion as endorsing organizations. Endorsements were featured on all project letterhead, on all pamphlets, and on the NPSAS:08 website. In addition, several of these organizations continued to promote the study throughout the data collection period in newsletters and other communications.

Early notification of sampled institutions had previously been found to have been effective in encouraging institutions to meet the requested deadlines in the previous round of NPSAS

(NPSAS:04). The early notification allowed schools time to arrange resources for participation within the schedule constraints of the study and time for any required internal review and approval procedures. The time also allowed institutions to address any potential obstacles to their participation. In accordance with these considerations, the start of the fall term was judged ideal for initiating contact with the institutions. The process of recruiting institutions and initiating coordinator contacts for NPSAS:08 began in September 2007. The earliest due dates for providing student lists occurred in January 2008. (Sample augmentation slightly delayed the initiation of contact with some institutions added for state representativeness.) For NPSAS:08, the collection of student enrollment lists proceeded in four stages: (1) training, (2) initial contact, (3) recruitment of institutions, and (4) collection of student lists.

Training. Institutional contactors were trained before the start of institutional recruitment on their role and responsibilities for NPSAS:08. During training, they received an overview of the study, detailed telephone scripts and frequently asked questions, and guidance on instructing institutions on how to complete the student enrollment list. In addition, institutional coordinators were trained in how to effectively gain the cooperation of institutions.

Initial contact. The first step in the institutional contacting process was to verify the contact information for the chief administrator, to whom the initial mailing would be directed. Institutional eligibility was also confirmed at this time. Institutions flagged as potentially ineligible—including closed institutions and institutions that indicated they were not Title IV eligible or open to the general public—were reviewed by project staff. Instances of sampled institutions that had merged with other institutions (sampled or not sampled), possible mission changes that might have affected the institution's sampling strata, and changes in name or address were also reviewed.

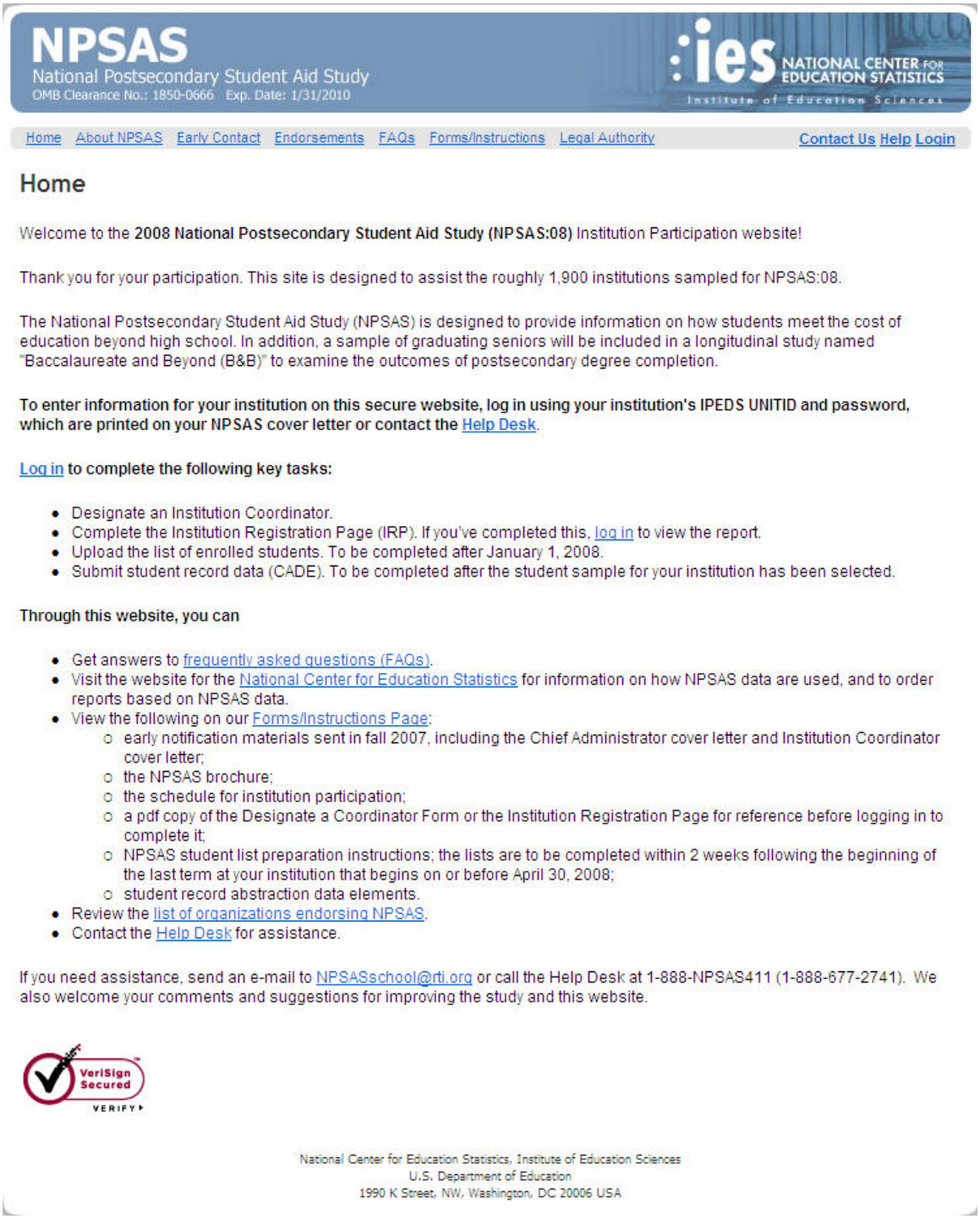
Recruitment of institutions. The recruitment of institutions began in September 2007. Chief administrators at sampled institutions for NPSAS:08 were sent the following materials (for a copy of all NPSAS:08 data collection materials, see appendix E):

- a cover letter, printed on NCES letterhead, providing background information on NPSAS and requesting that the chief administrator designate an institutional coordinator;
- a NPSAS:08 brochure summarizing the objectives of NPSAS and providing background information and selected findings;
- a copy of the NPSAS brochure sent to sampled students; and
- a flowchart containing the projected timeline of activities for NPSAS:08.

A mailing to the chief administrator's office requested the appointment of an institutional coordinator whose role would be to respond to requests for data and coordinate data production and delivery. Institutional contactors followed up with the chief administrator's office and obtained the name of an institutional coordinator by telephone when the institution had not logged into the website to designate one. Telephone prompts were supplemented with e-mail and, when necessary, mail prompts. Once an institutional coordinator was designated, he or she received a letter introducing the study and the role, together with the same brochures and time line sent to the chief administrator.

A website was developed to be used by institutions selected for participation in NPSAS:08 (figure 5). In addition to providing general information about NPSAS:08, this website served as a central repository for all study documents and instructions. Additionally, the institutions were able to securely upload student enrollment lists at the site.

Figure 5. The NPSAS:08 institutional website home page: 2008



NPSAS
National Postsecondary Student Aid Study
OMB Clearance No.: 1850-0666 Exp. Date: 1/31/2010

ies NATIONAL CENTER FOR EDUCATION STATISTICS
Institute of Education Sciences

Home About NPSAS Early Contact Endorsements FAQs Forms/Instructions Legal Authority Contact Us Help Login

Home

Welcome to the **2008 National Postsecondary Student Aid Study (NPSAS:08)** Institution Participation website!

Thank you for your participation. This site is designed to assist the roughly 1,900 institutions sampled for NPSAS:08.

The National Postsecondary Student Aid Study (NPSAS) is designed to provide information on how students meet the cost of education beyond high school. In addition, a sample of graduating seniors will be included in a longitudinal study named "Baccalaureate and Beyond (B&B)" to examine the outcomes of postsecondary degree completion.

To enter information for your institution on this secure website, log in using your institution's IPEDS UNITID and password, which are printed on your NPSAS cover letter or contact the Help Desk.

Log in to complete the following key tasks:

- Designate an Institution Coordinator.
- Complete the Institution Registration Page (IRP). If you've completed this, log in to view the report.
- Upload the list of enrolled students. To be completed after January 1, 2008.
- Submit student record data (CADE). To be completed after the student sample for your institution has been selected.

Through this website, you can

- Get answers to frequently asked questions (FAQs).
- Visit the website for the National Center for Education Statistics for information on how NPSAS data are used, and to order reports based on NPSAS data.
- View the following on our Forms/Instructions Page:
  - early notification materials sent in fall 2007, including the Chief Administrator cover letter and Institution Coordinator cover letter;
  - the NPSAS brochure;
  - the schedule for institution participation;
  - a pdf copy of the Designate a Coordinator Form or the Institution Registration Page for reference before logging in to complete it;
  - NPSAS student list preparation instructions; the lists are to be completed within 2 weeks following the beginning of the last term at your institution that begins on or before April 30, 2008;
  - student record abstraction data elements.
- Review the list of organizations endorsing NPSAS.
- Contact the Help Desk for assistance.

If you need assistance, send an e-mail to NPSASschool@rti.org or call the Help Desk at 1-888-NPSAS411 (1-888-677-2741). We also welcome your comments and suggestions for improving the study and this website.

VeriSign Secured VERIFY

National Center for Education Statistics, Institute of Education Sciences
U.S. Department of Education
1990 K Street, NW, Washington, DC 20006 USA

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08) website.

Visitors to the home page were provided the following links:

- *About NPSAS* linked to information on the study's mandate and research objectives, with a link provided to NCES reports from previous study cycles.

- *Early contact* linked to information about the initial stage of institutional contacting for NPSAS:08.
- *Endorsements* linked to a complete list of all 26 endorsing organizations.
- *Forms/Instructions page* linked to lists of the institutional forms required for participation in the NPSAS:08 study and the instructions for completing them.
- *Frequently asked questions (FAQs)* linked to questions and answers concerning all stages of data collection for both components of NPSAS:08.
- *Legal authority* linked to the legislation that mandates and authorizes the data collection effort for NPSAS and that protects the confidentiality of the data.

The webpage contained a status screen that indicated the stages of institutional data collection that were completed (denoted by a check mark). This allowed institutions to easily recognize and select those stages that were not yet completed.

Once designated, the institutional coordinator was asked to complete an Institution Registration Page (IRP). The form asked for basic information about the institution, such as term structure, and confirmed the deadline for providing a student list, which was initially set for 2 weeks after the start of the term which included the date April 30.[18] The IRP also gave the institutional coordinator the opportunity to review the data elements requested on the list and to alert project staff to any potential difficulties (e.g., institutional prohibitions against releasing certain items or policies requiring an institutional review board [IRB] review of data collection).

Collection of student lists. Calls to institutional coordinators were conducted to prompt timely completion of the IRP and delivery of the student enrollment lists. Institutional coordinators also received telephone, e-mail, and mail prompts reminding them of upcoming deadlines. A help desk telephone number and e-mail address were also provided for institutions that required assistance on project tasks.

On request, institutions were provided with a letter affirming RTI's IRB had approved the study. Institutional review at the institutional level was not required for participating institutions because RTI had received approval from its own IRB; however, supporting materials were provided to institutions that wanted to conduct their own review before participating. A customized packet was prepared for each requesting institution and mailed to the coordinator. The packet included copies of questionnaires and complete descriptions of relevant survey procedures, including information about confidentiality and informed consent.

Because the six-state oversample was selected after the field test data collection, 20 institutions within the six-state oversample were sampled for both the field test and the full-scale studies. After initial contact, mailings alerted the schools to their inclusion in the full-scale sample, and project staff further addressed any concerns about burden by offering support and assistance and by working with the institution on a realistic schedule.

Both state systems and large for-profit systems with centralized record keeping were given the option of reporting for their constituent institutions, whenever feasible, at the system level. This provision greatly increased the efficiency of data collection and reduced burden for individual institutions. The project coordinator worked with these systems directly to manage any unique requests.

[18] The date April 30 was selected to include virtually all students enrolled prior to the summer term without delaying data collection.

2.4.2 Student Enrollment List Acquisition

Complete instructions for preparing the student list were provided for the institutional coordinators on the NPSAS website; instructions were clarified in follow-up telephone conversations as necessary. In those subsequent telephone contacts, contractor staff worked closely with the institutional coordinator to determine the most reasonable list of student information that could be provided by the institution.

Due dates for providing the enrollment list of students for NPSAS:08 were based on the term structure of each institution. Institutions were encouraged to submit an electronic list by uploading it to the secure website. The data items requested for each student were as follows:

- full name;
- institutional student ID;
- Social Security number;
- educational level (undergraduate, master's, doctoral, other graduate, or first-professional status);
- class level of undergraduates;
- baccalaureate indicator (a *baccalaureate recipient* was defined as one who received or would potentially receive a baccalaureate degree during the NPSAS year [July 1, 2007, and June 30, 2008]);
- major;
- six-digit CIP code for major;
- date of birth;
- indicator of student's being older or younger than age 18;[19]
- local address;
- local telephone number;
- campus e-mail;
- permanent address;
- permanent telephone number; and
- permanent e-mail.

Because of schedule constraints on data collection, data processing, and file development, student record abstractions (CADE) and student interviews were conducted concurrently, and student-locating data were collected from institutions when the student enrollment lists used for sample selection were collected. These arrangements allowed both student record abstraction and student interviewing to begin almost immediately after sample selection. From the few institutions unwilling to provide locating data for all students on their enrollment lists, locating data were requested for just the sample students immediately after the sample was selected.

[19] The indicator of a student's being older or younger than age 18 was requested when the student's date of birth was not provided.

Date of birth was requested on the lists to identify minors so that those who were sampled would not be interviewed until age 18. If the date of birth indicated that the student would not be 18 years old by the end of the academic year (June 30, 2008), then the student was assumed to be concurrently enrolled in high school and therefore ineligible.

In prior NPSAS studies that identified a B&B cohort, lists of potential baccalaureate recipients were collected with the student list of all enrolled undergraduates and graduates/first professionals. Unfortunately, these baccalaureate lists often could not be provided until late in the spring or in the summer, after baccalaureate recipients could be positively identified. To help facilitate earlier receipt of lists, the enrollment lists for 4-year institutions included an indicator (B&B flag) of students who had received or would potentially receive a baccalaureate degree during the NPSAS year (*yes, no, don't know*). Institutions were instructed to make this identification before spring graduation to avoid delaying provision of the lists. The enrollment lists for 4-year institutions also included an indicator of class level for undergraduates (*first year, second year, third year, fourth year,* or *fifth year*) that was used when the B&B flag was not provided for any students. These two indicators were used instead of requesting a separate list of baccalaureate recipients.

Because a disproportionately large number of baccalaureate recipients are business majors, major field and CIP code were requested on the lists to allow for undersampling of business majors. CIP codes were also used to identify and oversample baccalaureate recipients who were STEM majors; this group is of analytic interest in B&B.

Follow-up with institutional coordinators was conducted by telephone, mail, and e-mail. Telephone prompts to the institutional coordinators were made for institutions that had not provided lists. E-mail reminders that encouraged participation were sent to institutional coordinators when deadlines were pending. A refusal-conversion letter was mailed to institutions that had not responded.

As enrollment lists were received, they were reviewed for completeness, readability, and accuracy. At a minimum, the lists had to include locating information for each student, either specific contact information or a Social Security number that could be used in the locating process. Before actual student sampling, several checks were implemented on quality and completeness of provided student enrollment lists. Institutions providing lists that failed at least one of these checks were called so that detected problems could be rectified. Completeness or quality checks failed if any of the following conditions existed:

- education level—undergraduate, master's, doctoral, other graduate, or first-professional status—was not included or was unclear;
- baccalaureate recipients or graduating seniors were not identified (unless the institution was less than 4 year, was a graduate school, or explicitly indicated that no such students existed in the school);
- major fields of study or CIP codes were not clearly identified for baccalaureates; or
- the number of students listed was inconsistent with the most recent IPEDS data.

Several checks on quality and completeness of student lists were implemented before the sample students were selected. Institutions providing lists that failed these checks were contacted so the detected problems could be resolved. Enrollment lists failed quality control checks under the following conditions:

- baccalaureates were not identified in 4-year institutions (unless the institution enrolled only graduate/first-professional students or explicitly indicated that no baccalaureates existed in the school); or
- student level—undergraduate, master's, doctoral, other graduate, or first professional—was not clearly identified.

Quality checks on counts for baccalaureate students were performed separately from checks on counts for the rest of the students. The unduplicated potential baccalaureate counts were checked against the unduplicated counts of bachelor's degrees awarded according to the IPEDS Completions Survey. The check failed if the count for any unduplicated list was 50 percent less than or 75 percent greater than the IPEDS count. The list counts were expected to exceed the IPEDS counts because the list counts were of expected baccalaureate recipients, while the IPEDS counts were of actual baccalaureate recipients. The unduplicated counts of other undergraduates, graduates, and first-professional students were checked against the unduplicated annual enrollment counts from the IPEDS Fall Enrollment Survey.[20] The check failed if the count for any unduplicated list differed by at least 50 percent from the IPEDS count.

Sometimes state postsecondary systems and private multicampus institutions were able to provide enrollment lists for all their sampled institutions from a central office. In these instances, a *lead institution* was appointed, and a coordinator was designated to report for all sampled institutions. System-wide offices also provided other data collection assistance, and some institutions—particularly within the six oversampled states—actively encouraged their campuses to participate. More than 250 institutions reported as part of a multicampus system.

A point of contact was identified in each of the six states with representative samples of undergraduates at the state level. These individuals were regularly updated on the participation status of institutions within their states. They also assisted with ongoing efforts to encourage institutional participation by contacting the chief administrators and institutional coordinators at sampled institutions.

2.4.3 Administrative Data for Sample Members

A portion of the student data for the NPSAS:08 was obtained from various administrative databases, including two U.S. Department of Education databases: the CPS and the National Student Loan Data System (NSLDS). Other data sources included the National Student Clearinghouse (NSC), ACT, and the College Board. These diverse, sometimes overlapping, data sources provided information that often could not be reliably or directly collected from institutions or students. They also provided a way to complete certain data that were also gathered by means of student record abstraction or the student interview but were missing for individual sample members (e.g., demographics, characteristics, enrollment information). Overlapping data sources sometimes served to check or confirm the accuracy of similar information from other sources.

Central Processing System. To reduce institutional burden, information related to student applications for federal financial aid was obtained from the CPS. Students enter financial status information about themselves and their families on the U.S. Department of Education's Free Application for Federal Student Aid (FAFSA) form. The information is then converted to an electronic format, analyzed, and provided to requesting institutions and other approved parties. As in NPSAS:2000 and NPSAS:04, CPS assigned RTI a "special designation code" allowing access to

[20] The IPEDS Fall Enrollment Survey collects fall enrollment counts as well as an unduplicated annual enrollment count.

the FAFSA data. Under this procedure, financial aid application data were requested through a standard and secure online Federal Data Request process. The CPS was accessed daily to download the requested 2007–08 data as the student sample was selected. At the conclusion of data collection, the CPS was accessed to obtain updated information for the 2007–08 and 2008–09 academic years.

National Student Loan Data System. Student-level data on the type and amount of Pell Grants, ACGs, SMART Grants, and federal student loans were obtained from the NSLDS database. The electronic data interchange with NSLDS was performed twice during the data collection period; it was performed once again after data collection had ended, in order to submit the most up-to-date data possible for matching and to obtain the most up-to-date federal grant and loan data. A successful match with the NSLDS database required that the student have a valid application record in the database. The accessed NSLDS data files included both information for the year of interest and a complete federal grant or loan history for each student. The data transfer was secured through an NCES system that required a login and used Secure Sockets Layer (SSL) technology.

National Student Clearinghouse. Data on institutions attended, enrollment dates, and degree completions for the student sample were obtained from the NSC StudentTracker service. An individual student record would match with the NSC only if the student's institution was a participant in the NSC. The StudentTracker request for data was performed once toward the end of data collection to obtain the most updated personally identifying data for the match. Multiple records per student matched were returned, including historical records, although the period of interest for NPSAS was only the 2007–08 academic year. An account was established with NSC that permitted secure delivery and receipt of files over encrypted Secure File Transfer Protocol connections.

ACT. To obtain valuable admissions test data, a file merge with ACT was performed. Student ACT scores and survey data from the most recent test record between the 2001–02 and 2006–07 academic years were received. This data match was performed once toward the end of data collection in order to send the most updated personally identifying data to use as matching criteria. As with NSLDS, the data transfer was secured through an NCES system that required a login and used SSL technology.

SAT Reasoning Test. To obtain SAT test scores and questionnaire data, a file merge with the College Board was performed. Recovered database records spanned high school graduation years 2002–07. If multiple test records were found in the database, only the most recent record was returned. As with ACT, file merging was performed toward the end of data collection, and the file transfers were secured through an NCES system that required a login and used SSL technology.

2.4.4 Data Abstraction From Student Records

Pre-data-collection and support activities. The first step in the CADE record abstraction process involved sending the student sample to CPS to obtain student financial aid application data. On completion of CPS matching (typically a 24-hour turnaround), a number of data elements retrieved from CPS were preloaded into the CADE database, initializing the CADE system for each institution. These preloaded data elements included an indicator of whether the student had been matched successfully to the CPS system, as well as selected CPS variables for use in CADE software edit checks. In addition, to assist in identifying common types of financial aid received by students, the CADE system was customized for each institution with preloaded names of institutional financial aid programs and up to 12 state financial aid programs.

Once CADE was initialized for a particular institution, an informational packet on the CADE system was sent to the designated institutional coordinator. These packets included instructions for accessing the NPSAS:08 institutional website and a list of the data elements. In addition, call center staff called to notify institutions that the CADE data collection had begun. Referencing daily status reports that summarized the progress made by institutions that had selected self-CADE and data-CADE, calls were made periodically to the institutional coordinators of schools that had not yet completed student record abstraction. Institutions using the field-CADE option were notified by mail and contacted by the field data collector, at which time an appointment was made for the field data collector to visit the institution.

Packets sent to self-CADE institutions included the *NPSAS 2008, National Postsecondary Student Aid Study Web-CADE User's Guide*, which included complete specifications, instructions, and system requirements for self-CADE submission. The user's guide also discussed the study's confidentiality procedures and included information on alternative methods of data submission. Packets sent to data-CADE institutions included specific instructions on how to construct the requested data files.

The NPSAS:08 institutional website allowed institutional staff to access an electronic list of sample members from their institution, which enabled the institution to create programs to provide the requested data for the sampled students. Several system features of the website—including help text and an e-mail generator for problem reports—assisted institutions with data entry. To help address any institutional staff questions or problems at any point in the student record abstraction process, a help desk number was established. The help desk ensured that institutional and project staff worked together to correct any data collection issues before submission of student record data.

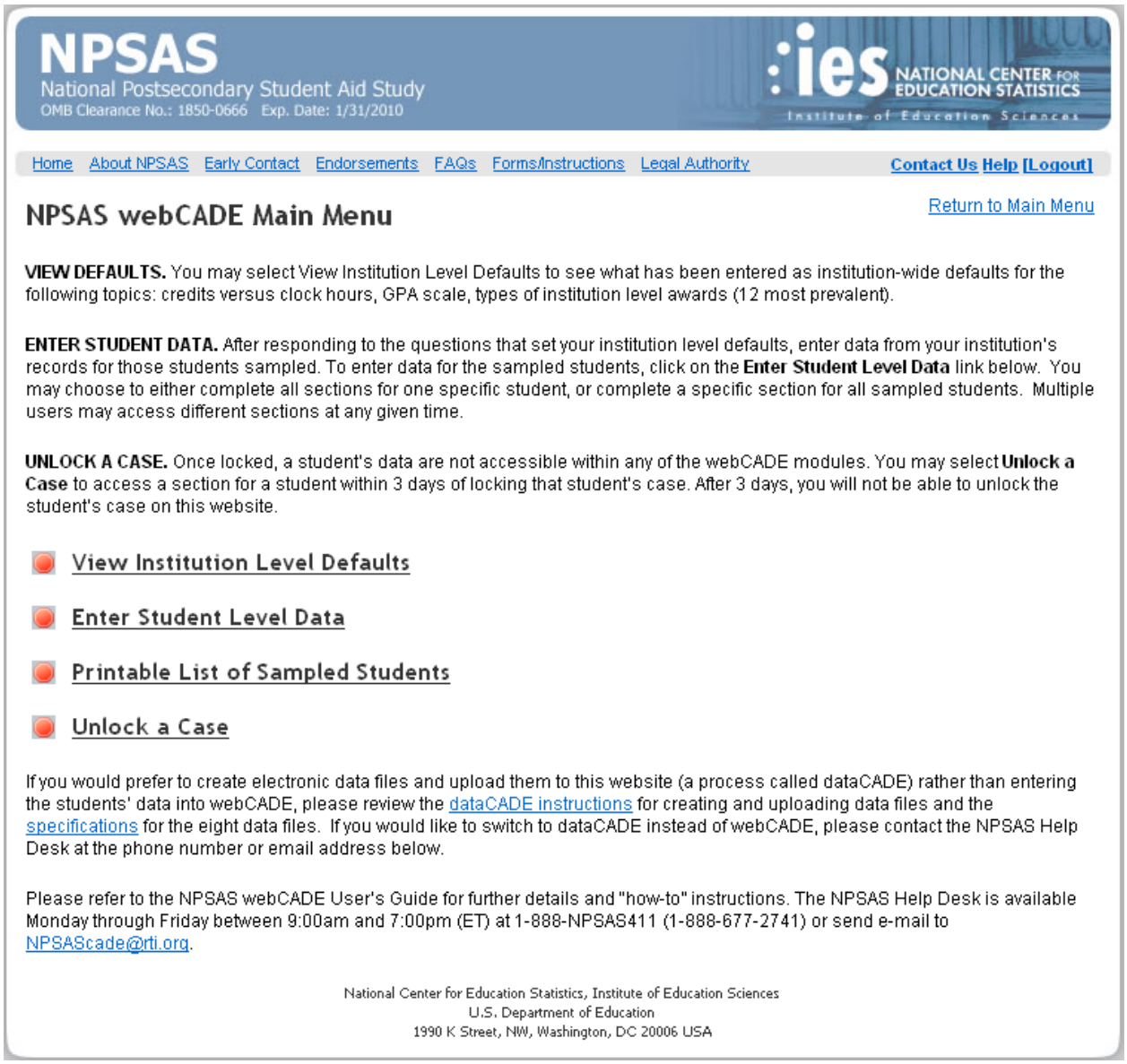
Training for field staff who were involved in abstracting data and entering data into field-CADE for institutions included the following: a description of the study's objectives and schedule, an explanation of how the financial aid process worked on campuses, an outline of procedures for working with the institutional coordinator and other staff at the institutions, and instructions for locating records. The training also included a review of, and practice with, each section of the CADE instrument and electronic transmission of completed cases. Procedures for contacting supervisors and other administrative tasks were discussed. A laptop computer was provided to each trainee for use during training and subsequent field work. Field interviewer training also emphasized the importance of data security and procedures to maintain confidentiality.

Data collection. Institutional record data for sampled students were collected according to procedures resembling those successfully tested and implemented during NPSAS:04. As mentioned, institutions could choose one of three modes of student record abstraction: (1) self-CADE, in which institutional staff entered data directly into the web-based system; (2) data-CADE, or provision of student record information in data files, according to specifications; or (3) field-CADE, in which trained field data collectors used laptops to abstract student record data into the CADE system.

Visitors to the web-CADE web page (i.e., institutional coordinators) were first asked to provide institution-specific information on their use of clock or credit hours, their grade point average scale, and the names of institutional grants and scholarships. After completing these items, the institutional coordinators entered all data for each student by clicking on the *Enter Student Level Data* link (figure 6). The student data could be entered in one or multiple sessions. Finally, the user locked each complete case to indicate that it was ready for processing. If cases were locked in error, the user could unlock a case, so long as it had not been locked for more than 3 days. (After 3 days,

the user would have to call the help desk for any data changes.) The website also provided the help desk telephone number and e-mail address.

Figure 6. NPSAS:08 web-CADE main menu: 2008



SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08) website.

As an alternative to keying data into the web-based CADE application (self-CADE), institutions—particularly those with large sample sizes and system reporters—were given the option of submitting data files containing student record data. This method of data abstraction was first used in NPSAS:2000. File specifications for uploading comma-separated or delimited flat files were provided to institutions that chose this option (for a copy of these file specifications, see appendix F). Data-CADE consisted of eight data files (including student-level, term-level, and aid award–level files) that were similar in structure to the database underlying the web-based CADE application (self-CADE). On completion of the data-CADE file preparation, institutions submitted

their data files through a secure server on the NPSAS:08 institutional website. On submission, an automated quality control system processed the files and instantly reported to the institutions any problems in the data (e.g., incorrect student identification variables, lack of term-level data for sample students, incorrect file names) so they could be corrected and resubmitted.

Institutions were also given the option of having a field data collector visit and provide student record data-entry services at no expense. Field data collectors used laptops, with a local version of the web-based CADE application, to enter data abstracted from student records. All features in the web version were present in the laptop version, including real-time edit features to help detect out-of-range or inconsistent entries. In addition, data previously obtained from CPS were preloaded into the system before data collection began. On completion of data entry, the field data collectors transmitted the data to the same database used by the web-based CADE application, maintaining all completed student records in one location.

2.4.5 Student Interviewing

Staff training. The mixed-mode design of NPSAS:08 data collection required the development of four separate training programs: supervisor training, help desk training, telephone interviewer training, and tracer training. Separate training sessions were held for each of these groups (table 8).

Table 8. NPSAS:08 training sessions for interviewing and tracing staff: 2008

Staff trained	Number of training sessions[1]	Time period	Number of staff trained
Supervisors	3	January–July 2008	26
Help desk agents	1	February 2008	19
Telephone interviewers	15	March–August 2008	482
Tracing supervisors and tracing specialists	3	March–May 2008	72

[1] Number of training sessions indicates the number of training sessions held for unique groups of trainees.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

- *General training.* Before attending project training, telephone interviewers completed 4 hours of comprehensive general training, which outlined call center rules and procedures, sample member refusal-conversion techniques, and an overview of the CATI-CMS. The general training included an overview of survey research, an introduction to survey data collection methods, and responsibilities of telephone interviewers, including confidentiality requirements and the rights of survey respondents. During their general training, telephone interviewers also learned general interviewing skills, such as enunciation and maintaining the proper pace. Telephone interviewers completed their general training either in the Raleigh, North Carolina, Call Center with a supervisor, or online with the use of an interactive website (for a list of general training modules, see appendix G).
- *Supervisor, help desk, and telephone interviewer training.* The interviewer training program was designed to maximize the trainees' active participation. Materials included a training agenda, an interviewing manual, a list of FAQs, a pronunciation guide, and handouts featuring help text for selected NPSAS questionnaire forms. Interviewer training sessions consisted of lectures, demonstrations, and hands-on practice exercises with the student

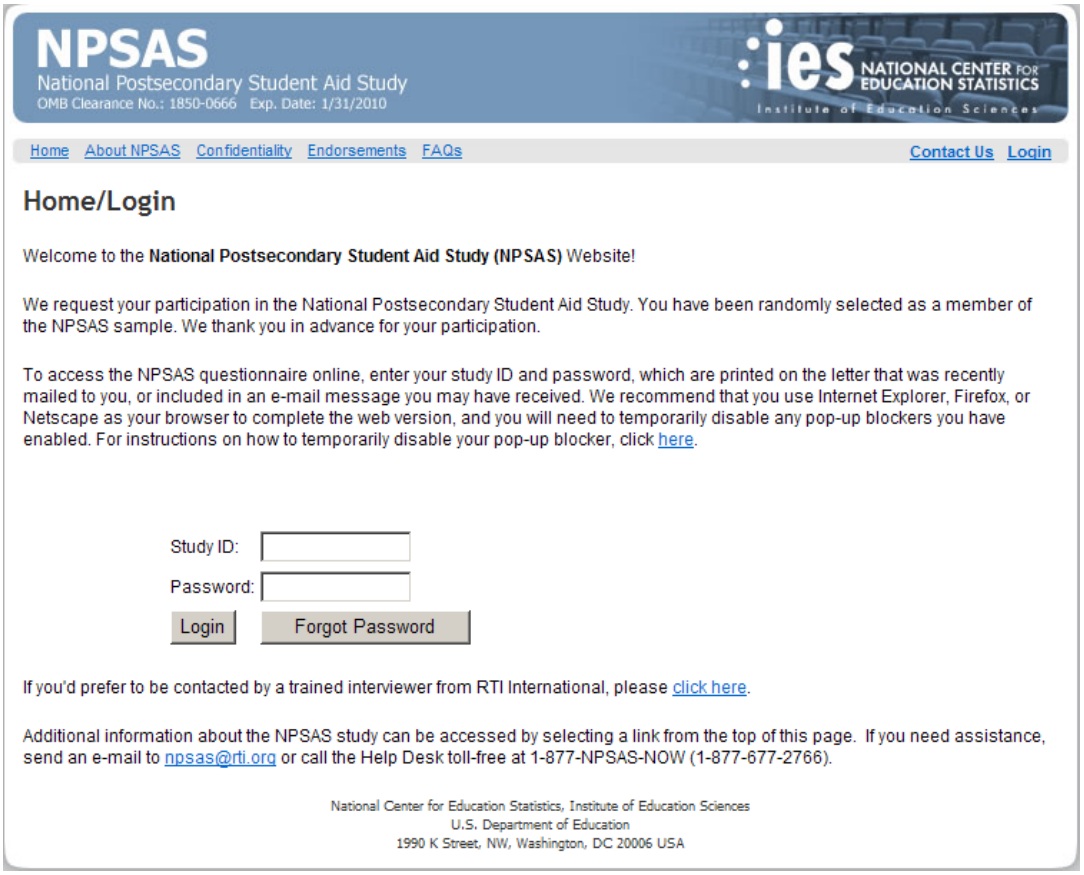
instrument and online-assisted coders. Sessions for supervisors, help desk staff, and telephone interviewers included a study overview, an introduction to data collection procedures, a review of the confidentiality requirements, a demonstration interview, an in-depth review of the instrument, a review and discussion of active listening techniques, hands-on practice exercises with the instrument and open-ended coders. Additionally, interviewers were trained to use the CATI-CMS and were given opportunities to practice in CATI and to ask questions. Help desk agents received the same training as telephone interviewers but also received specialized training on the project-specific help desk application and on solutions to common problems. The help desk application was designed to assist agents in documenting problems reported to the study help desk hotline (for a copy of the telephone interviewer training agenda and the training manual's table of contents, see appendix G).

- *Tracer training*. New tracers received 12 hours of classroom training and 4 to 8 hours of additional mentoring when they arrived on the production floor. The training included information on RTI's history and devoted time to the protection of human study subjects. To receive permission to conduct tracing activities, tracers had to pass certification requirements for protection of study subjects and their information. Training included an introduction to tracing methods, an overview of confidentiality and security matters in the protection of sample members, training in the use of proprietary databases, and instruction on the use of the tracing CMS. Tracers were also trained to use probing to obtain information helpful in locating sample members (for a copy of the tracing training agenda, see appendix G).
- *Additional training*. In addition to general sessions, help desk agents and select telephone interviewers received a brief training on conducting calls to prompt randomly selected sample members who had not completed the interview by a specified date to do so. Other additional training modules, such as refusal training, were conducted as necessary. Additionally, quality circle meetings were conducted as an extension of the training program and for continual quality improvement. Telephone interviewers and supervisory staff were given the opportunity to ask questions, and as needs were identified, additional training topics were highlighted and addressed in meetings and subsequent posted quality circle notes.

Student website. The student website for NPSAS:08 provided background and general information about the study, the sample, the study sponsor, the study contractor, and confidentiality assurances (figure 7). In addition to the information available on the website, links to other relevant websites (e.g., NCES and RTI) were provided. The student website also provided a way for sampled students to securely log in and complete the interview from the NPSAS home page.

The initial login page provided the link to the web instrument. The login required entry of a specific study username and password, which were provided in the lead letter and subsequent e-mail messages. Sample members could also obtain their study username and password by clicking on the *Forgot Password* button (to send a request to the project's e-mail address) or by contacting a help desk agent at the NPSAS toll-free number. The initial login page also provided a way for respondents to submit their telephone number so they could be contacted by an interviewer.

Figure 7. NPSAS:08 student website: 2008



SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08) website.

The student website also presented a variety of information about the study. Sample members visiting it were provided the following links from the home page:

- *About NPSAS* linked to information on the study's mandate and research objectives, with a link provided to NCES reports from previous study cycles.
- *Confidentiality* linked to information regarding the regulations in place to protect sample members' confidentiality.
- *Endorsements* linked to a complete list of all 26 endorsing organizations.
- *FAQs* linked to questions and answers concerning the NPSAS:08 study.

Web interviews. Sample members were first introduced to the web interview through the lead letter packet. During the first 3 weeks after a sample member's invitation to participate, he or she was able to complete a web interview or to call the help desk to complete a telephone interview. The begin and end dates for this initial period varied because the sample was selected on a flow basis. A subset of sample members was selected to receive a prompting telephone call approximately 10 days after the beginning of their initial 3-week data collection period. The determination of which sample members were included in the prompting sample was made with the use of paradata (information that has been collected about the data collection process) to target cases as appropriate (for more details on prompting, see section 3.5.3). The prompting call was to remind sample

members that they had been selected to participate in the NPSAS:08 study and to encourage them to log in to the study website and complete the web interview during the early completion period, while they were still eligible to receive the $30 incentive.

The website was accessible 24 hours a day, 7 days a week, throughout the data collection period, enabling sample members to complete interviews online at their convenience. E-mail prompts periodically were sent to all sample members (for whom a valid e-mail address was available) during the 3-week period to encourage participation and remind them of the address for the study website and the toll-free help desk phone number. Sample members were assured of the confidentiality of their responses and the voluntary nature of the survey, and they were told that they could decline to answer any survey question. The convenience of the web interview was emphasized, especially that it could be completed any time from any location with internet access and that respondents could discontinue and resume the interview as necessary.

Help desk operations. Anticipating the first student calls after the introductory mailing, the NPSAS:08 help desk opened in conjunction with the first student mailout so as to provide support to web respondents. Help desk staff were available to assist sample members who had questions or problems accessing and completing the web interview. A toll-free help line was established to accept incoming help desk calls. If technical difficulties prevented sample members from completing the web interview, help desk agents, who were also trained to conduct telephone interviews, would encourage sample members to complete a telephone interview rather than attempt the web interview.

The help desk application documented all incoming calls from sample members. In addition to this primary documentation function, it provided the following:

- information required to verify a sample member's identity;
- login information allowing a sample member to access the web interview;
- systematic documentation of each call; and
- a means for tracking calls that could not be resolved immediately.

The help desk application also provided project staff with reports on the types and frequency of problems experienced by sample members and a means for monitoring the resolution status of all help desk inquiries.

Telephone interviews. After the early response phase expired for each case, the production-interviewing phase began, during which attempts were made to locate and interview study sample members who had not yet completed an interview. Once the interviewee was located, an attempt was made to conduct the full interview by telephone, although the web version was available throughout the data collection period.

All calls were managed by the CATI-CMS, the system used by the call center to track outbound telephone calls and case status. This system included an automated call scheduler that assigned cases to interviewers by case priority and time of day. Case assignment was designed to maximize the likelihood of contacting and interviewing sample members, according to when the sample member was likely to be available. For each case, a call roster prioritized sample member names and telephone numbers for the interviewers. New roster lines were added as dictated by the results of CATI tracing and intensive tracing results.

Some cases required special treatment. To manage those students who initially refused to participate (including those for whom locator sources acted as "gatekeepers"), some interviewers were trained in refusal-conversion techniques. Sample members and locator sources who spoke only Spanish, primarily located in Puerto Rico, were assigned to bilingual CATI interviewers.

An abbreviated interview was developed in part as a tool for Spanish translation and as a tool for nonresponse conversion near the end of the data collection period. Sample members who had been identified as Spanish-speaking were given the opportunity to take the NPSAS:08 student interview by telephone, in Spanish, through the use of the Spanish abbreviated interview. Sample members who had been identified as hard to reach or as refusals were offered the abbreviated interview, which took approximately 10 minutes to complete, as opposed to the approximately 25 minutes for the full interview. Once the sample member was offered the abbreviated interview, it could be completed either by telephone or on the Web. Sample members were offered the same incentive for completing the abbreviated interview.

Use of incentives. Sample members were offered an incentive of $30 to complete a web interview during the early response phase. Sample members were eligible to receive the incentive, regardless of participation mode, provided they completed the interview before their incentive expiration date. No incentive was offered during the production-interviewing phase.

Toward the end of data collection, a different incentive plan was used with particular types of nonrespondents: (1) sample members who initially refused the interview; (2) sample members for whom intensive tracing yielded a good mailing address but no telephone number; and (3) sample members identified as hard to reach (i.e., those with 15 or more call attempts, where contact had been established with the sample member and no "hard" appointment was pending). The incentive offer consisted of a letter from the project director on RTI letterhead or an e-mail tailored to the specific type of nonrespondent. Sample members were promised a $30 check if they completed the interview, regardless of the mode they used to do so. The incentive letters were mailed on a flow basis as respondents met one of the three criteria for designation as a target nonrespondent.

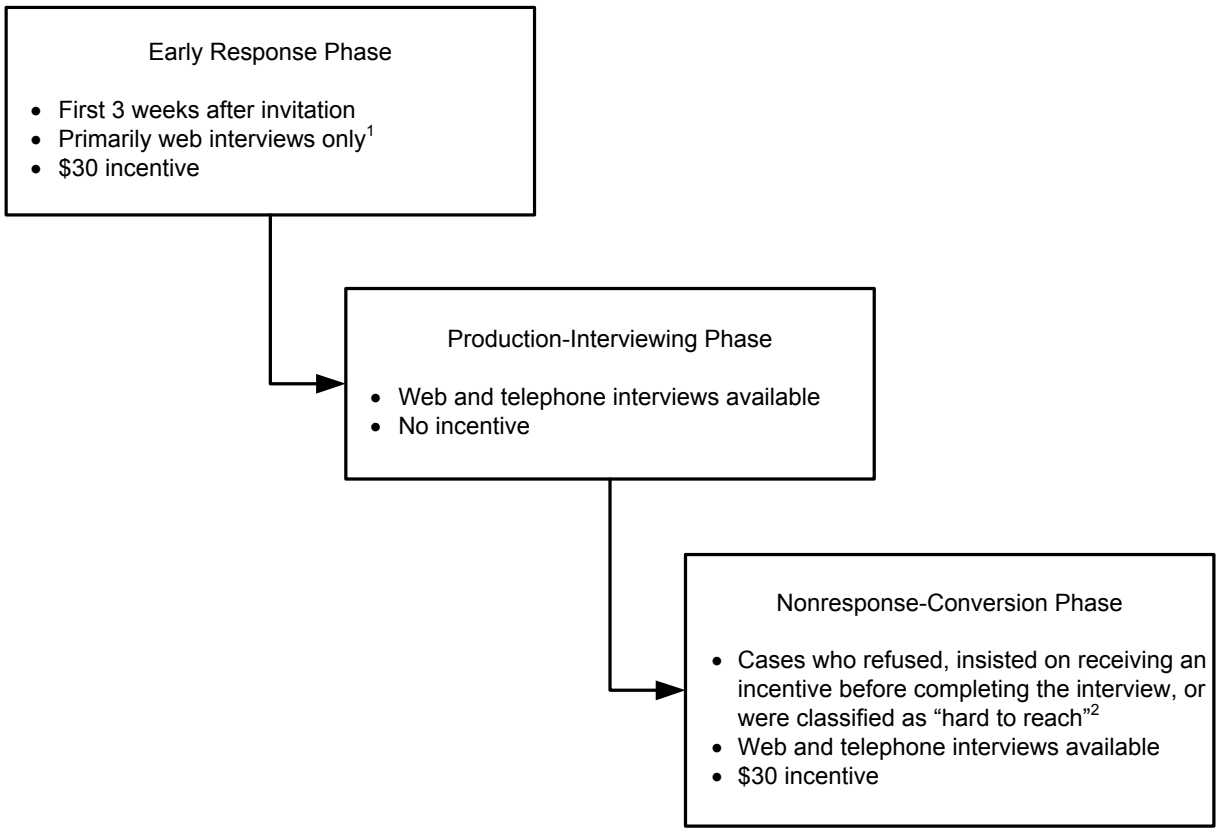
Near the end of data collection, as part of a final effort to convert sample members who still had not responded to the previous incentive offers, all pending cases received a final letter requesting participation by the end of the data collection period. Postcards and letters reemphasized the importance of the study and offered all remaining nonrespondents a check for $30 if they completed the interview.

2.4.6 Phases of Student Interviewing

The student interview design for the NPSAS:08 consisted of three phases (figure 8). The first phase, the *early response phase*, allowed sample members to complete the student interview over the Web. This phase lasted approximately 3 weeks from the time that the student was sent the data collection announcement letter. Because students were sampled for NPSAS:08 on a flow basis, the dates for the early response phase varied by institution. Sample members who completed the interview during this phase received an incentive of $30. No outbound calls were made during this phase; however, sample members were able to call the help desk to request a telephone interview.

The second phase of data collection was the *production-interviewing phase*. Telephone interviewers made outbound calls to early response phase nonrespondents to complete the interview over the telephone. No incentive was offered to respondents during this phase.

Figure 8. Data collection phases for NPSAS:08: 2008



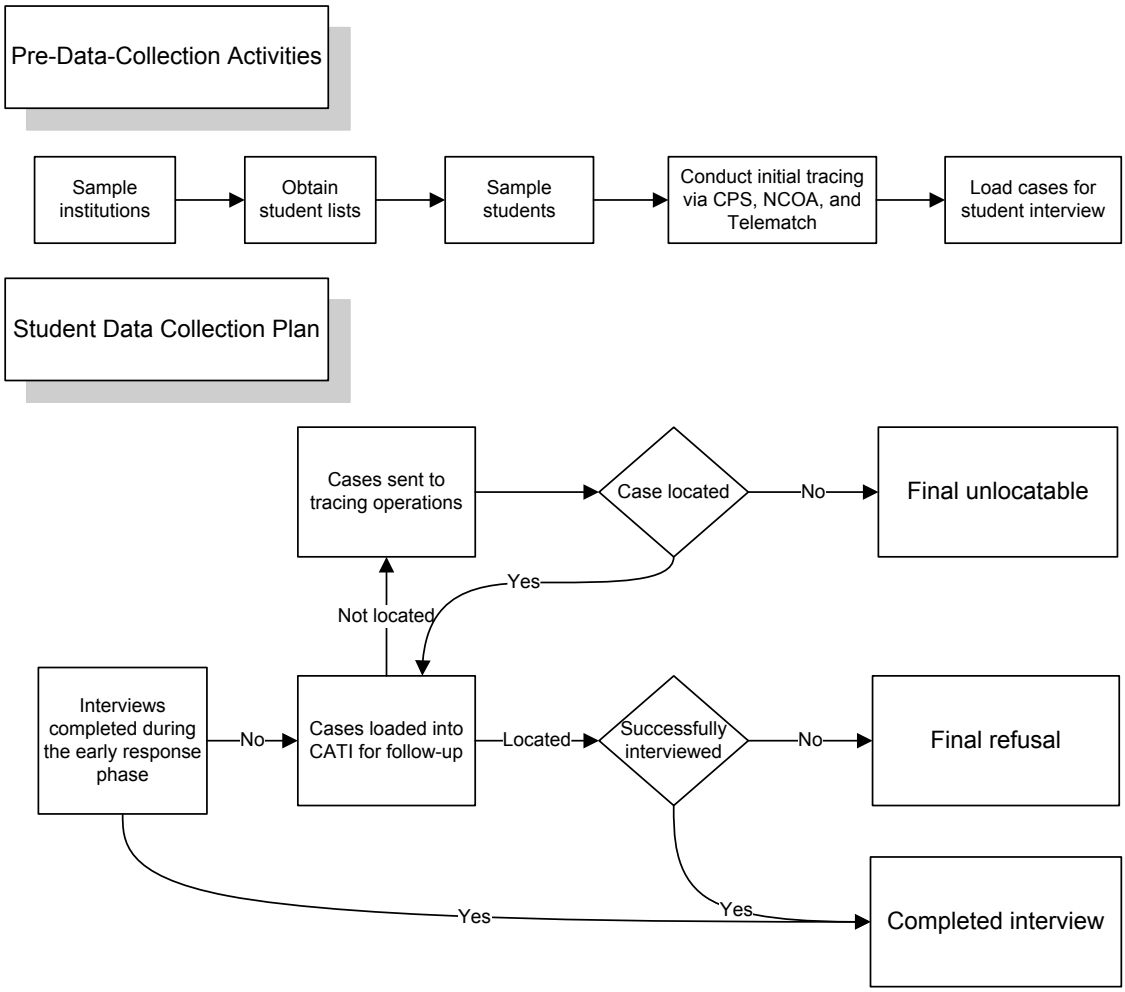
[1] Sample members could also call the help desk to complete a telephone interview during the early response phase.
[2] A sample member became "hard to reach" once he or she was called at least 15 times with minimal or no contact or was not locatable in the Tracing Operations Unit (TOPS).
NOTE: The web interview was available throughout the course of data collection.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

The final phase of data collection was the *nonresponse-conversion phase*. Once a sample member was classified as a refusal or as a "hard-to-reach" case (i.e., was called at least 15 times with minimal or no contact or was not locatable in RTI's Tracing Operations Unit [TOPS]) or insisted on receiving the incentive before completing the interview, that member became eligible for a nonresponse-conversion incentive. Sample members who completed the interview during this phase were offered an incentive of $30. Regardless of their interview phase, sample members could access the web interview throughout the entire data collection period.

2.4.7 Locating and Contacting Students

The data collection design for the NPSAS:08 student interviews involved several steps (figure 9). After sampling, initial tracing attempts were made by several locating databases. Sample members who were difficult to locate were sent to intensive tracing. Once located, sample members were contacted to complete the student interview.

Figure 9. NPSAS:08 data collection plan: 2008



NOTE: CATI = computer-assisted telephone interviewing; CPS = Central Processing System; NCOA = National Change of Address.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

Batch searches. After receiving enrollment lists from participating institutions, batch locating activities were conducted so that address and telephone information for the selected sample members could be updated. Initially, information received from institutions was entered into the NPSAS:08 locator database. This database served as a central repository for all locating information obtained for sample members. Several databases were used to ensure that the student locating information provided by the institutions was kept current.

First, cases with a valid Social Security number were sent to CPS for matching. The information obtained from CPS was compared with the information already obtained from the institutions, and any changes were loaded into the locator database. Next, all cases with one or more valid addresses were sent to the U.S. Postal Service National Change of Address (NCOA) database. The NCOA database consists of change-of-address data submitted to the U.S. Postal Service. It is updated every 2 weeks, and records are stored for 4 years. This new address information provided another update for the locator database. Finally, all student addresses and telephone numbers were sent to Telematch for any telephone number changes. Telematch offers a computerized residential

telephone number updating service consisting of more than 150 million directory assistance consumer and business listings. The service uses name, street address, and zip code as search criteria and returns a telephone number for each match.

Mailing. After addresses were updated, a notification mailing went to all sample members, inviting them to complete the web interview. The initial student mailing contained a lead letter and brochure sent by U.S. Postal Service Priority Mail. The materials contained information about the study, responses to commonly asked questions, information about confidentiality and security, contact information for project and NCES staff and the NSPAS:08 help desk, and details on how to access the web instrument with sample members' unique usernames and passwords (for a copy of all notification materials sent to sample members, see appendix E). Because the student sample was so large, letters were sent approximately twice a week on a flow basis, depending on when student information was received from the institution and when batch tracing procedures for the case were completed.

Locating during interviewing. Telephone contacting of cases that did not complete an interview during the early response period began immediately after the early completion expiration date, although selected sample members may have received a prompting call halfway through their early collection period. When assigned a web nonrespondent case, the interviewer called the telephone number that the computer-assisted telephone interviewing case management system (CATI-CMS) designated as having the greatest likelihood of reaching the sample member and attempted an interview. If the sample member could not be reached at this number, the interviewer asked the person who answered the call how to contact the sample member. If this approach did not lead to contact with the sample member, the interviewer initiated tracing procedures (using all other available information for the sample member and other contact persons to attempt location of the sample member). When all tracing options available to the interviewer were exhausted, the case was assigned to RTI's TOPS for intensive tracing.

Intensive tracing. The most difficult cases to locate were sent to TOPS for intensive tracing. These cases included those that had no telephone number to be loaded into the CATI-CMS and those that were designated a "dead end" in CATI (i.e., no more telephone numbers were available for the case). TOPS accessed both public-domain and proprietary locating databases. The proprietary databases provided real-time access to credit bureau databases (TransUnion and Experian), which contain current address and telephone listings for the majority of consumers with a credit history. TOPS also accessed a variety of other sources that provided information on current addresses, telephone numbers, and the names and telephone numbers of any contacts who might be able to assist in finding the sample member. Whenever these sources provided information that indicated a sample member was not available for the study (e.g., deceased, incarcerated, or out of the country), the case was finalized, and no further contact efforts were made.

TOPS used a two-tiered intensive tracing strategy to locate NPSAS:08 sample members. The first tier (TOPS-1) identified sample members with Social Security numbers to trace them through consumer databases. If a search generated a new telephone number, that case was sent back to CATI for telephone interviewing. If a new address but not a new telephone number was generated, tracers called directory assistance or accessed other databases to obtain telephone numbers. This first level of effort minimized the time that cases were in tracing and unavailable for CATI efforts.

All remaining cases (those lacking new information from the Social Security number search) underwent a more intensive level of tracing in the second tier (TOPS-2). The second tier of tracing activities included the following:

- accessing directory assistance for telephone number listings at various addresses;
- using electronic reverse match databases to obtain the names and telephone numbers of contacts and then calling those contacts;
- contacting the current or last known residential sources, such as landlords, current residents, tax assessors, realtors, and other business establishments, including institutional or alumni offices associated with previous addresses; and
- conducting web searches to attempt to locate sample members and contacts.

Tracers checked new leads produced by these steps to confirm the addresses and telephone numbers for the sample members. When information for a case was confirmed, the case was returned to CATI for telephone interviewing. If TOPS located a new e-mail address for a sample member, the information was loaded into the database for future e-mail correspondence to nonrespondents. Cases that could not be located (e.g., no working telephone numbers were available or numbers for relevant neighborhood sources were unpublished) were reviewed by supervisors and, as necessary, were finalized as unlocatable.

Nonresponse conversion. Once sample members were classified as refusals or as hard to reach (i.e., they had refused at least once, were called at least 15 times with minimal or no contact), they moved into the third phase of data collection: nonresponse conversion. Once a case entered the nonresponse-conversion phase, the sample member was sent an additional letter informing him or her of eligibility for a $30 incentive. Final requests for participation were sent to all remaining nonrespondents approximately 10 days before the data collection period ended. These final requests were sent by Federal Express, except when the destinations were post office boxes, campus boxes, or Puerto Rico addresses, in which case they were sent by U.S. Postal Service Priority Mail.

2.5 Systems

2.5.1 Hatteras and the Instrument Development and Documentation System

The NPSAS:08 instruments were designed with the use of Hatteras, an RTI-developed web-based survey system. Hatteras has two major components: (1) a web-based survey editor used to design the instrument, and (2) the survey engine to deploy the instrument to the web browser. Final documentation of the survey instrument was completed with use of the Instrument Design and Documentation System (IDADS) documentation module. IDADS was developed by RTI to facilitate the production of final data files.

The Hatteras survey editor allowed project staff to develop, review, and modify changes to the NPSAS:08 instrument specifications and code. Some of the many features of the survey editor include searching, reviewing, commenting on, updating, exporting, and importing information associated with instrument development. All comments associated with each form, as well as a historical account of all changes requested by both project staff and NCES, were automatically saved in the database.

The Hatteras survey engine allowed changes to be made to the NPSAS:08 instrument easily. The engine created ASP.NET, version 2.0, web pages from the specifications created in the survey editor. When changes were made, these web pages required only simple table updates rather than extensive programmer intervention. Programmers were required only for changes that involved more than question wording changes.

Once the survey was in production, the instrument specifications were imported into the IDADS documentation module. All data from Hatteras, including question wording, variable, and value labels for each item, were imported into IDADS. Additional documentation information was entered with the use of IDADS, including technical descriptions of items, such as variable types (alpha or numeric), the respondent to whom the item was administered, variable prefix, data source for the variable, and frequency distributions for response categories. The documentation module was used to generate instrument facsimiles and deliverable electronic codebook (ECB) input files.

2.5.2 Integrated Management System

All elements of the data collection process were monitored through an Integrated Management System (IMS). The IMS is a comprehensive set of desktop tools designed to give project staff and NCES easy access to a centralized repository for project data and documents. The NPSAS:08 IMS comprises several modules: the management module, the Receipt Control System (RCS), and the web-based CATI-CMS.

The *management module* of the IMS is a secure website that contains tools to assist project staff and the NCES project officer in managing the study. Important study management information, such as the current project schedule, monthly progress reports, daily data collection reports and status reports, project plans and specifications, key project information and deliverables, instrument specifications, staff contacts, the project bibliography, and a document archive are located in the management module of the IMS. The IMS website also has a download area from which the client and subcontractors can retrieve files when necessary.

The *RCS* refers not only to the collection of case and locator data, but also to the group of integrated systems that throughout data collection monitored the data and performed regularly scheduled processes, such as tracing, locating, and mailing. With the RCS integrated systems, project staff were able to perform stage-specific activities, track status of cases, identify problems early, and implement solutions effectively. RCS locator data were used for a number of daily tasks related to sample maintenance. All RCS data were stored in SQL tables and programmed in ASP.NET, version 2.0. The most widely used RCS systems are as follows:

- The Mailout program is a web-based application used interactively to produce mail-merge files based on sample member status.
- The Email-out program is a web-based application that runs either interactively or in batch mode to produce e-mail messages to specific groups of individuals and according to specified criteria.
- The CSQuery program is a web-based application that allows administrators to see all case information stored in the RCS tables, such as current status, when mailouts or e-mail messages were sent, contacts, and addresses.
- The Tracing program is a web-based interactive application that allows users to produce flat files to send to external vendors of locating services, such as Telematch or NCOA. The user can select a group of cases to send on the basis of their status, and the file is formatted according to the specific tracing company's specifications. The application provides links to the various tracing company websites for secure data upload and download.

- The Mail-Return program allows users to document receipt of letters returned by the U.S. Postal Service because of incorrect or forwarded addresses. Users of this program select the case, choose the specific mailout returned, and either enter the new address or indicate the old address as bad. The case is then ready for the next mailing if a valid address exists for it.

The RCS interacts with the CMS and TOPS databases, sending locator data between the three systems as necessary, generally as overnight processes. For example, RCS overnight processes checked for event codes indicating that a new student sample had been selected from a school's enrollment list and then automatically loaded the case data into the RCS tables. On completion of loading, the status for each case newly sampled would be flagged as requiring transfer to the CPS. The next morning, project staff would send the cases ready for the CPS; when the cases returned from the CPS, their status would be set so that, the next night, the RCS overnight processes would move the cases to the next phase of processing, batch locating. Other examples of overnight processes run by the RCS include the following: reminder e-mail messages and mailings based on time since last contact, IMS report updates, preloading of instruments, and downloading survey data from NCES servers.

The CMS is the technological infrastructure that connected the various components of the CATI system, including the NPSAS:08 instrument, utility screens, databases, call scheduler, report modules, links to outside systems, and other system components. The system used a call scheduler to assign cases to interviewers in a predefined priority order. In addition to delivering appointments to interviewers at the appropriate time, the call scheduler calculated the priority scores (the order in which cases must be called, according to preprogrammed rules), sorted cases in nonappointment queues, and computed time zone adjustments to ensure that cases were not called outside the specified calling hours. The call scheduler also permitted callbacks to be set and assigned status codes to the cases. Each NPSAS:08 case contained one or more roster lines that detailed specific contact information for a sample member (e.g., home and work telephone numbers); the call scheduler used a call algorithm to determine, in accordance with previous call results, which roster line should be called next.

2.5.3 Variable Tracking System

The central mechanism for constructing input files for the NCES ECB was a software application called the *Variable Tracking System* (VTS). The VTS tracked and stored documentation for student records, student interviews, administrative databases, and derived variables required for the ECB and the Data Analysis System. This documentation included weighted and unweighted variable distributions, variable labels, value codes and labels, and a text field describing the development and source of each variable and, if applicable, the programming code used to construct it. The VTS automatically produced, according to NCES specifications, the input files for the ECB and the Data Analysis System.

Chapter 3. Data Collection Outcomes

Presented in this chapter are the data collection results for the 2007–08 National Postsecondary Student Aid Study (NPSAS:08), including institutional response rates; student response rates; student record matching to administrative databases and institutional records; and student locating, tracing, and interviewing outcomes. Results are presented by institutional characteristic and student type, and the overall completeness of data is presented and assessed.

3.1 Institutional Participation

Eligible sample institutions were asked to participate in two stages of NPSAS:08 by (1) providing a comprehensive list of enrolled students for sample selection, and (2) providing data from student records for the sampled students. Consequently, the potential for institutional nonresponse existed at these two points in the survey process. Rates of institutional response for both the national and the state-representative samples were assessed.

3.1.1 National Institution Sample

Of the total institution sample (n = 1,960), about 1,940 institutions (99 percent) were determined to be eligible to participate in NPSAS:08 (table 9). Approximately 1,730 institutions (89 percent) provided student enrollment lists for use in the second stage of sampling (e.g., selecting the student sample). List provision rates (among eligible institutions) varied by type of institution, ranging from 80 percent for private, for-profit, less-than-2-year institutions to 94 percent for public, 4-year, non-doctorate-granting institutions. Weighted participation rates were calculated in accordance with the institutional probabilities of selection and enrollment.[21] The overall weighted participation rate was 90 percent.[22]

[21] In addition to the probabilities of selection, the participation rates accounted for the institution enrollment from the 2004–05 Integrated Postsecondary Education Data System (IPEDS). The unit of analysis for NPSAS is a student, so factoring the enrollment into the participation rates indicates the percentage of students represented by the participating institutions. The weighted response rates are based on the base weight and can be interpreted as the estimated percentages of institutions in the population that would have provided a usable student sampling list, if they had been asked.

[22] When the weighted response rates differ from the unweighted response rates, it is because of the nonresponding institutions' having, on average, a lower or higher weight than the responding institutions. NPSAS:08 was designed to produce efficient estimates only at the student level. Institutions were selected with probabilities proportional to size; therefore, weighted institution-level estimates are subject to a high level of sampling variation.

Table 9. Numbers of sampled, eligible, and participating institutions and enrollment list participation rates, by institutional characteristic: National sample: 2008

Institutional characteristic[1]	Sampled institutions	Eligible institutions	Institutions providing lists		
			Number	Unweighted percent	Weighted percent[2]
All institutions	1,960	1,940	1,730	89.0	90.1
Institutional level					
Less-than-2-year	130	120	100	82.6	83.2
2-year	570	560	510	89.7	90.7
4-year non-doctorate-granting	700	700	630	89.7	91.9
4-year doctorate-granting	560	560	500	88.8	88.6
Institutional control					
Public	960	960	880	91.9	91.2
Private not-for-profit	650	640	560	87.4	86.7
Private for-profit	350	340	290	83.6	88.2
Type of institution					
Public less-than-2-year	20	20	20	90.9	93.2
Public 2-year	450	450	410	91.7	91.2
Public 4-year non-doctorate-granting	200	200	190	94.4	95.4
Public 4-year doctorate-granting	290	290	260	90.7	89.2
Private not-for-profit less-than-4-year	20	20	20	84.2	84.7
Private not-for-profit 4-year non-doctorate-granting	370	370	320	88.2	87.9
Private not-for-profit 4-year doctorate-granting	260	260	230	86.5	85.9
Private for-profit less-than-2-year	100	90	70	80.4	81.0
Private for-profit 2-year-or-more	260	250	210	84.8	90.2

[1] Institutional characteristics are based on data from the sampling frame, which was formed from the 2004–05 IPEDS and freshened from the 2005–06 IPEDS.
[2] The weight described in this column is a base weight.
NOTE: Percentages are based on the unrounded count of eligible institutions within the row under consideration. Detail may not sum to totals because of rounding. IPEDS = Integrated Postsecondary Education Data System.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

3.1.2 State Institution Samples

For the state-representative samples, list provision rates (among eligible institutions) varied by state and type of institution, ranging from 67 percent for Illinois private, for-profit, degree-granting institutions to 100 percent for Georgia and Minnesota public, 2-year and 4-year institutions (table 10). The weighted participation rates ranged from 74 percent for Texas private, not-for-profit, 4-year institutions to 100 percent for Georgia and Minnesota public, 2-year and 4-year institutions.

Table 10. Numbers of sampled, eligible, and participating institutions and enrollment list participation rates, by institutional characteristic: State samples: 2008

Institutional characteristic[1]	Sampled institutions	Eligible institutions	Institutions providing lists		
			Number	Unweighted percent	Weighted percent[2]
California					
Public 2-year	70	70	60	82.4	83.6
Public 4-year	30	30	30	87.5	88.8
Private not-for-profit 4-year	40	30	30	79.4	75.2
Private for-profit degree-granting	30	30	30	83.3	87.4
Georgia					
Public 2-year	50	50	50	100.0	100.0
Public 4-year	20	20	20	100.0	100.0
Private not-for-profit 4-year	30	30	30	89.3	91.7
Private for-profit degree-granting	20	20	20	82.6	79.9
Illinois					
Public 2-year	30	30	30	90.0	91.9
Public 4-year	10	10	10	91.7	93.2
Private not-for-profit 4-year	30	30	30	93.3	92.4
Private for-profit degree-granting	30	30	20	66.7	90.0
Minnesota					
Public 2-year	30	30	30	100.0	100.0
Public 4-year	10	10	10	100.0	100.0
Private not-for-profit 4-year	30	30	20	88.9	80.7
Private for-profit degree-granting	30	30	30	96.8	97.4
New York					
Public 2-year	40	40	30	89.5	96.8
Public 4-year	40	40	40	92.1	93.4
Private not-for-profit 4-year	60	60	50	91.1	94.7
Private for-profit degree-granting	30	30	30	92.9	95.2
Texas					
Public 2-year	30	30	30	93.3	94.5
Public 4-year	40	40	40	92.1	89.1
Private not-for-profit 4-year	30	30	30	83.3	74.1
Private for-profit degree-granting	30	30	20	85.7	89.6

[1] Institutional characteristics are based on data from the sampling frame, which was formed from the 2004–05 IPEDS and freshened from the 2005–06 IPEDS.
[2] The weight described in this column is a base weight.
NOTE: Detail may not sum to totals because of rounding. Percentages are based on the eligible institutions within the row under consideration. IPEDS = Integrated Postsecondary Education Data System.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

3.2 Study Respondents

Student-level data for NPSAS:08 were collected from various sources: student records by means of computer-assisted data entry (CADE); student interviews; and administrative federal and private databases such as the Central Processing System (CPS), the National Student Loan Data System (NSLDS), National Student Clearinghouse (NSC), ACT files, and SAT Reasoning Test

(SAT) files. For NPSAS:08, key variables were identified across these data sources to determine the minimum requirements to support the analytic objectives of the study. Sample members who met these minimum requirements were classified as *study respondents*. Specifically, a study respondent was any sample member who was determined to be study eligible, according to the criteria delineated in chapter 2, and, at a minimum, had valid data from any source for the following:

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

3.2.1 National Student Sample

The final national sample numbered 137,800 students (table 11). Approximately 96 percent of the final sample (n = 132,800) was eligible for NPSAS. On completion of data collection, 96 percent of the eligible sample (n = 127,700) was determined to have sufficient data to meet the definition of study respondent. The unweighted student response rates (among eligible students) varied by type of institution, ranging from 90 percent for students from public, less-than-2-year institutions to 99 percent for students from private, for-profit, 2-year-or-more institutions. Weighted response rates were calculated on the basis of the institutional weights and student probabilities of selection.[23] The overall student weighted response rate was 96 percent.[24]

[23] The weighted response rates are based on the base weight and can be interpreted as the estimated percentages of students in the population that would have responded if they had been asked.

[24] When the weighted response rates differ from the unweighted response rates, it is because of the nonresponding students' having, on average, a lower or higher weight than the responding students.

Table 11. Numbers of sampled and eligible students and response rates, by institutional characteristic: National sample: 2008

Institutional characteristic[2]	Sampled students	Eligible students[3]	Study respondents[1] Unweighted percent	Study respondents[1] Weighted percent[4]
All students	137,800	132,800	96.2	95.7
Institutional level				
Less-than-2-year	8,820	7,950	95.0	96.7
2-year	43,460	40,770	93.3	92.5
4-year non-doctorate-granting	37,930	37,140	97.8	97.6
4-year doctorate-granting	47,590	46,940	97.6	97.6
Institutional control				
Public	87,470	84,240	95.3	94.9
Private not-for-profit	32,760	31,950	90.0	88.9
Private for-profit	17,570	16,610	92.8	92.2
Type of institution				
Public less-than-2-year	1,730	1,480	90.0	88.9
Public 2-year	39,340	37,010	92.8	92.2
Public 4-year non-doctorate-granting	16,120	15,850	98.0	98.1
Public 4-year doctorate-granting	30,280	29,910	97.3	97.4
Private not-for-profit less-than-4-year	2,080	1,790	97.0	97.7
Private not-for-profit 4-year non doctorate-granting	14,200	13,930	97.3	96.8
Private not-for-profit 4-year doctorate-granting	16,480	16,230	98.0	97.8
Private for-profit less-than-2-year	6,610	6,050	96.1	97.6
Private for-profit 2-year-or-more	10,960	10,560	98.5	98.7

[1] A study respondent is defined as any eligible sample member for whom sufficient key data were obtained from one or more sources, including the student interview, institutional records, and the U.S. Department of Education's Central Processing System.
[2] Institutional characteristics are based on data from the sampling frame formed from the 2004–05 IPEDS and freshened from the 2005–06 IPEDS.
[3] Sample member eligibility was determined during the student interview or from institutional records in the absence of a student interview.
[4] The weight described in this column is a base weight.
NOTE: Percentages are based on the unrounded count of eligible students within the row under consideration. Details may not sum to totals because of rounding. IPEDS = Integrated Postsecondary Education Data System.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

3.2.2 State Student Samples

The unweighted and weighted student response rates (among eligible students) varied by state and type of institution, ranging from 83 percent for California public, 2-year institutions to almost 100 percent for Illinois private, for-profit, degree-granting institutions (table 12).

Table 12. Numbers of sampled, eligible undergraduate students and response rates, by institutional characteristic: State samples: 2008

Institutional characteristic[2]	Sampled undergraduate students	Eligible undergraduate students[3]	Undergraduate study respondents[1]		
			Number	Unweighted percent	Weighted percent[4]
California					
Public 2-year	5,380	4,880	4,070	83.5	83.5
Public 4-year	3,870	3,850	3,690	96.0	96.5
Private not-for-profit 4-year	1,980	1,940	1,810	93.3	93.9
Private for-profit degree-granting	1,070	1,030	1,020	99.0	99.5
Georgia					
Public 2-year	4,170	4,020	3,880	96.5	94.2
Public 4-year	2,120	2,090	2,080	99.3	99.4
Private not-for-profit 4-year	1,510	1,490	1,470	98.7	97.8
Private for-profit degree-granting	1,670	1,610	1,600	99.3	99.4
Illinois					
Public 2-year	2,770	2,580	2,190	84.7	91.8
Public 4-year	1,860	1,830	1,780	97.2	97.8
Private not-for-profit 4-year	2,230	2,180	2,150	98.9	98.8
Private for-profit degree-granting	1,460	1,410	1,410	99.6	99.7
Minnesota					
Public 2-year	4,640	4,400	4,190	95.1	95.8
Public 4-year	2,290	2,270	2,200	96.8	96.4
Private not-for-profit 4-year	1,290	1,290	1,280	99.3	99.1
Private for-profit degree-granting	1,340	1,300	1,260	97.1	95.8
New York					
Public 2-year	4,200	4,070	3,960	97.4	97.2
Public 4-year	3,160	3,130	3,090	98.6	98.4
Private not-for-profit 4-year	3,450	3,390	3,330	98.1	98.1
Private for-profit degree-granting	1,550	1,440	1,410	97.6	96.5
Texas					
Public 2-year	2,430	2,270	2,100	92.5	93.3
Public 4-year	3,160	3,120	3,010	96.6	96.2
Private not-for-profit 4-year	1,280	1,250	1,240	98.6	98.9
Private for-profit degree-granting	1,330	1,290	1,270	98.1	99.1

[1] A study respondent is defined as any eligible sample member for whom sufficient key data were obtained from one or more sources, including student interview, institutional records, and the U.S. Department of Education's Central Processing System.
[2] Institutional characteristics are based on data from the sampling frame formed from the 2004–05 IPEDS and freshened from the 2005–06 IPEDS.
[3] Sample member eligibility was determined during the student interview or from institutional records in the absence of a student interview.
[4] The weight described in this column is a base weight.
NOTE: Detail may not sum to totals because of rounding. Percentages are based on the eligible students within the row under consideration. IPEDS = Integrated Postsecondary Education Data System.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

3.3 Data Collection Results

3.3.1 Student Record Matching

Central Processing System. Reported here are the results of matching and downloading student data from the U.S. Department of Education's CPS (table 13). The CPS contains data provided to the U.S. Department of Education by students and their families when they complete the Free Application for Federal Student Aid (FAFSA); therefore, successful matching to CPS can occur only for sample members who were federal student financial aid applicants for the years requested.

The initial CPS matching process began after the student sample had been selected for an institution but before student record abstraction and student interviewing had begun. Matching was against the CPS data for the 2007–08 financial aid year. Because data obtained from the CPS were used to determine study response status, match rates are presented for all eligible sample members for whom a Social Security number (SSN) was available. Not all sample members were submitted to the CPS for matching (see table 13), primarily because some institutions did not provide student SSNs and last names. After CADE data collection, some student cases not initially matched to the CPS were resubmitted, either because of a newly obtained SSN or evidence in the institutional records that the student had applied for federal student aid for the 2007–08 academic year.

The overall matching rate for the 2007–08 CPS data was about 69 percent. Match rates varied by type of institution, ranging from 56 percent for public, less-than-2-year institutions to 90 percent for private, for-profit, less-than-2-year institutions. Approximately 48 percent of graduate/first-professional students matched to the 2007–08 CPS, while 71 percent of total undergraduate students, potential baccalaureate students and other undergraduate students matched. As part of the undergraduate aid process, nearly all institutions required undergraduate aid applicants to file a FAFSA to determine their eligibility for federal Pell Grants, federal campus-based aid, and federal loans. Graduate and first-professional students were not ordinarily required to file a FAFSA unless they were applying specifically for federal loans, the only type of federal aid generally available to graduate students; graduate students often apply directly to their institution or department for fellowships and assistantships, which are usually not based on financial need and do not require the completion of the federal financial aid forms on which CPS matching is based.

The NPSAS:08 sampled students were matched to the 2008–09 CPS files, as well. It was expected that fewer students would match to the 2008–09 CPS files, primarily because some students completed their postsecondary education during the 2007–08 NPSAS year. Overall, 71 percent of eligible students matched to either the 2007–08 CPS or the 2008–09 CPS, and 37 percent matched to both data files (see table 13).

National Student Loan Data System. Because data obtained from NSLDS were not required to classify a sample member as a study respondent, results of the matching to NSLDS loan and grant files are based only on study respondents themselves (table 14). Successful matching to NSLDS could occur only for sample members who have received federal loans, Pell Grants, Science and Mathematics Access to Retain Talent (SMART) Grants, or Academic Competitiveness Grants (ACGs). NSLDS files are historical, so information about receipt of such loans and grants was available not only for the NPSAS study year, but also for any applicable prior years (where applicable). Consequently, historical match rates reported for eligible study respondents do not necessarily reflect the current NPSAS year.

In total, 79,790 study respondents (63 percent of eligible study respondents) were matched to the historical loan database. NSLDS loan match rates ranged from 43 percent for students who attended public, 2-year institutions to 90 percent for private, for-profit, 2-year-or-more institutions. Pell Grant matches were obtained for 62,890 study respondents (49 percent). The Pell Grant match rate ranged from 35 percent for private, not-for-profit, 4-year, doctorate-granting institutions to 81 percent for private, for-profit, less-than-2-year institutions. ACG or SMART Grant matches were acquired for 11,780 study respondents (9 percent of those submitted). The match rate ranged from less than 1 percent for public, less-than-2-year institutions to 15 percent for public, 4-year, doctorate-granting institutions.

National Student Clearinghouse. Results are reported for the matching to the National Student Clearinghouse (NSC), which provides information on postsecondary enrollment, degree, and certificate records on behalf of participating postsecondary institutions (table 15). Match results are based on enrollment and degree records for the 2007–08 academic year. An individual student record was able to match to the NSC only if the student's institution was a participant in the NSC. Conversely, the study respondents matched to the NSC for both their NPSAS sampled institution and any other participating institutions attended during the 2007–08 year.

In total, 93,890 study respondents (74 percent) matched to the NSC for their NPSAS sampled institution. The match rate ranged from zero percent for both public less-than-2-year and private, for-profit less-than-2-year institutions to 87 percent for public, 4-year, non-doctorate-granting institutions. Matches to institutions other than the sample members' NPSAS institutions were obtained for 16,060 study respondents (13 percent). The match rate ranged from 6 percent at private, not-for-profit, less-than-4-year institutions and private, for-profit, less-than-2-year institutions to 18 percent at private, for-profit, 2-year-or-more institutions. A small percentage (8 percent) of study respondents matched to the NSC for both their NPSAS institution and at least one other institution.

ACT and SAT Reasoning Test. ACT survey data and scores came from the most recent test record for each matched sample member between the 2001–02 and 2006–07 academic years. In total, 29,590 study respondents (23 percent) matched to the ACT database (table 16). The match rate ranged from 8 percent for students sampled from private, for-profit, less-than-2-year institutions to 33 percent for students sampled from public, 4-year, doctorate-granting institutions. Match rates also varied by student type: about 25 percent of undergraduate students had an ACT record on file for the matched years, whereas only 6 percent of the graduate students had a similar record in the database.

The most recent student records of SAT and questionnaire data were obtained for high school graduation years 2002–07. SAT data records were acquired for 48,600 study respondents (38 percent; see table 16). Rates of matched records ranged from 11 percent of students from public, less-than-2-year institutions to 55 percent of students from private, not-for-profit, 4-year, doctorate-granting institutions.

The rates for study respondents matching to both SAT and ACT or to either SAT or ACT are also reported. In total, 11,670 study respondents (9 percent) matched to both test databases, and 66,530 (52 percent) matched to either one or the other.

Table 13. Results of Central Processing System matching for 2007–08 and 2008–09, by institutional characteristic and student type: 2008

Institutional characteristic and student type[1]	Eligible students[2]	Matched to 2007–08		Matched to 2008–09		Matched to both years		Matched to either year	
		Number[3]	Percent	Number[3]	Percent	Number[3]	Percent	Number[3]	Percent
All students	118,850	81,560	68.6	47,290	39.8	44,200	37.2	84,650	71.2
Institutional level									
Less-than-2-year	7,260	6,050	83.3	1,970	27.1	1,890	26.0	6,140	84.5
2-year	35,700	22,630	63.4	13,320	37.3	12,060	33.8	23,900	66.9
4-year non-doctorate-granting	34,190	25,400	74.3	15,010	43.9	14,350	42.0	26,060	76.2
4-year doctorate-granting	41,700	27,470	65.9	16,980	40.7	15,900	38.1	28,550	68.5
Institutional control									
Public	73,320	46,910	64.0	29,190	39.8	26,810	36.6	49,280	67.2
Private not-for-profit	29,730	20,740	69.8	12,160	40.9	11,590	39.0	21,310	71.7
Private for-profit	15,800	13,910	88.0	5,940	37.6	5,790	36.6	14,060	89.0
Type of institution									
Public less-than-2-year	1,290	720	55.7	230	17.7	200	15.5	740	57.9
Public 2-year	31,980	19,190	60.0	11,990	37.5	10,750	33.6	20,430	63.9
Public 4-year non-doctorate-granting	14,000	9,960	71.2	6,320	45.2	5,970	42.6	10,310	73.7
Public 4-year doctorate-granting	26,070	17,040	65.4	10,650	40.9	9,900	38.0	17,790	68.2
Private not-for-profit less-than-4-year	1,720	1,340	77.6	550	31.6	510	29.7	1,370	79.6
Private not-for-profit 4-year non-doctorate-granting	13,310	9,630	72.4	5,660	42.5	5,430	40.8	9,860	74.1
Private not-for-profit 4-year doctorate-granting	14,700	9,770	66.5	5,960	40.5	5,650	38.5	10,070	68.5
Private for-profit less-than-2-year	5,590	5,010	89.6	1,640	29.3	1,580	28.3	5,070	90.6
Private for-profit 2-year-or-more	10,210	8,900	87.2	4,300	42.1	4,210	41.2	9,000	88.1
Student type									
Total undergraduate	106,080	75,400	71.1	43,690	41.2	40,840	38.5	78,240	73.8
Potential B&B student	22,350	15,760	70.5	4,700	21.0	4,210	18.8	16,250	72.7
Other undergraduate	83,730	59,640	71.2	38,990	46.6	36,640	43.8	61,990	74.0
Graduate/first-professional	12,770	6,160	48.2	3,600	28.2	3,350	26.3	6,410	50.2

[1] Both institutional and student classifications were verified to correct classification errors on the sampling frame.
[2] Includes all eligible students for whom apparently legitimate Social Security numbers were obtained either before or during computer-assisted data entry.
[3] The number presented reflects the total number of matches of those submitted and may include students who were classified as study nonrespondents.
NOTE: Detail may not sum to totals because of rounding. B&B = Baccalaureate and Beyond.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

Table 14. Results of National Student Loan Data System matching, by institutional characteristic and student type: 2008

Institutional characteristic and student type[2]	Study respondents[3]	Matched to NSLDS loan[1]		Matched to NSLDS Pell[1]		Matched to NSLDS ACG or SMART[1]	
		Number	Percent[4]	Number	Percent[4]	Number	Percent[4]
All students	127,700	79,790	62.5	62,890	49.2	11,780	9.2
Institutional level							
Less-than-2-year	7,560	5,390	71.4	5,730	75.8	50	0.7
2-year	38,030	18,330	48.2	19,970	52.5	1,310	3.4
4-year non-doctorate-granting	36,320	25,960	71.5	18,840	51.9	4,260	11.7
4-year doctorate-granting	45,790	30,110	65.7	18,350	40.1	6,150	13.4
Institutional control							
Public	80,290	43,970	54.8	37,710	47.0	7,600	9.5
Private not-for-profit	31,200	21,830	70.0	13,080	41.9	3,560	11.4
Private for-profit	16,220	13,980	86.2	12,100	74.6	620	3.8
Type of institution							
Public less-than-2-year	1,330	650	49.0	730	54.4	#	0.1
Public 2-year	34,330	14,890	43.4	16,920	49.3	1,260	3.7
Public 4-year non-doctorate-granting	15,520	9,880	63.7	7,720	49.8	1,900	12.2
Public 4-year doctorate-granting	29,100	18,550	63.7	12,340	42.4	4,430	15.2
Private not-for-profit less-than-4-year	1,730	1,010	58.3	1,170	67.6	30	1.5
Private not-for-profit 4-year non-doctorate-granting	13,560	9,920	73.2	6,370	47.0	1,850	13.7
Private not-for-profit 4-year doctorate-granting	15,910	10,900	68.5	5,540	34.8	1,680	10.6
Private for-profit less-than-2-year	5,820	4,600	79.0	4,700	80.8	40	0.8
Private for-profit 2-year-or-more	10,400	9,390	90.2	7,400	71.2	580	5.5
Student type							
Total undergraduate	113,190	70,220	62.0	59,150	52.3	11,680	10.3
Potential B&B student	23,630	17,070	72.2	12,210	51.7	4,160	17.6
Other undergraduate	89,560	53,150	59.3	46,940	52.4	7,520	8.4
Graduate/first-professional	14,510	9,570	66.0	3,740	25.8	100	0.7

Rounds to zero.
[1] Both institutional and student classifications were verified to correct classification errors on the sampling frame.
[2] Matching was completed on historical files.
[3] Study respondents were eligible sample members who met minimum data requirements across various data sources.
[4] Percentages are based on the number of eligible students within the row under consideration.
NOTE: Detail may not sum to totals because of rounding. ACG = Academic Competitiveness Grant; B&B = Baccalaureate and Beyond; NSLDS = National Student Loan Data System; SMART = Science and Mathematics Access to Retain Talent.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

Table 15. Results of National Student Clearinghouse matching, by institutional characteristic and student type: 2008

Institutional characteristic and student type[1]	Study respondents[2]	Matched for the NPSAS school		Matched for another school		Matched to both		Matched to either	
		Number	Percent[3]	Number	Percent[3]	Number	Percent[3]	Number	Percent[3]
All students	127,700	93,890	73.5	16,060	12.6	9,810	7.7	100,150	78.4
Institutional level									
Less-than-2-year	7,560	#	#	450	6.0	#	#	450	6.0
2-year	38,030	28,790	75.7	5,600	14.7	3,520	9.3	30,870	81.2
4-year non-doctorate-granting	36,320	26,440	72.8	5,210	14.3	2,690	7.4	28,960	79.7
4-year doctorate-granting	45,790	38,670	84.4	4,800	10.5	3,600	7.9	39,870	87.1
Institutional control									
Public	80,290	67,120	83.6	10,640	13.3	7,360	9.2	70,400	87.7
Private not-for-profit	31,200	23,930	76.7	3,250	10.4	2,230	7.1	24,960	80.0
Private for-profit	16,220	2,840	17.5	2,170	13.4	220	1.4	4,790	29.5
Type of institution									
Public less-than-2-year	1,330	#	#	90	7.0	#	#	90	7.0
Public 2-year	34,330	28,690	83.6	5,420	15.8	3,550	10.3	30,560	89.0
Public 4-year non-doctorate-granting	15,520	13,450	86.6	2,090	13.4	1,360	8.8	14,170	91.3
Public 4-year doctorate-granting	29,100	24,980	85.8	3,050	10.5	2,450	8.4	25,570	87.9
Private not-for-profit less-than-4-year	1,730	410	23.4	100	5.5	20	1.4	480	27.6
Private not-for-profit 4-year non-doctorate-granting	13,560	10,250	75.6	1,780	13.1	1,110	8.2	10,910	80.5
Private not-for-profit 4-year doctorate-granting	15,910	13,280	83.5	1,380	8.7	1,090	6.9	13,570	85.3
Private for-profit less-than-2-year	5,820	#	#	350	6.0	#	#	350	6.0
Private for-profit 2-year-or-more	10,400	2,840	27.3	1,820	17.5	220	2.1	4,440	42.7
Student type									
Total undergraduate	113,190	82,810	73.2	14,930	13.2	9,160	8.1	88,580	78.3
Potential B&B student	23,630	19,490	82.5	2,540	10.8	1,680	7.1	20,350	86.1
Other undergraduate	89,560	63,320	70.7	12,390	13.8	7,480	8.4	68,230	76.2
Graduate/first-professional	14,510	11,080	76.4	1,130	7.8	650	4.5	11,560	79.7

Rounds to zero.
[1] Both institutional and student classifications were verified to correct classification errors on the sampling frame.
[2] Eligible sample members who met minimum data requirements across various data sources were classified as study respondents.
[3] Percentages are based on the number of eligible students within the row under consideration.
NOTE: Detail may not sum to totals because of rounding. B&B = Baccalaureate and Beyond.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

Table 16. Results of ACT and SAT Reasoning Test matching, by institutional characteristic and student type: 2008

Institutional characteristic and student type[1]	Study respondents	Matched to ACT		Matched to SAT		Matched to both		Matched to either	
		Number	Percent[2]	Number	Percent[2]	Number	Percent[2]	Number	Percent[2]
All students	127,700	29,590	23.2	48,600	38.1	11,670	9.1	66,530	52.1
Institutional level									
Less-than-2-year	7,560	570	7.5	880	11.6	80	1.0	1,370	18.1
2-year	38,030	5,920	15.6	9,250	24.3	1,530	4.0	13,640	35.9
4-year non-doctorate-granting	36,320	9,540	26.3	14,360	39.5	3,360	9.3	20,540	56.5
4-year doctorate-granting	45,790	13,570	29.6	24,110	52.7	6,700	14.6	30,980	67.7
Institutional control									
Public	80,290	19,760	24.6	30,740	38.3	7,090	8.8	43,400	54.1
Private not-for-profit	31,200	8,280	26.5	15,060	48.3	4,230	13.6	19,100	61.2
Private for-profit	16,220	1,560	9.6	2,810	17.3	340	2.1	4,020	24.8
Type of institution									
Public less-than-2-year	1,330	120	8.7	140	10.6	10	1.0	250	18.4
Public 2-year	34,330	5,640	16.4	8,810	25.7	1,480	4.3	12,980	37.8
Public 4-year non-doctorate-granting	15,520	4,360	28.1	6,620	42.6	1,240	8.0	9,740	62.7
Public 4-year doctorate-granting	29,100	9,640	33.1	15,170	52.1	4,360	15.0	20,440	70.2
Private not-for-profit less-than-4-year	1,730	150	8.6	250	14.6	50	3.1	350	20.1
Private not-for-profit 4-year non-doctorate-granting	13,560	4,290	31.7	6,040	44.5	1,880	13.9	8,450	62.3
Private not-for-profit 4-year doctorate-granting	15,910	3,840	24.1	8,770	55.1	2,300	14.4	10,310	64.8
Private for-profit less-than-2-year	5,820	450	7.8	700	12.0	60	1.1	1,090	18.7
Private for-profit 2-year-or-more	10,400	1,110	10.6	2,110	20.3	280	2.7	2,930	28.2
Student type									
Total undergraduate	113,190	28,730	25.4	43,540	38.5	11,220	9.9	61,050	53.9
Potential B&B student	23,630	7,670	32.5	12,600	53.3	3,660	15.5	16,610	70.3
Other undergraduate	89,560	21,060	23.5	30,940	34.5	7,550	8.4	44,440	49.6
Graduate/first-professional	14,510	870	6.0	5,070	34.9	450	3.1	5,480	37.8

[1] Both institutional and student classifications were verified to correct classification errors on the sampling frame.
[2] Percentages are based on the number of eligible students within the row under consideration.
NOTE: Detail may not sum to totals because of rounding. Files were matched to the student's most recent ACT score and survey data record for the 2001–02 through 2006–07 academic years and SAT scores for high school graduation years 2002–07. B&B = Baccalaureate and Beyond; SAT = SAT Reasoning Test.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

3.3.2 Student Record Abstraction

Institutions were given three options for abstracting information about sampled students from institutional records. The first option was for the institution's staff to use the web-based CADE application (self-CADE), the second option was for the institution's staff to create data files and upload them to the NPSAS website (data-CADE), and the third option was to have a trained field data collector visit the institution and abstract the data (field-CADE).

Of the 1,670 institutions that provided student record data, most (63 percent) did so by self-CADE (table 17). Data-CADE was the next most used method, with 36 percent of CADE completions being submitted as electronic data files. Field data collectors performed the record abstraction from the remaining 1 percent of CADE completions.

Table 17. Student record abstraction method: 2008

	Institutions providing CADE		Total students[1]	
CADE abstraction method	Number	Percent[2]	Number	Percent[2]
Total	1,670	100.0	130,410	100.0
Abstraction method				
Self-CADE	1,050	62.9	61,880	47.4
Data-CADE	600	35.9	68,110	52.2
Field-CADE	20	1.2	430	0.3

[1] The total represents the number of students sampled from institutions for which CADE data were provided and includes students who were classified as study nonrespondents.
[2] Percentage of total number of eligible institutions/students.
NOTE: Detail may not sum to totals because of rounding. CADE = computer-assisted data entry.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

The trend for an increase in data-CADE and a decrease in field-CADE, seen in NPSAS:04, continued in NPSAS:08. The rate at which institutions chose data-CADE in NPSAS:08 was significantly higher than the rate in NPSAS:04 at 36 percent, as opposed to 21 percent ($\chi = 8.78$, $p < .05$). For NPSAS:08, the data-CADE option was made available earlier in the collection period, which may have contributed to this increase; moreover, most of the participating institutional systems used data-CADE. Although self-CADE was used more often than data-CADE (63 percent, as opposed to 36 percent), a higher percentage of student data was received via data-CADE than via self-CADE (52 percent, as opposed to 47 percent). This finding indicated that the institutions with larger sample sizes preferred data-CADE as their mode of abstraction.

From NPSAS:04 to NPSAS:08, a corresponding decrease in the use of field-CADE was found: 1 percent of institutions chose to use field CADE in NPSAS:08, as opposed to 13 percent in NPSAS:04 ($\chi = 13.13$, $p < .05$). Field-CADE use likely decreased because an increased number of institutional staff now have internet access and familiarity with web applications. A climate increasingly emphasizing data security may also explain the decrease. Some institutional staff reported feeling hesitant about field data collectors' abstracting student data in person.

A student record was considered to represent CADE record completion if it had nonmissing data for one or more of the following variables:

- receipt of financial aid;
- enrollment (attendance status and enrollment intensity);

- tuition;
- degree program; or
- race.

An institution was classified as having completed CADE if sufficient data were obtained for at least one sample student. CADE completion rates are reported in table 18. Overall, 96 percent (weighted) of the participating institutions (those that provided enrollment lists from which a student sample could be selected) completed CADE.

CADE completion rates ranged from 94 percent (weighted) for institutions choosing field-CADE to 98 percent for those using data-CADE. CADE completion rates varied by type of institution as well, ranging from 92 percent for private, for-profit, less-than-2-year institutions to 100 percent for private, not-for-profit, less-than-4-year institutions.

Table 18. Institution-level computer-assisted data entry completion rates, by institutional characteristic and abstraction method: 2008

Institutional characteristic and abstraction method[1]	Institutions providing lists	Institutions providing CADE		
		Number	Unweighted percent[2]	Weighted percent[3]
Total	1,730	1,670	96.5	96.1
Institutional level				
Less-than-2-year	100	90	94.0	93.0
2-year	500	480	95.8	97.6
4-year non-doctorate-granting	630	610	97.1	96.6
4-year doctorate-granting	500	490	96.8	96.4
Institutional control				
Public	880	850	96.8	96.5
Private not-for-profit	560	540	96.1	95.4
Private for-profit	290	280	96.2	96.1
Type of institution				
Public less-than-2-year	20	20	95.0	97.5
Public 2-year	410	390	96.1	95.9
Public 4-year non-doctorate-granting	190	180	98.4	97.6
Public 4-year doctorate-granting	260	250	96.9	97.1
Private not-for-profit less-than-4-year	20	20	100.0	100.0
Private not-for-profit 4-year non-doctorate-granting	320	310	95.7	93.8
Private not-for-profit 4-year doctorate-granting	230	220	96.4	95.7
Private for-profit less-than-2-year	70	70	93.2	91.9
Private for-profit 2-year-or-more	210	210	97.2	99.4
Abstraction method				
Self-CADE	1,090	30	96.5	95.5
Data-CADE	620	600	96.8	97.6
Field-CADE	20	20	87.0	93.8

[1] Institutional characteristics were verified (where possible) to correct classification errors on the sample frame.
[2] Percentage of institutions providing lists.
[3] The weight described in this column is a base weight.
NOTE: Detail may not sum to totals because of rounding. CADE = computer-assisted data entry.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

Student-level CADE completion rates are reported by institutional type and student type (table 19). Overall, the student-level CADE completion rate (the percentage of NPSAS-eligible sample members for whom a completed CADE record was obtained) was 96 percent (weighted). Weighted student-level completion rates ranged from 87 percent for private, not-for-profit, less-than-4-year institutions to 99 percent for public, 4-year, non-doctorate-granting institutions. Weighted completion rates by student type were 96 percent for undergraduate and 97 percent for graduate and first-professional students.

Table 19. Student-level computer-assisted data entry completion rates, by institutional characteristic and student type: 2008

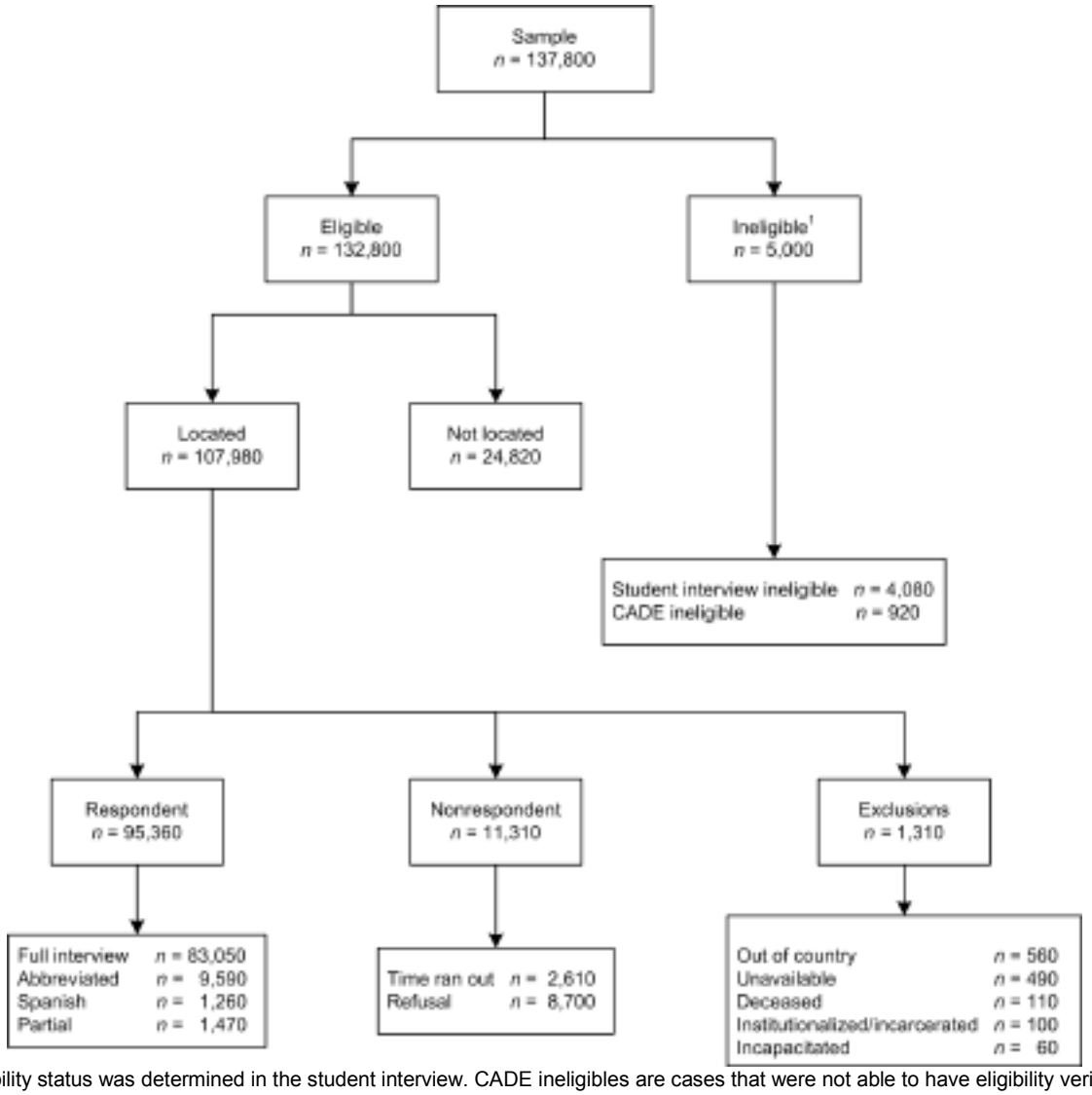
		CADE completes[1]		
Institutional characteristic and student type[2]	Eligible students[3]	Number	Unweighted percent	Weighted percent[4]
Total	132,800	126,620	95.3	96.2
Institutional level				
Less-than-2-year	7,950	7,210	90.7	92.0
2-year	40,770	38,410	94.2	95.1
4-year non-doctorate-granting	37,140	36,130	97.3	97.2
4-year doctorate-granting	46,940	44,880	95.6	97.0
Institutional control				
Public	84,240	80,690	95.8	96.3
Private not-for-profit	31,950	30,300	94.9	95.1
Private for-profit	16,610	15,630	94.1	97.4
Type of institution				
Public less-than-2-year	1,480	1,310	88.1	92.1
Public 2-year	37,010	35,210	95.1	95.2
Public 4-year non-doctorate-granting	15,850	15,650	98.8	99.0
Public 4-year doctorate-granting	29,910	28,530	95.4	97.1
Private not-for-profit less-than-4-year	1,790	1,570	88.0	87.3
Private not-for-profit 4-year non-doctorate-granting	13,930	13,190	94.6	93.8
Private not-for-profit 4-year doctorate-granting	16,230	15,550	95.8	96.3
Private for-profit less-than-2-year	6,050	5,500	90.9	91.6
Private for-profit 2-year-or-more	10,560	10,130	95.9	98.8
Student type				
Total undergraduate	118,030	112,450	95.3	96.1
Potential B&B student	24,290	23,360	96.2	96.8
Other undergraduates	93,740	89,080	95.0	96.0
Graduate/first-professional	14,770	14,180	96.0	96.7

[1] Eligible students who met the criteria for qualification as a CADE completion, which required an indication of financial aid receipt, enrollment status, tuition, degree program, or race/ethnicity in the CADE instrument. Numbers presented here may include students who were classified as study nonrespondents.
[2] Both institutional characteristics and student classifications were verified (where possible) to correct classification errors on the sample frame.
[3] Students determined to be eligible in CADE or the student interview (or both).
[4] The weight described in this column is a base weight.
NOTE: Detail may not sum to totals because of rounding. B&B = Baccalaureate and Beyond; CADE = computer-assisted data entry.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

3.3.3 Student Locating and Response Rates

Of the eligible student sample members, 81 percent (n = 107,980) were located (figure 10). An additional 5,000 sample members were determined to be ineligible for the study. Of the eligible sample members, 95,360 (72 percent) completed either a full interview (n = 83,050), or an English- or Spanish-language abbreviated interview (n = 9,590 for English and n = 1,260 for Spanish), or enough of the questionnaire that the interview could be considered partially completed (n = 1,470). (Questionnaires with a completed enrollment section but incomplete survey were considered *partial interviews*.)

Figure 10. NPSAS:08 student locating results: 2008



[1] Eligibility status was determined in the student interview. CADE ineligibles are cases that were not able to have eligibility verified within the student interview.
NOTE: Detail may not sum to totals because of rounding. CADE = computer-assisted data entry.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

3.4 Student Locating

Typically, students change residences throughout their college careers, often without updating their local contact information. Because locating a group as mobile as the NPSAS:08 student sample presented special challenges, various approaches were adopted to locate and interview the students in the NPSAS sample: an initial mailing and e-mail to all students, follow-up letters and e-mail messages to nonrespondents, telephone tracing (calling local and permanent numbers, as well as any other numbers obtained), and intensive tracing for especially hard-to-reach students (i.e., using consumer databases, web searches, and crisscross or reverse directories). Of the 132,800 eligible sample members, 81 percent were located (table 20). The highest location rates were for students attending public and private, not-for-profit, 4-year, doctorate-granting institutions (both 87 percent), while the lowest location rates were among those students attending private, for-profit, less-than-2-year institutions (64 percent). Finding graduate/first-professional students proved easier, with 91 percent of these students being successfully located, compared with 84 percent of B&B eligible students ($\chi = 16.71$, $p < .01$).

Table 20. Student locating, by institutional characteristic and student type: 2008

Institutional characteristic and student type[1]	Total	Located	
		Number	Unweighted percent
Total	132,800	107,980	81.3
Institutional level			
Less-than-2-year	7,950	5,210	65.5
2-year	40,770	30,930	75.9
4-year non-doctorate-granting	37,140	31,020	83.5
4-year doctorate-granting	46,940	40,830	87.0
Institutional control			
Public	84,240	68,950	81.8
Private not-for-profit	31,950	27,490	86.0
Private for-profit	16,610	11,550	69.5
Type of institution			
Public less-than-2-year	1,480	1,070	71.9
Public 2-year	37,010	28,390	76.7
Public 4-year non-doctorate-granting	15,850	13,480	85.0
Public 4-year doctorate-granting	29,910	26,020	87.0
Private not-for-profit less-than-4-year	1,790	1,290	72.4
Private not-for-profit 4-year non-doctorate-granting	13,930	12,040	86.4
Private not-for-profit 4-year doctorate-granting	16,230	14,160	87.3
Private for-profit less-than-2-year	6,050	3,860	63.8
Private for-profit 2-year-or-more	10,560	7,690	72.8
Student type			
Total undergraduate	118,030	94,600	80.2
Potential B&B student	24,290	20,470	84.3
Other undergraduate	93,740	74,140	79.1
Graduate/first-professional	14,770	13,380	90.6

[1] Both institutional characteristics and student types were verified to correct classification errors on the sampling frame.
NOTE: Numbers may not sum to totals due to rounding. Excludes 5,000 cases determined to be ineligible for the study.
B&B = Baccalaureate and Beyond.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

3.4.1 Database Batch Tracing Before Data Collection

To locate students for NPSAS:08, institutions were asked to provide local and permanent telephone numbers and addresses for students on their enrollment lists. This information was then confirmed or updated by matching to three locating databases: CPS, National Change of Address (NCOA), and Telematch.

The CPS database contained information from students who had applied for federal student aid. If a student was matched to the CPS database, additional locating information often was available, including new (or previous) local and permanent addresses and telephone numbers, contact information for the student's parents or guardians, and information about other possible contacts. Of the 123,020 cases sent to and processed through CPS before data collection ended, 80,610 (66 percent) were returned with new or confirmed information (table 21).

Table 21. Batch processing record match rates, by tracing source: 2008

Method of tracing	Number of records sent	Number of records matched	Percent matched[1]
CPS	123,020	80,610	65.5
NCOA	135,240	9,860	7.3
Telematch	137,800	65,550	47.6
Accurint Batch	32,970	10,550	32.0

[1] Percentage is based on the number of records sent for batch tracing. Since records were sent to multiple tracing sources, multiple record matches were possible.
NOTE: Detail may not sum to totals because of rounding. CPS = Central Processing System; NCOA = National Change of Address.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

All student address information from the institutions and from CPS was sent to NCOA for updating. NCOA, a database of change-of-address data submitted to the U.S. Postal Service, contains almost 113 million records. NCOA records are updated every 2 weeks and stored for 4 years. Of the 135,240 cases sent to NCOA for processing, 9,860 (7 percent) were returned with updated addresses.

To obtain telephone numbers matching the student addresses, all contact information obtained to this point was then sent to Telematch for batch processing. Telematch uses name, street address, and zip code as search criteria and returns either a telephone number update or confirmation, or an indication that no telephone match was found. Of the 137,800 cases sent to Telematch, about one half (48 percent) were returned with new or confirmed telephone numbers. Whenever telephone leads for a case were exhausted in the computer-assisted telephone interviewing case management system (CATI-CMS), the case was sent also to Accurint Batch. Accurint Batch is a flexible search vendor capable of providing a variety of contact information at low cost per case. In addition to verifying the provided address and phone number, this batch search can provide active/verified addresses and phone numbers; the name associated with located phone numbers; the best SSN for the subject, up to five aliases (or AKAs, "also known as"); dates of birth (DOBs); and previous/unverified addresses and phone numbers, with dates the subject was associated with the address, and deceased date, if applicable. This vendor provides an indicator that the phone number returned has been verified as active and belonging to a subject in the past 24 hours. Our experience has shown that Accurint Batch is becoming increasingly known for providing accurate, up-to-date phone numbers. Of the 32,970 cases sent to Accurint Batch, 10,550 (32 percent) were returned with new or confirmed telephone numbers.

3.4.2 Intensive Tracing During Data Collection

Sample members who did not complete an interview and who had not been reached were assigned to RTI's Tracing Operations Unit (TOPS) for intensive tracing. Intensive tracing entailed searches of public and proprietary databases, the Web, and various information directories. Overall, 20 percent of the eligible sample required intensive tracing (table 22). This rate ranged from a high of 33 percent for students from private, for-profit, less-than-2-year institutions to a low of 15 percent for students from private, not-for-profit, 4-year, doctorate-granting institutions.

Table 22. Students requiring intensive tracing, by institutional characteristic and student type: 2008

Institutional characteristic and student type[1]	Total	Cases requiring intensive tracing	
		Number	Percent
Total	132,800	27,140	20.4
Institutional level			
Less-than-2-year	7,950	2,400	30.2
2-year	40,770	10,560	25.9
4-year non-doctorate-granting	37,140	6,810	18.3
4-year doctorate-granting	46,940	7,370	15.7
Type of institution			
Public	84,240	17,410	20.7
Private not-for-profit	31,950	5,140	16.1
Private for-profit	16,610	4,590	27.6
Type of institution			
Public less-than-2-year	1,480	310	21.1
Public 2-year	37,010	9,710	26.2
Public 4-year non-doctorate-granting	15,850	2,600	16.4
Public 4-year doctorate-granting	29,910	4,790	16.0
Private not-for-profit less-than-4-year	1,790	400	22.4
Private not-for-profit 4-year non-doctorate-granting	13,930	2,300	16.5
Private not-for-profit 4-year doctorate-granting	16,230	2,440	15.0
Private for-profit less-than-2-year	6,050	1,990	32.9
Private for-profit 2-year-or-more	10,560	2,600	24.6
Student type			
Total undergraduate	118,030	24,890	21.1
Potential B&B student	28,290	3,830	15.8
Other undergraduate	93,740	21,060	22.5
Graduate/first-professional	14,770	2,250	15.2

[1] Both institutional characteristics and student types were verified to correct classification errors on the sampling frame.
NOTE: Detail may not sum to totals because of rounding. Excludes 5,000 cases determined to be ineligible for the study.
B&B = Baccalaureate and Beyond.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

Of the 27,140 sample members requiring intensive tracing, 15,900 (59 percent) were ultimately located, and 13,140 (48 percent) were interviewed (table 23).

Table 23. Locate and interview rates, by intensive tracing efforts: 2008

Intensive tracing status	Total	Located		Interviewed	
		Number	Percent	Number	Percent
Total	132,800	107,890	81.2	95,360	71.8
Intensive tracing required	27,140	15,900	58.6	13,140	48.4
No intensive tracing required	105,660	91,990	87.1	82,230	77.8

NOTE: Detail may not sum to totals because of rounding. Percentages are based on the total number of sample members within the row under consideration. Excludes 5,000 cases determined to be ineligible for the study.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

3.5 Interviewing Outcomes

Overall, 95,360 of approximately 132,800 eligible sample members (72 percent unweighted) completed either a full or partial NPSAS:08 student interview. Completion results are reported by institutional characteristic and student type (table 24). The weighted response rate was 71 percent overall and ranged from 56 percent for private, for-profit, less-than-2-year institutions to 77 percent for public, 4-year, doctorate-granting institutions.

Table 24. Student interview completion results, by institutional characteristic and student type: 2008

Institutional characteristic and student type[2]	Eligible students[3]	Completed interviews[1]		
		Number	Unweighted percent	Weighted percent[4]
Total	132,800	95,360	71.8	71.1
Institutional level				
Less-than-2-year	7,950	4,530	56.9	56.7
2-year	40,770	26,610	65.3	65.3
4-year non-doctorate-granting	37,140	27,680	74.5	73.1
4-year doctorate-granting	46,940	36,550	77.9	76.5
Institutional control				
Public	84,240	60,740	72.1	70.9
Private not-for-profit	31,950	24,490	76.6	75.9
Private for-profit	16,610	10,140	61.1	62.8
Type of institution				
Public less-than-2-year	1,480	900	61.0	58.5
Public 2-year	37,010	24,380	65.9	65.6
Public 4-year non-doctorate-granting	15,850	12,040	76.0	74.6
Public 4-year doctorate-granting	29,910	23,410	78.3	77.1
Private not-for-profit less-than-4-year	1,790	1,120	62.9	64.6
Private not-for-profit 4-year non-doctorate-granting	13,930	10,810	77.5	75.9
Private not-for-profit 4-year doctorate-granting	16,230	12,560	77.4	76.5
Private for-profit less-than-2-year	6,050	3,370	55.7	55.8
Private for-profit 2-year-or-more	10,560	6,770	64.1	64.2
Student type				
Total undergraduates	118,030	83,190	70.5	69.7
Potential B&B student[5]	24,290	18,000	74.1	75.9
Other undergraduate	93,740	65,190	69.5	68.9
Graduate/first-professional	14,770	12,180	82.4	79.5

[1] Eligible students who met the criteria for qualification as a student interview completion, which required completing at least a partial interview.
[2] Both institutional characteristics and student types were verified to correct classification errors on the sampling frame.
[3] Excludes 5,000 cases determined to be ineligible for the study.
[4] The weight described in this column is a base weight.
[5] Approximately 10 students completed the interview questions needed to determine B&B eligibility but did not meet the minimum requirements to be classified as an interview respondent.
NOTE: Detail may not sum to totals because of rounding. B&B = Baccalaureate and Beyond.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

3.5.1 Interview Completion by Data Collection Phase

As described in section 2.4.6, NPSAS:08 involved three phases of student interviewing: (1) the early response phase, (2) the production-interviewing phase, and (3) the nonresponse-conversion phase. The response rates for each phase of data collection are tabulated (table 25).

Table 25. Student interview response rates, by data collection phase, and by institutional characteristic and student type: 2008

Institutional characteristic and student type	Number eligible	Total complete			Data collection phase: Early response phase			Production-interviewing phase			Nonresponse-conversion phase		
		Number	Unweighted percent of eligible	Weighted percent of eligible	Number	Unweighted percent of total	Weighted percent of total	Number	Unweighted percent of total	Weighted percent of total	Number	Unweighted percent of total	Weighted percent of total
Total	132,800	95,360	71.8	71.1	45,870	34.5	34.7	10,020	7.5	8.4	39,480	29.7	28.0
Institutional level													
Less-than-2-year	7,950	4,530	56.9	56.7	1,120	14.0	13.8	340	4.3	4.1	3,070	38.6	38.8
2-year	40,770	26,610	65.3	65.3	10,640	26.1	27.2	3,190	7.8	9.3	12,780	31.3	28.9
4-year non-doctorate-granting	37,140	27,680	74.5	73.1	13,870	37.4	36.4	2,560	6.9	7.7	11,250	30.3	29.0
4-year doctorate-granting	46,940	36,550	77.9	76.5	20,240	43.1	42.5	3,920	8.4	8.2	12,390	26.4	25.8
Institutional control													
Public	84,240	60,736	72.1	70.9	29,320	34.8	34.4	6,980	8.3	9.2	24,440	29.0	27.3
Private not-for-profit	31,950	24,490	76.6	75.9	13,300	41.6	40.9	2,400	7.5	7.8	8,780	27.5	27.1
Private for-profit	16,610	10,140	61.1	62.8	3,250	19.6	24.7	640	3.9	3.5	6,260	37.7	34.7
Institutional type													
Public less-than-2-year	1,480	900	61.0	58.5	290	19.3	18.7	120	8.4	8.1	490	33.3	31.8
Public 2-year	37,010	24,380	65.9	65.6	9,950	26.9	27.6	2,970	8.0	9.5	11,450	31.0	28.6
Public 4-year non-doctorate-granting	15,850	12,040	76.0	74.6	6,230	39.3	38.6	1,230	7.8	9.1	4,580	28.9	27.0
Public 4-year doctorate-granting	29,910	23,410	78.3	77.1	12,860	43.0	42.7	2,640	8.8	8.9	7,910	26.5	25.5
Private not-for-profit less-than-4-year	1,790	1,120	62.9	64.6	350	19.8	18.7	130	7.2	6.5	640	35.9	39.4
Private not-for-profit 4-year non-doctorate-granting	13,930	10,810	77.6	75.9	5,830	41.9	39.5	1,030	7.4	7.6	3,950	28.3	28.8
Private not-for-profit 4-year doctorate-granting	16,230	12,560	77.4	76.5	7,110	43.8	43.5	1,250	7.7	8.0	4,200	25.9	25.0
Private for-profit less-than-2-year	6,050	3,370	55.7	56.8	780	12.8	13.5	210	3.5	4.0	2,380	39.4	38.3
Private for-profit 2-year-or-more	10,560	6,770	64.1	64.2	2,470	23.4	26.9	430	4.1	3.4	3,870	36.7	33.9
Student type													
Total undergraduates	118,030	83,190	70.5	69.6	38,950	33.0	32.8	8,570	7.3	8.3	35,660	30.2	28.5
Potential B&B students[1]	24,290	18,000	74.1	76.0	9,870	40.6	41.3	1,690	7.0	7.9	6,430	26,5	26.7
Other undergraduate	93,740	65,190	69.5	69.0	29,080	31.0	32.0	6,880	7.3	8.3	29,230	31.2	28.7
Graduate/first-professional	14,770	12,180	82.4	79.5	6,910	46.8	45.5	1,440	9.8	9.0	3,820	25.9	25.0

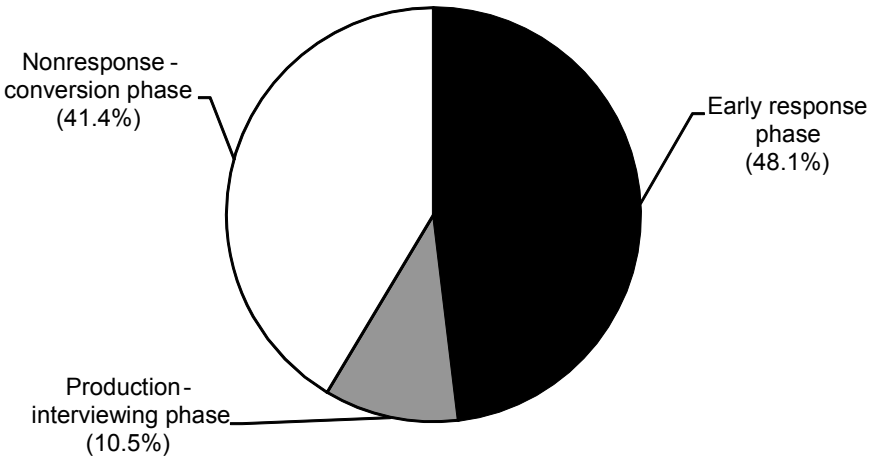
[1] Approximately 10 students completed the interview questions needed to determine B&B eligibility but did not meet the minimum requirements to be classified as an interview respondent.

NOTE: Numbers may not sum to totals due to rounding. Excludes 5,000 cases determined to be ineligible for the study. B&B = Baccalaureate and Beyond.

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

Early response phase. In the first phase, after initial locating of sample members, a period of 3 weeks was allotted for students to complete the web interview or to call the help desk to complete the interview by telephone. About 45,870 respondents (35 percent of the eligible sample) completed their interviews during the early response phase and therefore received the $30 incentive; these early completions represented 48 percent of all interviews completed for NPSAS:08 (figure 11). Moreover, 43 percent of 4-year doctorate-granting sample members completed the interview during the early response period, compared with only 14 percent of less-than-2-year sample members ($z = 19.19$, $p < .01$; table 25). Graduate students were more likely to complete the interview during this data collection phase than undergraduate students (47 percent, compared with 33 percent; $z = 20.17$, $p < .01$).

Figure 11. Percentage distribution of interview completions, by data collection phase: 2008



SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

An experiment was conducted in the NPSAS:08 field test to evaluate the effectiveness of prompting calls in increasing web interview response rates during the early response period. Results from this experiment showed that prompting calls increased web response rates during the early response period (Cominole et al. 2008). Approximately 22 percent of the sample members who were selected for prompting completed the field test interview during the early response period, compared with 18 percent of those who were not selected for prompting, $\chi^2(1, N = 2{,}110) = 9.58$, $p < .001$. Because prompting the entire NPSAS:08 sample would have been very costly, paradata were used to strategically select cases for prompting. The results for the NPSAS:08 sample members who received prompting calls are displayed in table 26.

Table 26. Results of prompting during the early response phase of data collection: 2008

Prompted	Total	Completed interview during the early response phase	
		Number	Percent
Total	5,350	840	15.6
Prompting successful	2,110	460	21.7
Spoke with sample member	980	310	31.7
Left message on voice mail	200	30	14.4
Left message with household member	940	120	12.8
Prompting unsuccessful	3,240	380	11.7

NOTE: Detail may not sum to totals because of rounding. Any case that was selected for prompting but completed the interview prior to being prompted or for whom no telephone number was available was excluded from the analysis.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

Any case that was selected for prompting but completed the interview before being prompted, or any for whom no telephone number was available, was excluded from the analysis. A prompting call was considered to be successful if either contact was made directly with the sample member, or a message was left with a contact or on voice mail. Of the 5,350 cases that received at least one prompting call, 2,110 were successfully prompted. Approximately 22 percent of the sample members who were successfully prompted completed the interview during the early completion phase, as opposed to the 12 percent of those who were not successfully prompted ($z = 2.87$, $p < .001$). Furthermore, sample members for whom contact was made directly during the prompting call had a 32 percent response rate during the early response period, as opposed to the 13 percent response rate for sample members whose prompting call resulted in a message left with a household member ($z = 4.27$, $p < .001$).

Production-interviewing phase. During the production-interviewing phase, interviewers telephoned sample members to conduct interviews. Approximately 10,020 interviews (8 percent of the eligible sample) were completed in this phase. No incentive was offered for completing interviews at this phase.

Nonresponse-conversion phase. The third phase of student interviewing involved the nonresponse conversion of refusals and the locating of students who were difficult to locate and to whom several calls had been placed without results. About 39,480 respondents (30 percent of the eligible sample) completed interviews during this phase and received a $30 incentive.

3.5.2 Interview Completion by Mode of Administration

Students responding to NPSAS:08 varied significantly in terms of the mode by which they completed the survey. Of all completed interviews, more were obtained by web respondents than by telephone respondents: 70 percent, as opposed to 30 percent weighted (table 27). Student interview completion rates, by institutional characteristic and student type, are reported (table 28). Sample members from less-than-2-year institutions had the highest percentage of telephone interview completions (63 percent). Graduate or first-professional students (78 percent) were more likely than total undergraduate students (68 percent) to complete a web interview ($z = 19.62$, $p < .001$).

Table 27. Student interview completion rates: 2008

Completion mode	Completed interviews		
	Number	Unweighted percent	Weighted percent[1]
Total	95,360	100.0	100.0
Web interviews	65,920	69.1	69.9
Early response period	43,730	45.9	47.0
With prompting	22,190	23.3	22.8
Telephone interviews	29,440	30.9	30.1

[1] The weight described in this column is a base weight.
NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

Table 28. Student interview completion rates by mode of completion, and by institutional characteristic and student type: 2008

Institutional characteristic and student type	Total	Interviewed			Mode of completion					
					Web			Telephone		
		Number	Percent	Weighted	Number	Percent	Weighted	Number	Percent	Weighted
Total	132,800	95,360	71.8	71.1	65,920	69.1	69.9	29,440	30.9	30.1
Institutional level										
Less-than-2-year	7,950	4,530	56.9	56.9	1,690	37.4	36.2	2,830	62.6	63.8
2-year	40,770	26,610	65.3	65.3	15,870	59.6	61.4	10,740	40.4	38.6
4-year non-doctorate-granting	37,140	27,680	74.5	73.1	19,720	71.2	71.0	7,960	28.8	29.0
4-year doctorate-granting	46,940	36,550	77.9	77.1	28,640	78.4	77.9	7,910	21.6	22.1
Institutional control										
Public	84,240	60,740	72.1	70.9	42,370	69.8	69.7	18,360	30.2	30.3
Private not-for-profit	31,950	24,490	76.6	75.9	18,400	75.1	75.1	6,090	24.9	24.9
Private for-profit	16,610	10,140	61.1	62.8	5,150	50.8	58.2	4,990	49.2	41.8
Type of institution										
Public less-than-2-year	1,480	900	61.0	58.5	400	44.4	44.6	500	55.6	55.4
Public 2-year	37,010	24,380	65.9	65.6	14,840	60.9	62.1	9,540	39.1	37.9
Public 4-year non-doctorate-granting	15,850	12,040	76.0	74.6	8,790	73.0	72.7	3,250	27.0	27.3
Public 4-year doctorate-granting	29,910	23,410	78.3	77.1	18,340	78.3	78.1	5,070	21.7	21.9
Private not-for-profit less-than-4-year	1,790	1,120	62.9	64.6	530	47.2	43.5	590	52.8	56.5
Private not-for-profit 4-year non-doctorate-granting	13,930	10,810	77.5	75.9	7,980	73.9	72.9	2,820	26.1	27.1
Private not-for-profit 4-year doctorate-granting	16,230	12,560	77.4	76.5	9,890	78.7	78.9	2,670	21.3	27.1
Private for-profit less-than-2-year	6,050	3,370	55.7	55.8	1,220	36.3	36.8	2,150	63.7	63.2
Private for-profit 2-year-or-more	10,560	6,770	64.1	64.2	3,930	58.0	62.1	2,840	42.0	37.9
Student type										
Total undergraduates	118,030	83,190	70.5	69.6	56,440	67.8	68.2	26,750	32.2	31.8
Potential B&B student[1]	24,290	18,000	74.1	75.9	14,020	77.9	77.5	3,980	22.1	22.5
Other undergraduate	93,740	65,190	69.5	68.9	42,410	65.1	67.1	22,780	34.9	32.9
Graduate/first-professional	14,770	12,180	82.4	79.5	9,490	77.9	78.2	2,690	22.1	21.8

[1] Approximately 10 students completed the interview questions needed to determine B&B eligibility but did not meet the minimum requirements to be classified as an interview respondent.

NOTE: Detail may not sum to totals because of rounding. All percentages are unweighted and based on the number of eligible students within the row under consideration.

B&B = Baccalaureate and Beyond.

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

3.6 Completeness of Data Records Among Study Respondents

A study respondent was any eligible student for whom sufficient data were obtained from one or more sources. The sources used to define response status were institutional student records (CADE), the student interview, and the U.S. Department of Education's CPS.

Most study respondents exceeded classification criteria to be study respondents. Overall, data completeness was high (table 29). Approximately 71 percent of the study respondents had both student interview and CADE student record data. Slightly less than 50 percent of study respondents had data from all three of the primary data sources. Nearly all of the study respondents had student record data (CADE) from their NPSAS institution (as shown in the first, second, third, and fifth rows of table 29). Additionally, almost two thirds of respondents had a federal aid application in the CPS database for the 2007–08 academic year (as shown in the first, third, fourth, and seventh rows of table 29).

Table 29. Data completeness for NPSAS:08 study respondents, by data source: 2008

	Study respondents	
Sources of data	Number	Percent
Total	127,700	100.0
Student interview, CADE student record, and CPS data	59,270	46.4
Student interview and CADE student record data	31,850	24.9
CADE student record and CPS data	19,480	15.3
Student interview and CPS data	2,010	1.6
CADE student record data only	12,490	9.8
Student interview data only	1,820	1.4
CPS data only	780	0.6

NOTE: Detail may not sum to totals because of rounding. CADE = computer-assisted data entry; CPS = Central Processing System.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

This page intentionally left blank.

Chapter 4. Evaluation of Field Operations and Data Quality

Presented in this chapter are the results of data-quality analyses and an account of procedures that ensured quality control in the 2007–08 National Postsecondary Student Aid Study (NPSAS:08). Efforts to identify students eligible for the Baccalaureate and Beyond Longitudinal Study (B&B) are also discussed.

4.1 Enrollment List Collection

4.1.1 Early Contacting Activities

As learned from NPSAS:04, early effort to contact institutions accelerates the collection of student lists: lists for NPSAS:04 began arriving in January 2004, a month earlier than in previous cycles without early contacting. The spring 2003 initiation of early contacting was driven chiefly by the data collection schedule for a sister project in the National Study of Faculty and Students: the National Study of Postsecondary Faculty. Because NPSAS:08 lists were not due until January 2008, the September 2007 start date for recruitment presumably allowed institutions ample time to prepare for NPSAS data collection and resolve any impeding circumstances.

Interestingly, the cumulative flow of cases from January through June 2008 shows that data collection for 2008, compared with that for 2004, was further accelerated (table 30). Other factors, such as differences between the NPSAS:04 sample and the NPSAS:08 sample, may have contributed to this acceleration; nevertheless, the cumulative reported results suggest that a fall start date for recruitment of institutions does enable schools with January and February early due dates to provide their lists in a timely manner.

Table 30. Cumulative flow of enrollment list receipt: 2004 and 2008

	Cumulative percentage of lists received[1]	
Month	NPSAS:04	NPSAS:08
1	12.5	20.4
2	38.4	50.6
3	58.8	69.2
4	75.4	83.3
5	88.7	95.3
6	98.2	99.4
7	100.0	100.0

[1] Based on eligible institution sample of 1,360 institutions in 2004 and 1,730 institutions in 2008.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2003–04 National Postsecondary Student Aid Study (NPSAS:04), and 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

4.1.2 Institutional Participation

Institutional participation was evaluated for potential effects of prior NPSAS participation. Summary results of these analyses are shown in table 31. Among eligible institutions, the NPSAS:08 enrollment list provision rate among the 1,400 institutions that had previously participated in NPSAS was 91 percent, which is statistically different than the rate among institutions that had not previously participated (83 percent), $\chi^2(1, N = 1{,}940) = 29.74, p < .05$.

Table 31. NPSAS:08 institutional enrollment list participation, by prior NPSAS participation: 2008

Institutional characteristic[1]	Eligible institutions	No prior NPSAS participation			Participated in NPSAS at least once		
			Provided lists			Provided lists	
		Number	Number	Percent[2]	Number	Number	Percent[3]
All institutions	1,940	550	450	82.8	1,400	1,280	91.4
Institutional level							
Less-than-2-year	120	80	60	77.3	50	40	91.3
2-year	560	190	160	84.5	370	340	92.2
4-year non-doctorate-granting	700	220	190	84.2	480	440	92.3
4-year doctorate-granting	560	60	50	79.0	500	450	90.0
Institutional control							
Public	960	170	150	88.5	790	730	92.7
Private not-for-profit	640	160	130	80.0	480	440	89.9
Private for-profit	340	220	180	80.5	120	110	89.3
Type of institution							
Public less-than-2-year	20	10	10	87.5	10	10	92.9
Public 2-year	450	110	100	88.6	330	310	92.8
Public 4-year non-doctorate-granting	200	30	30	89.7	170	160	95.3
Public 4-year doctorate-granting	290	10	10	85.7	280	250	90.9
Private not-for-profit less-than-4-year	20	10	10	87.5	10	10	81.8
Private not-for-profit 4-year non-doctorate-granting	370	110	90	82.1	250	230	90.9
Private not-for-profit 4-year doctorate-granting	260	40	30	72.5	220	200	89.1
Private for-profit less-than-2-year	90	60	50	76.2	30	30	89.7
Private for-profit 2-year-or-more	250	160	130	82.3	90	80	89.1

[1] Institutional characteristics were verified by the institutions to correct classification errors on the sampling frame.
[2] Percentages are based on the count of eligible institutions with no prior NPSAS participation within the row under consideration.
[3] Percentages are based on the count of eligible institutions with prior NPSAS participation within the row under consideration.
NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

Institutional participation was also examined in terms of the 2005 Carnegie classification categories, as shown in table 32. Table 33 shows the number of Historically Black Colleges and Universities participating in the current and prior NPSAS rounds.

Table 32. Distribution of participating NPSAS:08 institutions, by 2005 Carnegie institutional classification: 2008

2005 Carnegie institutional classification	Number	Percent
All institutions	1,730	100.0
Not classified	150	8.6
Associate's—public rural-serving small	20	1.3
Associate's—public rural-serving medium	90	5.2
Associate's—public rural-serving large	70	3.9
Associate's—public suburban single campus	50	3.1
Associate's—public suburban multicampus	50	3.0
Associate's—public urban single campus	20	1.1
Associate's—public urban multicampus	80	4.8
Associate's—public special use	10	0.4
Associate's—private not-for-profit	10	0.4
Associate's—private for-profit	80	4.9
Associate's—public 2-year under-4-year	10	0.5
Associate's—public 4-year primarily associate's	10	0.3
Associate's—private not-for-profit 4-year primarily associate's	#	0.1
Associate's—private for-profit 4-year primarily associate's	20	1.0
Research (very high research activity)	80	4.5
Research (high research activity)	90	5.0
Doctoral/research universities	60	3.6
Master's (larger programs)	260	14.8
Master's (medium programs)	120	6.8
Master's (smaller programs)	70	4.1
Bachelor's colleges—arts and sciences	100	5.6
Bachelor's colleges—diverse fields	120	6.7
Bachelor's/associate's colleges	40	2.1
Special focus—theological	30	1.4
Special focus—medical	30	1.6
Special focus—other health professions	10	0.6
Special focus—engineering	#	0.2
Special focus—other technology	10	0.6
Special focus—business/management	20	1.3
Special focus—art, music, and design	30	1.6
Special focus—law	10	0.3
Special focus—other special focus	10	0.3
Tribal colleges	10	0.3

Rounds to zero.
NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

Table 33. NPSAS:08 participation of Historically Black Colleges and Universities: 2008

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

NOTE: Detail may not sum to totals because of rounding. HBCUs = Historically Black Colleges and Universities.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

4.1.3 Quality of Enrollment Lists

Each sample institution was asked to provide an electronic list of eligible students either by sending an encrypted, password-protected file attached to an e-mail message, or by uploading the list to a secure website. Institutions were allowed to fax a list to a secure location as a last resort—and only after the institution submitted a test of nonsensitive data to confirm that the fax was sent to the correct location; however, no faxed lists were received.

Once lists were received, they were evaluated in terms of appropriateness of format and documentation (relative to instructions provided), as well as for the accuracy of student counts (for a description of quality control procedures, see chapter 2). Table 34 presents the major types of discrepancies encountered. About 31 percent of the institutions provided lists with one or more such problems. The most common problem was that enrollment counts were out of bounds when compared with the Integrated Postsecondary Education Data System (IPEDS) (25 percent for all institutions). The check was not suspended or relaxed (unlike some prior rounds of NPSAS) because many of the institutions that were called about the discrepancy indicated that the enrollment list counts were, in fact, incorrect.

In the event that an enrollment list failed the quality control check, RTI staff contacted the institution to resolve the problem or obtain a new list. After any necessary revisions, all but three lists submitted were usable for selecting the student sample.[25]

[25] These institutions were classified as nonparticipants.

Table 34. Types of discrepancies encountered with student enrollment lists, by highest level of offering: 2008

Highest level of offering	Type of discrepancy encountered[1]	Number	Percent[2]
All institutions (*n* = 1,170)	None	1,190	69.1
	Count out of bounds[3]	430	24.8
	Unreadable file/list	10	0.8
	Could not identify student strata	30	1.4
	Multiple problems	40	2.2
	Other	20	0.9
Less-than-2-year (*n* = 100)	None	70	66.0
	Count out of bounds[3]	30	32.0
	Unreadable file/list	#	#
	Could not identify student strata	#	#
	Multiple problems	#	2.0
	Other	#	#
2-year (*n* = 500)	None	420	84.6
	Count out of bounds[3]	60	12.8
	Unreadable file/list	10	1.2
	Could not identify student strata	#	#
	Multiple problems	#	#
	Other	#	0.8
4-year non-doctorate-granting (*n* = 630)	None	400	63.4
	Count out of bounds[3]	190	29.9
	Unreadable file/list	#	0.5
	Could not identify student strata	10	1.8
	Multiple problems	20	2.9
	Other	#	0.6
4-year doctorate-granting (*n* = 500)	None	310	61.5
	Count out of bounds[3]	150	28.9
	Unreadable file/list	#	0.8
	Could not identify student strata	10	2.8
	Multiple problems	20	3.6
	Other	10	1.4

Rounds to zero.
[1] Categories are mutually exclusive, with an institution being included in only one category within highest level of offering.
[2] Percentages are based on the number of institutions within each institutional type.
[3] Out of bounds as compared with the 2004–05 IPEDS. (The 2005–06 IPEDS were used for institutions included due to freshening.)
NOTE: Detail may not sum to totals because of rounding. IPEDS = Integrated Postsecondary Education Data System.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

4.2 Student Record Abstraction

Procedures to abstract information from institutional student records by means of computer-assisted data entry (CADE) were first implemented in NPSAS:93. Since then, the procedures have improved for each study cycle, enhancing the effectiveness and user-friendliness of the approach, particularly for institutional staff. Most notably, these improvements include the web-based CADE system (web-CADE), used for self-administration by institutional staff and by field interviewers, and the option of submitting data as electronic files (data-CADE). In NPSAS:08, the self-CADE and field-CADE instruments were redesigned for a more user-friendly interface. Changes included the addition of a menu that allows the user to select any of the eight sections (rather than completing the sections sequentially) for any student and, once the section is selected, displays all items in it on one screen instead of requiring navigation to a new screen for every few items.

Other procedures facilitated the timeliness of CADE completion: (1) maintaining a help desk to assist institutions with data preparations, (2) offering data-CADE earlier in the data collection period, (3) scheduling calls to answer questions and to prompt self-CADE and data-CADE institutions to complete data abstraction, (4) prescheduling institutions for field staff, and (5) scheduling weekly conference calls with field staff to assess their progress.

4.2.1 Preloading Data Into the Computer-Assisted Data Entry System

To reduce the data-entry effort associated with institutional student record abstraction, specific elements were loaded into CADE records before data collection at the institutions (table 35). This preloading entailed customization of the financial aid award section of CADE to include nonfederal aid commonly awarded within the particular institution. This feature was first implemented during NPSAS:96; such customization has proven highly successful for NPSAS in two ways: (1) it has decreased burden by preventing institutions from having to enter certain data elements if these elements are already accurate, and (2) it has increased accuracy by checking specific data elements entered by the institution against preloaded data.

Data were preloaded from various sources, including IPEDS, the National Association of State Student Grant and Aid Programs State Aid Report, the institutional coordinator, and enrollment lists. The most extensive set of preloaded data came from the Central Processing System (CPS) for federal financial aid applicants and was used in two different ways. One use was to preload items to validate the data entered by users. Whenever a user's entry differed from what was preloaded from CPS, the user received a warning message indicating that the information entered was different from the information on record. The user could then choose to keep his or her original entry or change that entry. The variables involved in this use were citizenship status, veteran status, and student date of birth. The second way data were used was to prefill some items, allowing the user to confirm information in its unaltered state when it was correct and providing updates if necessary. These items were the student's address, phone number, driver's license number, driver's license state, dependency status, and expected family contribution.

Table 35. Source of data elements preloaded into the computer-assisted data entry system: 2008

CADE data element set	Data source
Institutional name and ID	IPEDS
Names of most common state financial aid awards	NASSGAP report
Names of most common institutional financial aid awards	Institutional coordinator
Institutional clock/credit hour indicator	Institutional coordinator
Institutional term names and dates	Institutional coordinator
GPA scale	Institutional coordinator
Student name, SSN from institutional records[1]	Enrollment list
Student type indicator (undergraduate/graduate/first-professional)	Enrollment list
Student local and permanent addresses	Enrollment list
Major CIP code	Enrollment list
Major text string	Enrollment list
Student date of birth, veteran status, and citizenship[2]	CPS record
Student address, phone number, driver's license number and state	CPS record
Student dependency and expected family contribution	CPS record
Flag indicating whether student matched to CPS	CPS record

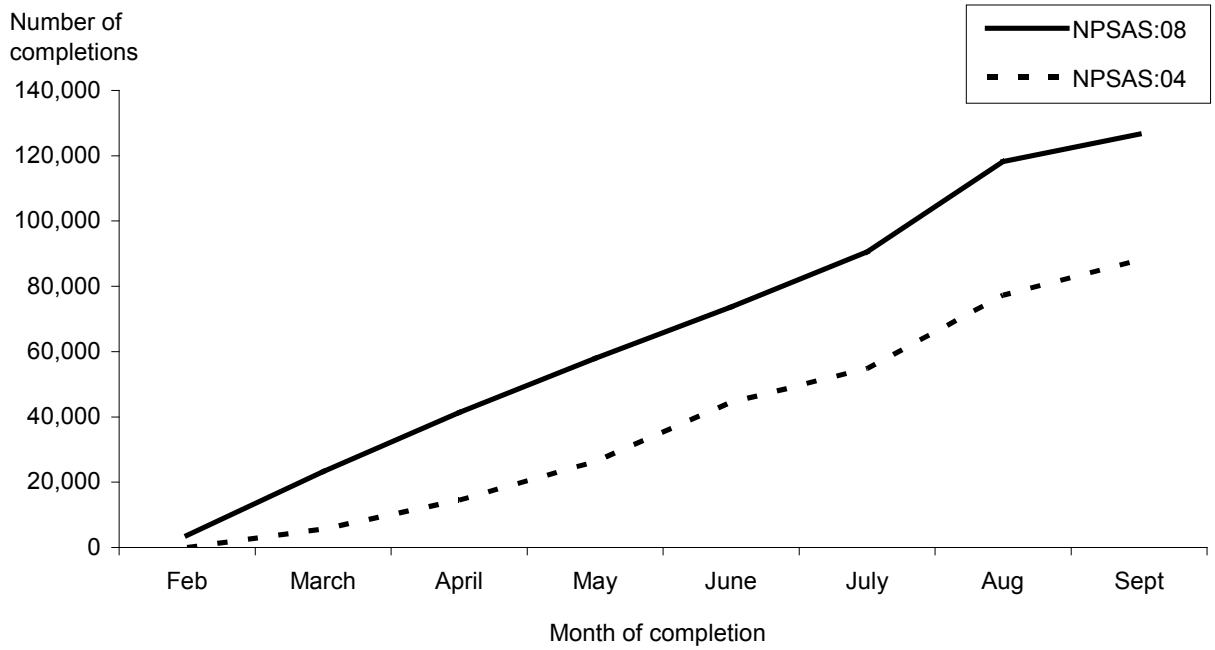
[1] SSN was not editable and was displayed in web-CADE if the student matched to CPS. If the student did not match to CPS, nothing was displayed, but a soft check was given if the user entered data that conflicted with what was on the enrollment list.
[2] These variables were not displayed on screen, but a confirmation box was displayed if the user entered data that conflicted with what was on the CPS record.
NOTE: CADE = computer-assisted data entry; CIP = Classification of Instructional Program; CPS = Central Processing System; GPA = grade point average; IPEDS = Integrated Postsecondary Education Data System; NASSGAP = National Association of State Student Grant and Aid Programs; SSN = Social Security number.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

4.2.2 Timeliness of Record Abstraction

The CADE system was prepared for each institution as enrollment lists were received, samples were selected, and matching to CPS was completed. Institutions that opted to provide data by means of self-CADE began receiving notification that the system was ready for them to begin data entry in early February 2008. The data-CADE specifications and system for uploading files were available to institutions in late February. By contrast, in NPSAS:04, data-CADE was unavailable to institutions until April. The first set of field-CADE data collectors began record abstraction activities in June 2008. Initialization of CADE systems continued through August 2008.

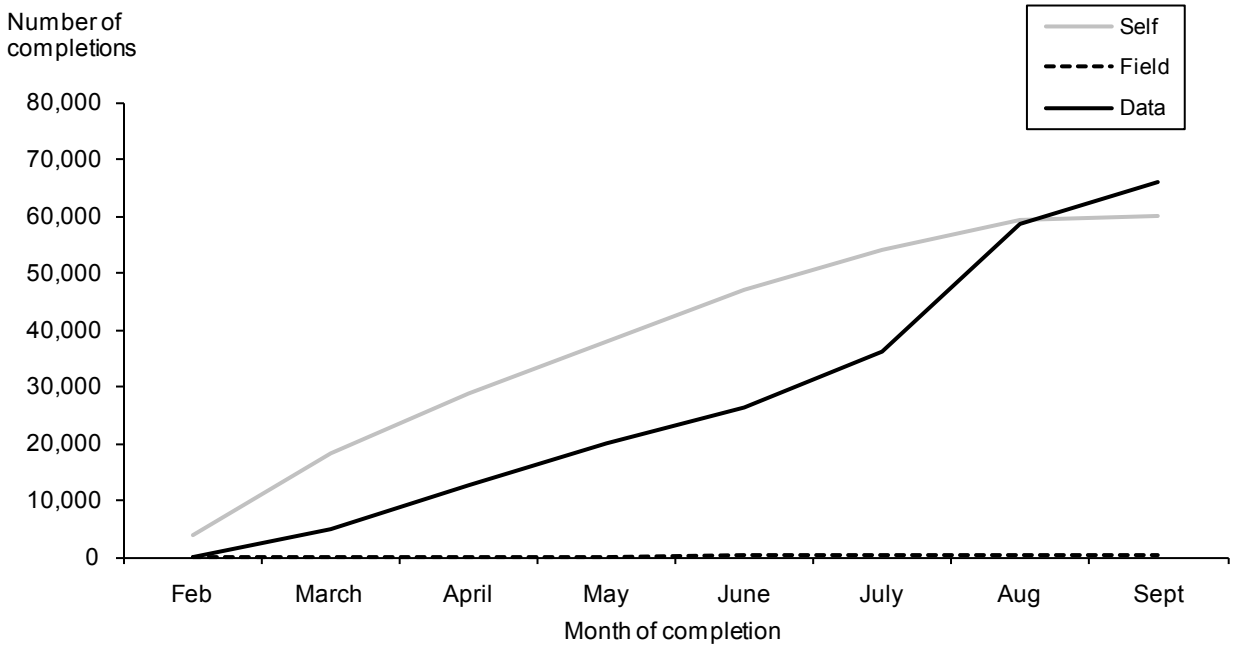
Data were collected on many more cases in NPSAS:08 than in NPSAS:04 (126,620 cases vs. 88,050 cases, respectively). Both NPSAS:08 and NPSAS:04 showed an increase in the number of CADE completions in late summer (figure 12). This increase was due primarily to the large number of institutions completing records by data-CADE (figure 13). Data-CADE was used by most institutional systems that provided data files for multiple institutions. By contrast, self-CADE showed a relatively steady flow of completed cases, which leveled off near the end of data collection.

Figure 12. Cumulative flow of computer-assisted data entry completions: 2004 and 2008



SOURCE: U.S. Department of Education, National Center for Education Statistics, 2003–04 National Postsecondary Student Aid Study (NPSAS:04), and 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

Figure 13. Cumulative flow of computer-assisted data entry completions, by abstraction mode: 2008



SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

4.2.3 Computer-Assisted Data Entry Data Completeness

As discussed in section 3.3.2, a student-level CADE completion required nonmissing data for at least one of the following critical variables:

- receipt of financial aid;
- enrollment (attendance status and enrollment intensity);
- tuition;
- degree program; or
- race.

According to this definition, 96 percent (weighted) of the eligible sample students were classified as CADE "completes" (see table 19). Unsurprisingly, because record-keeping systems vary dramatically among institutions, item-level response rates differed across data elements (table 36). Moreover, not all data elements were available at every institution; however, most of the key data elements showed a high percentage of item-level completeness.

Overall, item-level response rates were high—only two items had high rates of missing data: marital status and additional phone numbers. Student records frequently lack these items. Response rates varied somewhat by mode of abstraction; in general, data-CADE showed the highest rates of missing data. In contrast to the pattern in NPSAS:04, field-CADE had higher overall rates of item completion than self-CADE for all but three items (marital status, high school completion type, and race). Data-CADE showed the lowest item-level completion rates for all but three items (veteran status, race, and Hispanic descent). Both self-CADE and field-CADE used real-time edit checks and verifications, while data-CADE checks were performed after file creation, when the user uploaded the files. This difference may have contributed to the higher rate of item-level missing data for data-CADE.

Table 36. Computer-assisted data entry item-level completion rates, by method of abstraction: 2008

Data element	Number	Item response rates[1]			
		Total	Self-CADE	Field-CADE	Data-CADE
Total student record respondents	126,280	100.0	100.0	100.0	100.0
Student characteristics					
Date of birth	124,480	98.6	98.6	100.0	98.6
Gender	125,150	99.1	99.3	100.0	98.9
Marital status	57,530	45.6	55.0	44.4	37.0
Citizenship	124,340	98.5	100.0	100.0	97.1
Veteran status	118,140	93.6	87.2	96.4	99.3
High school completion type	85,780	76.3	77.4	76.1	75.4
Race	100,550	79.6	80.4	70.5	79.0
Hispanic descent	104,130	82.5	78.0	84.8	86.5
At least one phone number	115,920	91.8	93.9	95.4	89.9
At least two phone numbers	21,940	17.4	18.8	22.2	16.1
Enrollment					
Type of degree program	117,870	93.3	91.2	96.1	91.3
Student class level	119,050	94.3	95.7	98.3	92.9
Tuition jurisdiction classification	118,020	93.5	99.1	99.8	88.3
Total tuition amount	118,610	93.9	97.3	99.5	90.9
Financial aid[2]					
Any aid received (Y/N)	125,580	99.5	100.0	100.0	99.0
Federal aid received (Y/N)	125,570	99.4	99.9	100.0	99.0
State aid received (Y/N)	125,580	99.5	100.0	100.0	99.0
Undergraduate aid received (Y/N)	125,570	99.4	99.9	100.0	99.0
Graduate aid received (Y/N)[3]	126,270	100.0	100.0	100.0	100.0
Other aid received (Y/N)	125,570	99.4	99.9	100.0	99.0
Total financial aid amount[4]	125,570	99.4	99.9	100.0	99.0
Expected family contribution (EFC) amount	78,320	88.8	96.2	98.6	82.9

[1] Response rate is based on the number of students to whom the item applied.
[2] All financial aid gate items were logically coded for data-CADE students, based on the presence or absence of amounts in the nested items.
[3] The graduate aid gate question was logically imputed "No" if the student was an undergraduate.
[4] Total financial aid amount was computed by summing the amounts entered for each specific aid program. If the financial aid gate item was missing, then the financial aid amount item was also missing.
NOTE: Detail may not sum to totals because of rounding. CADE = computer-assisted data entry.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

4.3 Student Interviewing

4.3.1 Timing Analysis

The time students expended to complete the NPSAS:08 interview was examined, with emphasis on differences by completion mode, internet connection speed, and student type. To calculate the time burden, the NPSAS:08 student interview was developed with two time stamps embedded on each screen. The *start* timer recorded the time that the respondent's or interviewer's computer screen first displayed the web page. The *end* timer recorded the time that the respondent or the interviewer clicked the "Next" button to submit the responses from a particular page. From the two time stamp variables, an on-screen time and a transit time were calculated.

The timing analysis applied only to complete interviews with one-session logins. Multisession cases, including cases that stopped and later resumed the interview and those automatically logged off because of an extended period of inactivity, were excluded from analysis. In addition, both English- and Spanish-language abbreviated interviews were excluded. Outlier cases were defined as those whose time exceeded two standard deviations from the average time; these cases were identified both within sections and across the entire interview. Outliers, overall and by section, were excluded from analysis, so a case may have been identified as an outlier for a particular section but not for the total interview. Consequently, individual section times may not sum to the total interview time, and the number of cases for each section may not sum to the total. The number of cases per section varies also because sections F (Future Plans) and G (Locating) were applicable only to the B&B cohort.

The average interview time was calculated by adding each respondent's total interview completion time and dividing it by the total number of respondents. On average, respondents spent 26 minutes completing the NPSAS:08 student interview (table 37). Web interviews were significantly shorter (25 minutes) than telephone interviews (30 minutes), $t(36{,}000) = 68.92$, $p < .0001$. The amount of time spent both on-screen and in transit was significantly different, depending on the mode. Web interviews had a shorter average on-screen time than telephone interviews (22 minutes and 27 minutes, respectively), $t(34{,}000) = 93.86$, $p < .0001$. Conversely, telephone interviews had a shorter average transit time than web interviews (2.2 minutes and 3.4 minutes, respectively), $t(36{,}000) = 78.26$, $p < .0001$.

Table 37. Average interview time, by mode of administration: 2008

Mode of administration	Number of cases	Average on-screen time	Average transit time	Average completion time
Total	63,770	23.0	3.1	26.1
Web interview	47,280	21.6	3.4	25.0
Telephone interview	16,500	27.2	2.2	29.5

NOTE: Detail may not sum to totals because of rounding. All times are presented in minutes. Interview times are based on fully completed web and telephone interview cases only. Spanish- and English-language abbreviated interviews and partial cases were excluded from analysis. Also excluded from this analysis were *outlier* cases. An *outlier* was defined as any case whose completion time exceeded two standard deviations above or below the average time for a given section or for the total interview.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

The most time-consuming section to complete was the enrollment section, taking 7.4 minutes, on average, for all respondents (table 38). This outcome was expected because this section, critical for routing the respondent through the rest of the interview, was the lengthiest and most complex: It included items eliciting student status and eligibility, as well as enrollment information on additional schools attended during the 2007–08 academic year. Because the routing and question wording for the remainder of the interview were based on the responses to this section, the level of detail in the enrollment section was necessary.

Table 38. Average time to complete the NPSAS student interview, by interview section and completion mode: 2008

Interview section	All respondents		Web		Telephone	
	Number of cases	Average time	Number of cases	Average time	Number of cases	Average time
Total interview	63,770	26.1	47,280	25.0	16,500	29.5
Section A—Enrollment	63,480	7.4	47,240	7.1	16,250	8.5
Section B—Education Experiences	64,060	2.7	17,580	2.5	16,480	3.1
Section C—Financial Aid	63,620	4.2	47,100	4.0	16,520	4.9
Section D—Current Employment	63,880	5.6	47,490	5.2	16,400	6.7
Section E—Background	64,780	4.1	48,490	3.8	16,590	5.0
Section F—Future Plans	9,700	1.0	8,430	0.9	1,270	1.2
Section G—Locating	8,110	3.0	7,330	2.9	790	3.5

NOTE: All times are presented in minutes. Interview times are based on fully completed web and telephone interviews only. Spanish- and English-language abbreviated interviews and partial cases were excluded from analysis. Also excluded from this analysis were *outlier* cases. An *outlier* was defined as any case whose completion time exceeded two standard deviations above or below the average time for a given section or for the total interview. Outliers were identified separately for each section and for the total interview; therefore, individual section times do not sum to the total interview time.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

The education experiences section required about 2.7 minutes, on average, for all respondents to complete. Many items in this section pertained to subgroups of respondents, however. Items that applied to only undergraduate respondents included distance education, remedial classes, high school coursework, and frequency of changes in college major. Other items in the section were administered to only B&B eligible respondents (e.g., items such as transfer credits, studying abroad, and foreign languages). B&B eligible respondents who were undergraduate students spent an average of 3.2 minutes completing this section, whereas non–B&B eligible undergraduate respondents spent an average of 2.9 minutes (table 39). Graduate students who were B&B eligible respondents spent approximately 1.6 minutes completing this section, whereas non–B&B eligible graduate student respondents spent approximately 1.3 minutes.

Table 39. Average time to complete the education experiences section, by student type: 2008

Eligibility and student status	Number of cases	Section time
B&B eligible undergraduate student	11,300	3.2
Non–B&B eligible undergraduate student	43,110	2.9
B&B eligible graduate student	310	1.6
Non–B&B eligible graduate student	9,330	1.3

NOTE: All times are presented in minutes. Interview times are based on fully completed web and computer-assisted telephone interviewing cases only. Spanish- and English-language abbreviated interviews and partial cases were excluded from analysis. Also excluded from this analysis were *outlier* cases. An *outlier* was defined as any case whose completion time exceeded two standard deviations above or below the average time for a given section or for the total interview. B&B = Baccalaureate and Beyond.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

Overall, respondents spent, on average, 4.2 minutes to complete the section about financial aid. Items in this section focused on the application for federal student aid, types and amounts of aid received, graduate fellowships and assistantships, work-study programs, and a new set of items concerning knowledge about the Academic Competitiveness Grant and the national Science and Mathematics Access to Retain Talent Grant.

Completion of the current employment section took approximately 5.6 minutes. This section pertained to employment during enrollment in the study year and contained items about personal finances. Included were items related to the effects of employment on education, affordability of education without employment, reasons for working, spousal income, receipt of federal assistance, assets, and credit card burden. Table 40 shows the differences in section time by job status. Employed respondents spent an average of 6.5 minutes to complete this section, whereas respondents who were not employed spent 3.3 minutes, on average, to complete this section.

Table 40. Average time to complete the current employment section, by employment status: 2008

Job status	Number of cases	Section time
Employed	45,260	6.5
Not employed	18,630	3.3

NOTE: All times are presented in minutes. Interview times are based on fully completed web and computer-assisted telephone interviewing cases only. Spanish- and English-language abbreviated interviews and partial cases were excluded from analysis. Also excluded from this analysis were *outlier* cases. An *outlier* was defined as any case whose completion time exceeded two standard deviations above or below the average time for a given section or for the total interview.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

The background section collected demographic information about respondents and their families and required an average of 4.1 minutes to complete. It also contained items related to number of dependents, parent education, disability status, citizenship status, and community service work. Most items in this section applied either to all respondents or to subgroups based on age and student status.

The final two sections of the NPSAS:08 student interview, future plans and locating, applied only to B&B eligible respondents. The future plans section focused on postgraduation plans regarding employment, education, and teaching. This section required approximately 1.0 minute to complete. The locating section collected contact information for the follow-up B&B planned for 2009. This section required, on average, 3.0 minutes to complete. Contact information—such as parents' addresses, permanent addresses, and e-mail addresses—was collected in this section.

When interview completion time is examined, internet connection type and how type affects transit time should be considered. Respondents were asked to complete a small set of optional debriefing questions after the interview. Approximately 91 percent (n = 42,750) of respondents who answered a debriefing question concerning internet connection type indicated that they used a fast connection to complete the NPSAS:08 student interview (table 41). As expected, respondents using dial-up modems experienced much longer transit times (11.0 minutes) than those with a fast connection (3.2 minutes) $t(858) = -26.24$, $p < .0001$.

Table 41. Average time to complete the web student interview, by web connection type: 2008

Mode of administration	Number of cases	Average on-screen time	Average transit time	Average completion time
Total	46,960	21.6	3.4	25.0
Dial-up modem	840	22.9	11.0	33.9
Fast connection	42,750	21.5	3.2	24.6
Other connection type	650	24.3	4.7	29.0
Do not know connection type	2,720	22.6	3.6	26.2

NOTE: All times are presented in minutes. Detail may not sum to totals because of rounding. Interview times are based on fully completed web interviewing cases only. Spanish- and English-language abbreviated interviews and partial cases were excluded from analysis. Also excluded from this analysis were *outlier* cases. An *outlier* was defined as any case whose completion time exceeded two standard deviations above or below the average time for a given section or for the total interview. Outliers were identified separately for each section and for the total interview; therefore, individual section times do not sum to the total interview time.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

4.3.2 Call Counts

The outbound calling portion of NPSAS:08 was a considerable task. Approximately 500 interviewers were trained to make outbound calls to locate and interview sample members. Interviewers worked approximately 81,700 hours to complete a total of 29,440 interviews, which represents 2.78 hours per completed interview. Because the telephone interview required approximately 25 minutes to complete, the balance of the interviewer's time represents locating and contacting activities. All locating of sources and attempts to gain a sample member's interview were exhausted. Beyond these activities, the balance of the interviewer's time was spent on case maintenance, including tasks such as reviewing call histories, scheduling callbacks, providing comments, and updating case statuses.

During the NPSAS:08 data collection process, the number of calls placed to sample members varied according to several factors. Overall, slightly fewer than 1.5 million calls (1,465,500) were placed to NPSAS:08 sample members, for an average of 11 calls per sample member (table 42). NPSAS:08 respondents received an average of only 6 calls, as opposed to the 23 calls placed to nonrespondents $t(49{,}800) = 198.56$, $p < .0001$. Respondents who completed a telephone interview received approximately 4 fewer calls (for a total of 10 calls) than respondents who completed the web interview but required telephone follow-up $t(40{,}200) = 40.76$, $p < .0001$.

Table 42. Call counts, by interview status and mode of completion: 2008

Interview status and completion mode	Number of cases	Number of calls	Average calls per case
Total	132,800	1,465,500	11.0
Not interviewed	37,440	870,460	23.3
Interviewed	95,360	595,040	6.2
By completion mode			
Web interview—no telephone follow-up	44,800	†	†
Web interview—with telephone follow-up	21,120	297,780	14.1
Telephone interview	29,440	297,250	10.1

† Not applicable.
NOTE: Detail may not sum to totals because of rounding. This analysis does not include calls made to the 5,000 sample members determined to be ineligible for the study.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

The number of calls placed to sample members varied significantly according to type of institution, with sample members who attended less-than-2-year institutions receiving approximately 14 calls (table 43), $F(3, 133{,}800) = 530.62$, $p < .0001$, as opposed to the 10, 11, or 13 calls received by members attending the other types of institutions (4-year doctorate-granting, 4-year non-doctorate-granting, 2-year, respectively). Sample members who attended private, for-profit institutions received significantly more calls than sample members who attended either public institutions or private, not-for-profit institutions (13 calls, as opposed to 11 and 10), $F(2, 133{,}800) = 339.43$, $p < .0001$. Number of calls to sample members varied also according to institutional sector, ranging from 14 calls to sample members who attended private, for-profit, less-than-2-year, institutions to an average of 9 calls placed to sample members who attended private, not-for-profit, 4-year, non-doctorate-granting institutions.

Notable differences based on student type also existed. While B&B eligible sample members received approximately 10 calls, all other undergraduate sample members received 12 calls, a statistically significant difference, $t(38{,}700) = -15.32$, $p < .0001$. Significant differences were also found between the number of calls placed to undergraduates (11 calls, on average) and the number placed to graduate students (8 calls, on average), $t(19{,}900) = 31.73$, $p < .0001$.

Answering machines (including voice mail) provide both a way for study personnel to maintain contact with sample members when sample members are unable to answer the phone and a way for sample members to screen their incoming calls. Excluding sample members who completed interviews during the early response period, 56 percent of sample members who had at least one answering machine event completed the interview, compared with the 60 percent of sample members without an answering machine event who completed the interview ($z = -9.13$, $p < .01$). In addition to interview completions, call counts varied in relation to the number of answering machine events. Sample members who had at least one answering machine event were called approximately 20 times, compared with the 8 calls placed to sample members without an answering machine event, $t(30{,}000) = -107.01$, $p < .000$.

Table 43. Call counts, by institutional characteristic and student type: 2008

Institutional characteristic and student type	Total cases	Total calls to sample cases	Calls per case
Total	132,800	1,465,500	11.0
Institutional level			
Less-than-2-year	7,950	111,270	14.0
2-year	40,770	518,310	12.7
4-year non-doctorate-granting	37,140	389,980	10.5
4-year doctorate-granting	46,940	445,940	9.5
Institutional control			
Public	84,240	940,210	11.2
Private not-for-profit	31,950	308,270	9.7
Private for-profit	16,610	217,020	13.1
Type of institution			
Public less-than-2-year	1,480	20,420	13.8
Public 2-year	37,010	468,560	12.7
Public 4-year non-doctorate-granting	15,850	163,600	10.3
Public 4-year doctorate-granting	29,910	287,630	9.6
Private not-for-profit less-than-4-year	1,790	23,780	13.3
Private not-for-profit 4-year non-doctorate-granting	13,930	134,210	9.6
Private not-for-profit 4-year doctorate-granting	16,230	150,280	9.3
Private for-profit less-than-2-year	6,050	85,650	14.2
Private for-profit 2-year-or-more	10,560	131,370	12.5
Student type			
Total undergraduates	118,030	1,348,020	11.4
Potential B&B student	24,290	247,950	10.2
Other undergraduate	93,740	1,100,070	11.7
Graduate/first-professional	14,770	117,480	8.0

NOTE: Detail may not sum to totals because of rounding. This analysis does not include calls made to the 5,000 sample members determined to be ineligible for the study. B&B = Baccalaureate and Beyond.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

4.3.3 Online Coding

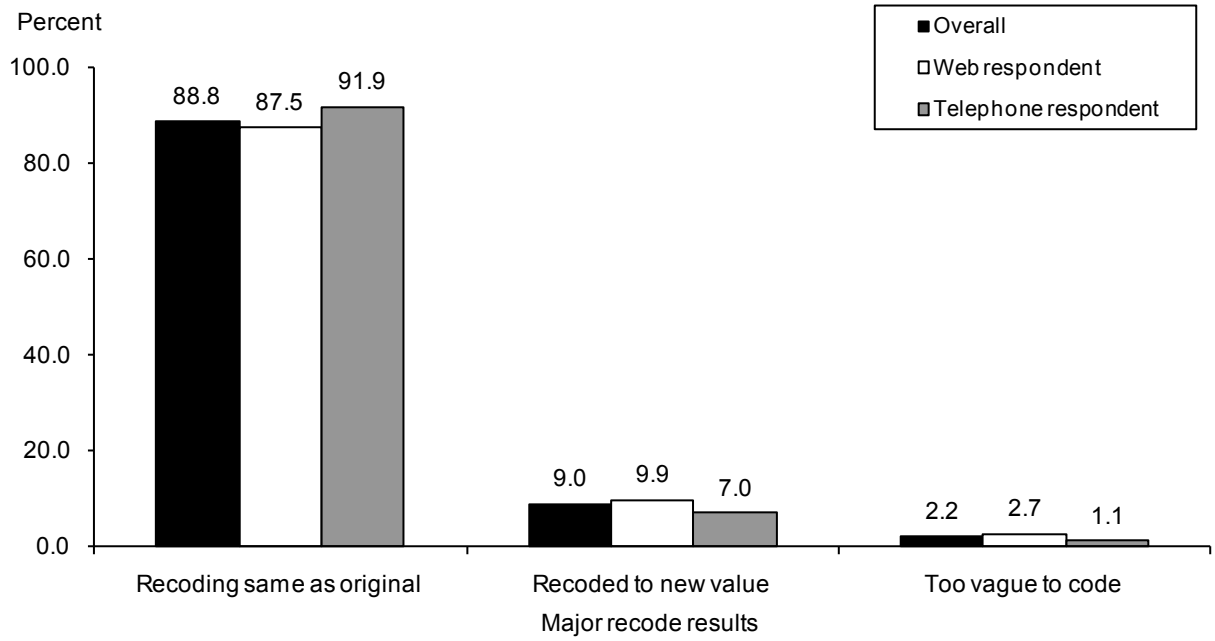
The online coding systems were used to categorize institutions attended and major fields of study into standardized categories. The major coding system used an assisted coder that derived a list of possible matches from the text string the respondent provided. If no areas matched, double dropdown boxes were used to categorize the major. The institutional names and enrollment history were collected in the student interview in a "looping" series of questions, which repeated for each school attended. As a result, students were able to report enrollment dates and intensity (e.g., full-time or part-time enrollment) for multiple postsecondary institutions. Section 2.3.2 includes a detailed description of the coders included in NPSAS:08.

As part of the data quality review of major coding, 10 percent of cases that were successfully coded during the interview were randomly selected for review after interview completion. Expert coders reviewed the selected cases by determining whether, based on the text string provided, the resulting code was appropriate. For the institutional coder, a recode review was not possible because the resulting code was based solely on the information (e.g. school name, city, state) provided and

there was no other source to confirm whether the correct institution had been selected. However, any case for which an institution was not successfully coded during the interview, went through a process called upcoding in which expert coders attempted to find the matching institution based on the information provided. Likewise, any majors for which a text string was provided but no code was selected were also upcoded.

Recoding. Major or field of study was the only coder that was recoded. Of all strings reviewed, 89 percent were coded correctly, 9 percent were recoded to a new value, and 2 percent were too vague to code (figure 14). Majors were coded more reliably by telephone interviewers (92 percent) than by web respondents (88 percent; $z = 5.0$, $p < .01$). This difference was likely due to the training and experience that the interviewers received with the coder.

Figure 14. Recode results for major field of study, by completion mode: 2008

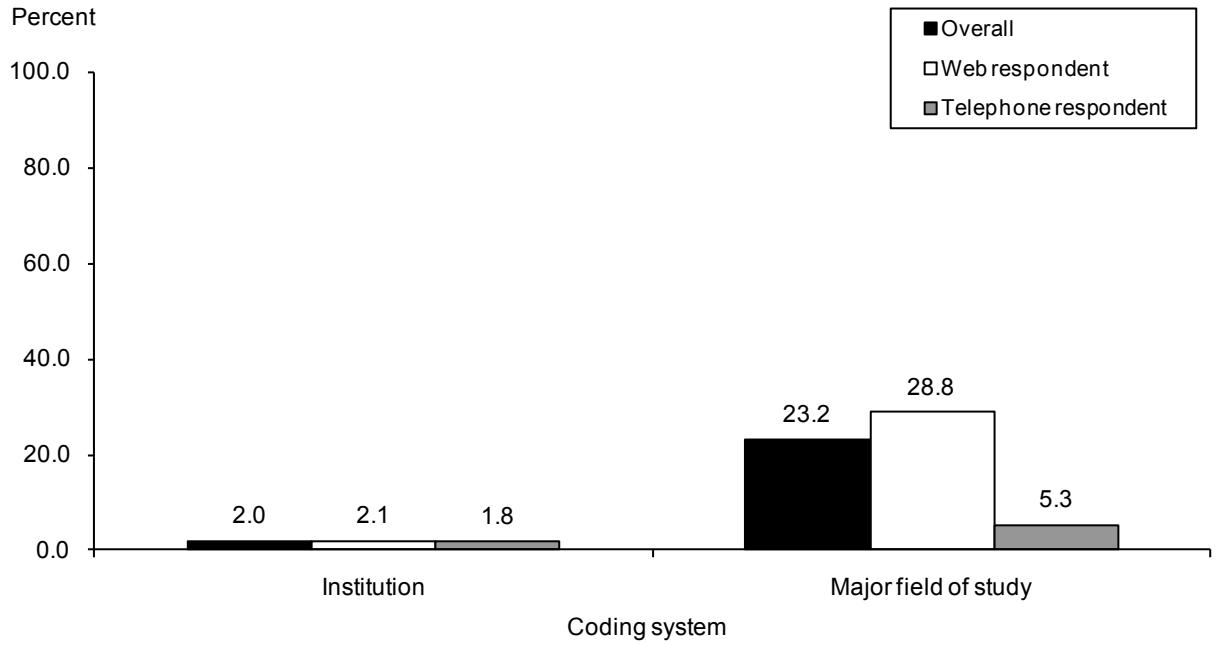


SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

Upcoding. In addition to the evaluation of interview coding accuracy, evaluation of the uncodable text strings was made for each coder. When major field of study or institutional codes were missing, expert coders used the text string (if provided) to determine the appropriate major or institutional code (figure 15).

Approximately 2 percent of all institutional codes reported in the interview required upcoding. While the difference between the institutional upcoding rates for web interviews (2 percent) and telephone interviews (2 percent) was very small, the difference was statistically significant ($z = -2.98$, $p < .01$). Majors had higher rates of upcoding than institutions. Overall, 23 percent of majors required upcoding. Web respondents were more likely to provide a major that required upcoding (29 percent) than telephone respondents (5 percent; $z = -21.40$, $p < .01$), likely because of the interviewers' training and experience.

Figure 15. Upcode results, by coding system and completion mode: 2008



SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

4.3.4 Help Text

The NPSAS:08 interview featured both general and screen-specific help text on all instrument screens. The general help text outlined basic information on internet browsers and response types (i.e., how to use a check box, dropdown, or radio button). The screen-specific help text defined instrument vocabulary, instructed sample members how to enter responses, and explained the kind of information requested for each form. Each help text screen also provided a toll-free telephone number to the NPSAS:08 help desk for further assistance.

The instrument contained counters that calculated the number of times that help text for each screen was accessed. The screen-level rate of help text access was analyzed overall and by mode of interview administration to identify screens that may have been problematic for users. Across both administration types, 1.0 percent of screen visits, on average, involved access to help text. Interviewers accessed help text more often than respondents (3 percent, as compared with less than 1 percent; $\chi = 15.77$, $p < .01$). Interviewers had been trained and encouraged to use help text.

Table 44 shows interview items where help text was accessed 5 percent or more of the times that the form was administered. Results were also examined by mode with the expectation that telephone respondents would encounter higher rates of help text usage because this feature of the instrument was emphasized during the telephone interviewer training. The form N8ASSOC had the highest rate of help text access: 15 percent. This form constituted follow-up questioning of students who reported they were working toward an associate's degree; it was to differentiate between Associate of Arts and Associate of Science degrees. Interviewers were more likely than web respondents to access help text for this form, with 40 percent of interviewers' visits to the form involving help text access, compared with only 1 percent of web respondents' visits ($\chi = 68.53$, $p < .01$).

Table 44. Interview items with high rates of help text access, by mode of administration: 2008

Item	Variable label	Total administered to	Total percent	Mode of administration: Web Percent	Mode of administration: Telephone Percent
N8ASSOC	Type of associate's degree	17,490	14.7	1.3	39.8
N8CKHOUR	Clock hour requirement	6,320	8.2	1.0	16.9
N8DRPCMP	Completed previous course/term	220	9.7	3.2	18.5
N8EARNS	Number of years/terms/months/weeks worked	32,800	11.9	0.7	49.0
N8HSCMP	Completing high school requirements for entire time enrolled at NPSAS	550	6.6	0.9	16.5
N8SCHJOB	School related job: work-study job	65,390	5.1	0.6	17.8
N8TAXTYP	Type of untaxed benefit received	6,880	7.2	0.3	17.0
N8WAERNS	School-related job: number of years/terms/months/weeks worked	7,010	9.4	0.7	43.7

NOTE: Detail may not sum to totals because of rounding. Rate of help text access applied to the interview screens that were administered to more than 50 respondents and on which help text was accessed at a rate of 5 percent or more.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

The item with the second-highest rate of help text access was N8EARNS, which asked respondents the number of terms, months, or weeks they worked in all kinds of jobs. The overall help text access rate for the form was 12 percent. The help text for this form was used primarily by interviewers (49 percent, compared with 1 percent for web respondents; $z = 114.03$, $p < .01$). N8WAERNS also asked respondents how many terms or months they had worked, but this item focused on school-related jobs (i.e., assistantships, fellowships, traineeship, work-study). Overall, 9 percent of respondents used help text for this form. This form was also mostly accessed by interviewers (44 percent, compared with 1 percent for web respondents; $z = -16.92$, $p < .01$).

N8CKHOUR asked respondents working on a certificate or diploma if their program required at least 3 months or 300 clock hours of instruction. The overall help text access rate for this form was 8 percent. This item's help text was accessed by 17 percent of telephone and 1 percent of web respondents ($z = 22.93$, $p < .01$).

On the form N8TAXTYP, recipients of untaxed benefits were asked to specify the type of benefit received. The overall help text usage rate for this screen was 7 percent. This form's help text was used primarily by telephone interviewers (17 percent, compared with less than 1 percent for web respondents; $z = 26.53$, $p < .01$).

The form N8DRPCMP, which asked those who dropped out during the current term if they completed a course in another term in the 2007–08 school year, had an overall help text usage rate of 10 percent. Telephone interviewers were more likely to use the help text for this form, compared with web respondents (19 percent, compared with 3 percent; $z = 3.74$, $p < .01$).

N8HSCMP and N8SCHJOB were the remaining two items for which help text was accessed at a rate of at least 5 percent. N8HSCMP asked respondents whether current enrollment was solely for completing high school requirements. The help text for this item was accessed primarily by interviewers (17 percent, compared with 1 percent for web respondents; $z = 7.12$, $p < .01$). The other item, N8SCHJOB, asked respondents if they had a work-study job or assistantship. This item was used in 18 percent of telephone interviewers and less than 1 percent of web interviews ($z = 87.89$, $p < .01$).

4.3.5 Item Nonresponse

Item-level nonresponse analysis focuses on the rates at which respondents failed to provide an answer to interview items. Missing data for items in the NPSAS:08 student interview were associated with several factors: (1) a true refusal to answer, (2) an unknown answer, (3) confusion over the question wording or response options, or (4) hesitation to provide a "best guess" response. Item nonresponse rates were based on the number of interview respondents to whom the item was applicable and of whom it was asked.[26] Overall, item-level nonresponse rates were low, with only 23 items out of approximately 500 having more than 10 percent of data missing (table 45).

Table 45. Item nonresponse for interview items with more than 10 percent of data missing, by section: 2008

Interview section and variable name	Variable label	Overall		Web		Telephone	
		Number administered to	Percent missing	Number administered to	Percent missing	Number administered to	Percent missing
Enrollment							
N8EXNMY	Date degree expected at NPSAS school	63,890	13.4	47,910	1.8	15,980	9.4
N8GPAEST	Estimate of cumulative GPA	5,990	11.8	3,290	12.9	2,710	10.5
N8MAJ1	Primary major: string	71,180	16.4	53,090	21.6	18,090	1.0
N8MAJ2	Secondary major: string	3,340	20.8	2,560	25.6	780	5.1
N8MJ2SPE	Secondary major: specific code	3,340	10.3	2,560	11.3	780	6.8
N8MJGPES	Estimate of cumulative GPA in major	6,300	14.3	3,380	14.5	2,920	14.0
N8NPDG	Degree completion in NPSAS year	1,580	25.5	1,360	24.7	220	30.5
Education experiences							
N8CMPTR	Cost of special equipment (computers/microscopes/tools)	82,550	13.6	61,000	17.2	21,550	3.4
N8OMJ1	Original major: string	9,730	15.7	7,660	19.6	2,070	1.2
Financial aid							
N8AMNEMP	Employer aid amount	9,540	38.9	7,680	43.2	1,870	20.9
N8AMNPMP	Aid from parent's employer amount	5,080	70.1	4,320	74.7	760	45.9
N8AMNPRV	Private organization aid amount	6,830	51.7	5,700	55.9	1,130	30.2
N8AMNVET	Veteran's benefits amount	5,470	65.1	4,520	71.6	960	34.6
N8GASSM	Other graduate assistantship amount	660	48.6	580	49.8	80	39.7
N8GFELM	Graduate fellowship amount	990	32.6	870	33.3	120	27.9
N8GLNEST	Amount borrowed for graduate loans: estimate	510	19.1	460	16.9	50	38.5
N8INAMT	College grant or scholarship amount	25,000	14.4	20,810	14.8	4,180	12.8
N8PRVEST	Amount of private or alternative loan: estimate	580	39.7	390	38.7	190	41.9
N8RASSM	Research assistantship amount	1,340	25.7	1,180	26.1	160	22.7
N8STAMT	State grant or scholarship amount	23,250	17.4	18,960	17.6	4,290	16.5
N8TASSM	Graduate teaching assistantship amount	1,390	25.2	1,230	25.6	170	22.2
N8TRNSM	Traineeship amount	390	79.8	340	83.4	50	55.1
N8ULNEST	Amount borrowed for undergraduate loans: estimate	4,780	25.0	4,050	22.8	730	36.9

NOTE: Detail may not sum to totals because of rounding. This table includes only items that were administered to at least 100 respondents. GPA = grade point average.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

[26] Partial interview completions and interview nonrespondents were excluded from this analysis.

Of the 1,580 students who received the item "degree completion in NPSAS year," approximately 26 percent failed to answer. Each of the three questions related to the coder for major field of study was found to have elevated levels of item nonresponse: Approximately 16 percent of respondents failed to enter a text string for the primary-major coder (N8MAJ1); 21 percent of respondents indicating a double major failed to enter a text string for their second majors (N8MAJ2).

Many respondents appeared reluctant to answer items that elicited information possibly deemed sensitive in nature, such as items regarding financial information. Eleven of the items, for example, focused on the specific amount of financial aid the respondent had received. One item was applicable only after the respondent had indicated receipt of a specific type of financial aid. High nonresponse on financial aid items was likely due to difficulty recalling specific amounts of financial aid received. The rates of item nonresponse for these items ranged from 14 percent for "college grant or scholarship amount" to 80 percent for "traineeship amount."

Items requesting estimates of loan amounts also showed high rates of item nonresponse. Each of these loan estimates followed a question asking the respondent to provide the specific amount of loans received. In order to receive the estimate item, the respondent had to leave the specific item blank. Estimate of amount borrowed for graduate loans had an item nonresponse rate of 19 percent, while the estimate of the amount borrowed for undergraduate loans had an item nonresponse rate of 25 percent. Moreover, approximately 40 percent of respondents did not provide an estimate for their amount of private or alternative loans.

Item-level nonresponse was analyzed by mode of interview administration to assess the impact of mode on item nonresponse. Fifteen items had nonresponse rates significantly different by mode. Only three items had higher rates of nonresponse among telephone interviews than among web interviews. For the item "date degree expected at NPSAS school," approximately 9 percent of telephone respondents failed to answer, compared with approximately 2 percent of web respondents ($\chi = 44.07, p < .01$). The remaining two items addressed the respondent's estimate of graduate and undergraduate loans. Telephone respondents were more likely (39 percent) than web respondents (17 percent) to omit an estimate for graduate loan amounts ($\chi = 3.70, p < .01$); for undergraduate loan estimates, the omission rates were 37 percent for telephone respondents and 23 percent for web respondents ($\chi = 8.10, p < .01$).

Twelve items showed significantly higher rates of nonresponse among web respondents than among telephone respondents. Seventeen percent of web respondents failed to answer the item "cost of special equipment," compared with 3 percent of the telephone respondents ($\chi = 50.81$, $p < .01$). Four of these items were related to coding of the respondents' primary or secondary major field of study. For respondents' primary major string, approximately 22 percent of web respondents failed to respond to the item, compared with 1 percent of telephone respondents ($\chi = 2.87, p < .01$). Three other items were found to have significantly higher rates of nonresponse for web interviewees than for telephone respondents: secondary major string ($\chi = 12.35, p < .01$), secondary major specific code ($\chi = 3.63, p < .05$), and original major string ($\chi = 20.42, p < .01$).

4.3.6 Conversion Text

To minimize item-level nonresponse, conversion text was used to encourage a respondent to provide an answer, particularly for critical items. The conversion text is intended to mimic the response conversion strategy attempted by an interviewer when a respondent refuses to answer a question. In the NPSAS:08 instrument, key items were identified to include conversion text. If left blank, these items were displayed again, with additional text emphasizing the importance of the item and sometimes with the addition of a "don't know" option. The "don't know" option was available to respondents only on conversion text items. Overall, conversion rates ranged from 25 percent to 100 percent (table 46). The form N8HSCMP, indicating whether high school requirements were completed, had the highest conversion rate (100 percent). Form N8PARNC, parent's income, had a conversion rate of 87 percent. A valid response was converted for 94 percent of respondents who saw the conversion text on the question asking the respondent's race (N8RAC1).

Table 46. Conversion rates for critical items, by type of response: 2008

Form name	Variable label	Cases viewing conversion text	Percent converted by providing a valid response	Percent converted by providing a "don't know" response	Total percent converted
N8HSCMP[1]	Completing high school requirements	20	100.0	#	100.0
N8INCOM[1]	Earnings in 2007	2,210	51.3	23.5	74.8
N8INCSP[1]	Spouse's earnings in 2007	810	29.5	34.7	64.1
N8NUMJOB	Number of jobs during NPSAS year	630	85.5	†	85.5
N8PARNC[1]	Parent's income in calendar year 2007	5,060	33.2	54.5	87.7
N8PRDG[1]	Earned prior degree/certificates	290	88.4	6.1	94.5
N8RAC1	Race	4,630	94.4	†	94.4
N8GPA[1,2]	Cumulative GPA	3,400	83.5	11.7	95.2
N8MJGNUM[1,2]	Cumulative GPA in major studies	3,880	88.2	9.5	97.8
N8HSGPA[1,2]	High school GPA	4,430	24.0	1.2	25.1
N8PRVAMT[1,2]	Private loans	630	58.1	29.5	87.6
N8UGLN[1,2]	Amount borrowed for undergraduate loans	16,370	22.3	5.9	28.2
N8GRLN[1,2]	Amount borrowed for graduate loans	1,420	29.8	4.6	34.4
N8NENRL[1]	Enrollment during the NPSAS year	760	74.4	15.0	89.4

† Not applicable.
# Rounds to zero.
[1] For these items, a "don't know" response option was added when the screen reloaded, in addition to text emphasizing the importance of the item.
[2] The conversion text for these items involved multiple forms in which the first form requested a specific amount or number and the second form provided the respondent a list of ranged categories.
NOTE: Detail may not sum to totals because of rounding. GPA = grade point average.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

The NPSAS:08 instrument included several items that used a multiform conversion method. For these questions, the initial question asked the respondent to provide a specific amount or number. To convert item nonresponse on these questions, a follow-up question provided categorical ranges to help the respondent classify his or her response. For example, if a sample member declined to provide a specific figure for income, the next item would provide the respondent with income ranges from which to select a response. Among the multiform conversion items, "Cumulative GPA in major studies" had the highest conversion rate at 98 percent, including the

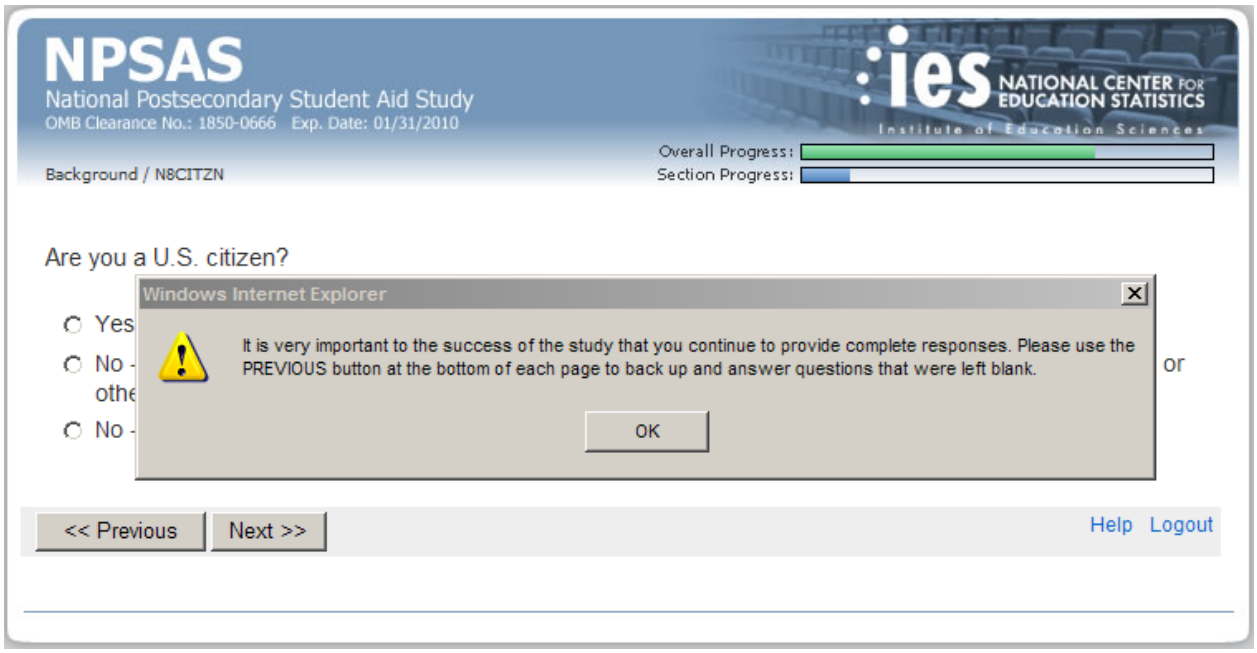
almost 10 percent who indicated "don't know." The lowest conversion rate among the multiform items was "high school GPA" at 25 percent.

4.3.7 Verification Checks

Certain forms within the NPSAS:08 student interview were embedded with verification messages that indicated when the respondent's answer was outside the expected range for that item. Once the validation message displayed, the answer could be corrected or accepted. When respondents were confronted with the warning message, 28 percent of responses were modified.

Throughout the course of the interview, respondents could decline to answer any question that they did not want to answer. Whenever three consecutive forms were left blank, a pop-up message reminded the respondent of the importance of the NPSAS:08 study and of providing complete information (figure 16). Respondents were also reminded that they could return to items previously left blank and answer them. Of the 95,360 respondents, 3,150 (3 percent) triggered this message at least once. Among those for whom the message was triggered, 98 percent provided a valid response on the form concerned.

Figure 16. Pop-up message displayed after three consecutive blank screens: 2008



SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08) website.

4.3.8 Help Desk

Help desk staff members provided live telephone assistance to students with technical questions and other concerns about the web interview. All help desk agents were trained to conduct telephone interviews as necessary. If technical difficulties precluded the sample member's logging in to the web interview, help desk agents encouraged the sample member to proceed with a telephone interview. In addition to taking live calls, help desk agents responded to voice mail messages left by respondents when the help desk was closed.

For project staff to gain a better understanding of the problems encountered by sample members attempting to complete the web interview, a program recorded each help desk incident that occurred during data collection. For each call-in occurrence, the help desk staff confirmed contact information for the sample member and logged the type of problem, a description of the problem and resolution, and the incident status (either pending or resolved). The help desk application documented all incoming calls from sample members. In addition to this primary documentation function, the help desk provided the following: information required to verify a sample member's identity to assist with login difficulties; login information allowing a sample member to access the web interview; and a means for tracking problems that could not be immediately resolved.

The help desk application also provided various reports on the type and frequency of problems experienced by sample members, as well as a way to monitor the resolution status of all help desk inquiries (table 47). Most problems reported by students who called the help desk were requests for study ID numbers or passwords (32 percent). Other problems involved difficulty disabling pop-up blockers (15 percent), website contact requests (19 percent), unavailability of the study website (2 percent), and general questions about the study (6 percent). The remaining incidents (21 percent) were miscellaneous situations that could not be coded into one of the principal categories.

Table 47. Help desk incident type: 2008

Type of incident	Total incidents recorded	Percent of total incidents
Total	2,965	100.0
Study ID/password	948	32.0
Questions about study	185	6.2
Browser settings	20	0.7
Website unavailable	58	2.0
Questionnaire content	30	1.0
Program error	64	2.2
Routing/skip patterns	12	0.4
Pop-up blocker	445	15.0
Update information	22	0.7
Website contact request[1]	555	18.7
Other	626	21.1

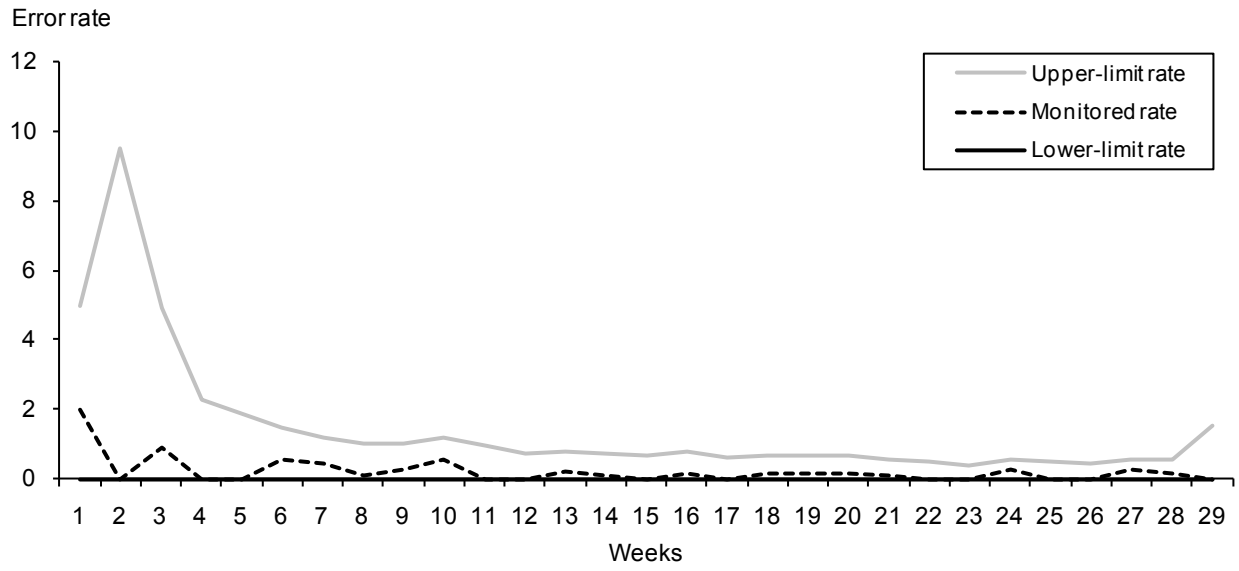
[1] Website contact requests occurred when a sample member sent a request through the NPSAS:08 student website to the help desk for contact. In the event that the request was for a new study ID or password, the incident was coded as a website contact request based on the way the incident was received.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

4.3.9 Telephone Interviews and Quality Assurance Monitoring

To ensure that telephone interviews were accurately administered and that their responses were being accurately coded, telephone interviewers were routinely monitored. The monitors had the ability to observe live interviews (with both visual and audio) without disturbing either the respondent or the interviewer. The monitor would observe up to 20 questions during each interview selected for review. Each question was assessed for two elements of interviewer performance: (1) correct delivery of the question, and (2) accurate keying of the response. To correctly gauge the accuracy and performance of the NPSAS:08 telephone interviewers, monitoring sessions were conducted throughout the data collection period and during all shifts.

Of the 23,042 items monitored, 93 question delivery errors and 31 data entry errors were observed. The overall error rates were low (typically less than 1 percent) and within control limits (figures 17 and 18).[27] The monitoring results were tabulated and posted as weekly and cumulative reports available to project staff. Error rates are typically higher at the beginning of data collection as telephone interviewers become familiar with the NPSAS:08 study and interview. Error rates typically become more stable over the course of the study.

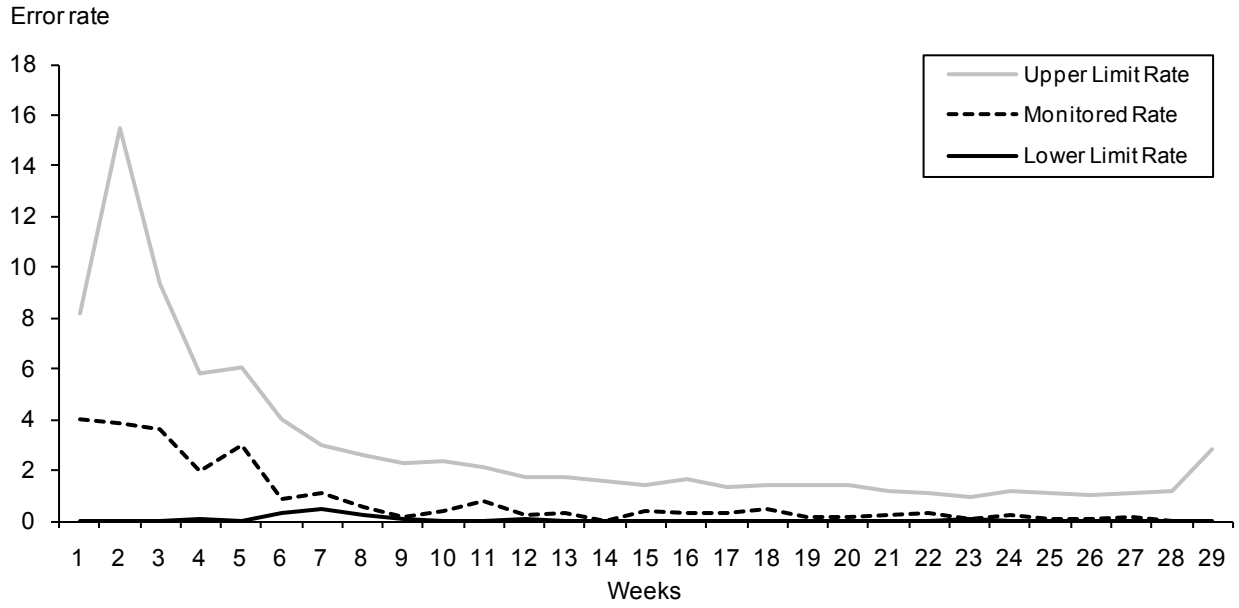
Figure 17. Question delivery error rates, by week: 2008



SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

[27] The upper and lower control limits were defined by 3 times the standard error of the proportion of errors to the number of questions observed for the period (+3 times the standard error for the upper limit; −3 times the standard error for the lower limit). Peaks in control limits can be attributed to variation in the number of observations across weeks.

Figure 18. Data entry error rates, by week: 2008



SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

Quality circle (QC) meetings provided an invaluable opportunity for open communication between call center staff and project staff to address the challenges encountered in data collection, to help ensure data quality, and to provide retrainings as necessary. Meetings were held weekly at the call center, and supervisors selected telephone interviewers to attend on a rotating basis. For interviewing staff unable to attend a given meeting, notes were distributed electronically to the call center supervisory staff, who distributed the notes to their staffs accordingly. Specifically, these meetings addressed elements of the interview and instrument design, addressed interview cooperation tactics, motivated the group in alignment with study goals, and offered exchange of feedback on data collection issues such as the following:

- necessary clarification of interview questions and item responses;
- NPSAS eligibility criteria;
- interviewers' submission of problem sheets;
- the importance of providing detailed case comments;
- help desk operations;
- Virtual Call Center (VCC) staff issues or concerns;
- methods of gaining cooperation from sample members and associated contacts (e.g., parents); and
- morale boosting and reinforcement of positive interviewing techniques.

Across the QC meetings, a variety of issues were addressed that reinforced specific content from training and contributed to prompt problem solving:

Item administration. Clarification of item responses and interpretation of meanings were addressed. Interviewers discussed items that caused confusion, and project staff provided explanations and made minor items and help text modifications when necessary. In addition, items that interviewers found difficult to administer were reviewed for further clarification.

Eligibility criteria. Because of the considerable complexity of the eligibility criteria, interviewers were reminded to allow eligibility determination to be made by the programmed instrument.

Problem sheets. Reporting problems when they occur is crucial to quality telephone interviewing. Interviewers were trained to report problems electronically and to provide detail, including the problem that occurred and the specific point in the interview at which it occurred. Review of problem sheets in QC meetings was a critical means by which staff learned to recognize and manage the different problems they would encounter. Issues that were brought to light through problem sheets were also discussed in QC meetings.

Computer-assisted telephone interviewing comments. During the QC meetings, telephone interviewers were reminded to add detailed comments to every record of their actions within the computer-assisted telephone interviewing case management system (CATI-CMS). These comments helped accurately record the event, as well as prepare the next interviewer who would be calling the sample member.

Help desk operations. Help desk staff who attended the QC meetings received specific information about current help desk operations. In addition, help desk agents were able to discuss current issues about which sample members were currently seeking assistance.

Cooperation from sample members. Discussions focused on the difficulty of gaining a sample member's trust during the initial phases of the call. Refusal-conversion strategies were discussed during QC meetings and adapted as necessary. Interviewers discussed successful approaches for calling refusals and hard-to-reach sample members. They also shared tips for overcoming concerns about individual items and found ways to benefit and learn from each other's experiences.

4.3.10 Virtual Call Center

In addition to a fully staffed call center located in Raleigh, North Carolina, NPSAS:08 employed a VCC. The VCC provided professionally trained interviewers with the tools to complete their work from home instead of in a traditional centralized call center. Interviewers worked on their own computers and connected remotely to the call center systems that would allow them to conduct interviews. A home visit was conducted for each interviewer to ensure that he or she was appropriately prepared, with a quiet, secure work area and sufficient computer equipment. VCC interviewers received the same training and preparation for work as interviewers who worked at the call center. Because all VCC interviewers were located in the local area, they were able to attend project training at the Raleigh call center. They did, however, complete their general training online before project training. To accommodate the VCC interviewers for QC meetings, conference calls were established so that staff could participate from their homes. Additionally, QC notes were posted online, eliminating the need to distribute physical copies to interviewers and allowing VCC interviewers to access the notes from a remote location.

To verify the quality of data being collected from staff using VCC technology, project staff regularly reviewed interview data. Overall, no substantial differences in either data quality or

interviewer productivity were found between interviews administered in the call center and those administered in VCC settings. On average, VCC interviewers took slightly longer to complete interviews than call center interviewers: 29 minutes for call center calls, as opposed to 31 minutes for the VCC, $t(778) = 7.64$, $p < .0001$. The rates of missing data for call center completions and VCC completions were also compared. Overall, approximately 4 percent of all items were found to have significantly different rates of missing data for the call center and VCC, indicating that interviewers used proper interviewing and probing techniques.

4.3.11 Respondent Profiles

Respondent characteristics impact interview experiences in different ways. In order to better understand how certain aspects affect outcomes, such as data collection phase, completion mode, or the completion type a respondent chooses, we examined various respondent characteristics, such as institutional type, age, and student status.

Data collection phases. Respondents completed the NPSAS:08 student interview in three data collection phases: the early response phase, the production-interviewing phase, and the nonresponse-conversion phase. Almost one half of the completed student interviews were obtained during the early response phase (48 percent; table 48). Graduate students were more likely (57 percent) than undergraduate students (47 percent) to complete the NPSAS:08 student interview during the early response phase ($\chi = 15.78$, $p < .01$). Respondents younger than 24 years old were more likely (50 percent) than respondents 30 years or older (44 percent) to complete the interview during the early response phase ($\chi = 11.18$, $p < .01$), while older respondents were more likely (44 percent) than the younger group (40 percent) to complete the interview during the nonresponse-conversion phase ($\chi = 6.6$, $p < .01$). Respondent patterns also differed by institutional type and control. Respondents from less-than-2-year institutions were more likely (68 percent) than those from 4-year, doctorate-granting institutions (34 percent) to complete the interview during the nonresponse-conversion phase ($\chi = 34.27$, $p < .01$). Women were more likely than men to complete the interview during the early response phase (49 percent, compared with 46 percent; $\chi = 5.95$, $p < .01$), while men were more likely than women to complete the interview during the nonresponse-conversion phase (42 percent, compared with 41 percent; $\chi = 2.84$, $p < .01$). Furthermore, respondents who reported that their race was White were more likely to complete the interview during the early response period than respondents who indicated their race was Black (57 percent, compared with 48 percent; $\chi = 13.24$, $p < .01$) but less likely than Asian respondents (57 percent, compared with 61 percent; $\chi = 4.71$, $p < .01$).

Completion type. Toward the end of the data collection period for NPSAS:08, an abbreviated interview was developed as a nonresponse-conversion tool. This interview was much shorter than the full interview, including only section A of the student instrument (for student instrument facsimile, see appendix C). Once data collection was completed, any case that had completed a sufficient portion of the student instrument was deemed a partial interview. Respondents who were older than 30 were more likely to complete an abbreviated interview (13 percent) than respondents younger than 24 (11 percent; $\chi = 2.88$, $p < .01$; table 49). The highest concentrations of abbreviated interviews were in the less-than-2-year institutions (32 percent) and private, for-profit institutions (22 percent).

Table 48. Characteristics of interview respondents, by data collection phase: 2008

Respondent characteristic	Total interviewed	Data collection phase: Early response phase Number	Early response phase Percent of total	Production-interviewing phase Number	Production-interviewing phase Percent of total	Nonresponse-conversion phase Number	Nonresponse-conversion phase Percent of total
Total	95,360	45,870	48.1	10,020	10.5	39,480	41.4
Completion mode							
Web	65,920	44,040	66.8	3,170	4.8	18,710	28.4
Telephone	29,440	1,820	6.2	6,850	23.3	20,770	70.6
Age							
Under 24	53,440	26,940	50.4	5,210	9.7	21,290	39.8
24 through 29	19,580	9,130	46.6	2,040	10.4	8,400	42.9
30 or older	22,350	9,800	43.8	2,770	12.4	9,790	43.8
Student status							
Undergraduate	82,750	38,690	46.8	8,540	10.3	35,530	42.9
Graduate	12,610	7,180	56.9	1,480	11.7	3,960	31.4
Institutional level							
Less-than-2-year	4,530	1,120	24.7	340	7.5	3,070	67.8
2-year	26,610	10,640	40.0	3,190	12.0	12,780	48.0
4-year non-doctorate-granting	27,680	13,870	50.1	2,560	9.3	11,250	40.6
4-year doctorate-granting	36,550	20,240	55.4	3,920	10.7	12,390	33.9
Institutional control							
Public	60,740	29,320	48.3	6,980	11.5	24,440	40.2
Private not-for-profit	24,490	13,300	54.3	2,400	9.8	8,780	35.9
Private for-profit	10,140	3,250	32.0	640	6.3	6,260	61.7
Race[1]							
White	57,550	33,010	57.4	6,700	11.6	17,850	31.0
Black	11,410	5,450	47.7	1,390	12.2	4,570	40.1
Asian	5,850	3,590	61.4	690	11.8	1,570	0.3
American Indian or Alaska Native	800	350	43.7	120	14.9	330	41.4
Native Hawaiian or Other Pacific Islander	770	290	37.8	140	17.5	350	44.7
Other[2]	3,430	1,450	42.2	540	15.6	1,450	42.2
Two or more races	3,120	1,630	52.1	380	12.2	1,110	35.6
Gender							
Male	38,240	17,740	46.4	4,340	11.4	16,160	42.3
Female	57,120	28,120	49.2	5,680	9.9	23,320	40.8

[1] Race was determined from the student interview and required a valid response to be included in the analysis.
[2] The "Other" race category is defined by respondents who selected only "Other Specify" on the Race form in the student interview.
NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

Table 49. Characteristics of interview respondent, by completion type: 2008

Respondent characteristic	Total interviewed	Full interview		Partial interview		Abbreviated interview	
		Number	Percent of total	Number	Percent of total	Number	Percent of total
Total	95,360	83,050	87.1	1,470	1.5	10,850	11.4
Completion mode							
Web interview	65,920	61,360	93.1	880	1.3	3,680	5.6
Telephone interview	29,440	21,690	73.7	580	2.0	7,170	24.3
Age							
Under 24	53,440	46,980	87.9	820	1.5	5,640	10.6
24 through 29	19,580	16,830	86.0	370	1.9	2,380	12.1
30 or older	22,350	19,240	86.1	280	1.2	2,830	12.7
Student status							
Undergraduate student	82,750	71,400	86.3	1,340	1.6	10,020	12.1
Graduate student	12,610	11,650	92.4	130	1.0	830	6.6
Institutional level							
Less-than-2-year	4,530	2,960	65.4	100	2.2	1,470	32.4
2-year	26,610	22,880	86.0	510	1.9	3,220	12.1
4-year non-doctorate-granting	27,680	24,080	87.0	400	1.4	3,200	11.5
4-year doctorate-granting	36,550	33,130	90.6	460	1.3	2,970	8.1
Institutional control							
Public	60,740	53,780	88.5	860	1.4	6,100	10.0
Private not-for-profit	24,490	21,600	88.2	370	1.5	2,520	10.3
Private for-profit	10,140	7,670	75.6	240	2.3	2,240	22.0
Gender							
Male	38,240	33,220	86.9	590	1.5	4,430	11.6
Female	57,120	49,820	87.2	880	1.5	6,420	11.2

NOTE: Detail may not sum to totals because of rounding. Race was not included as a characteristic in this table because only respondents who completed section E (Background) received the race item.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

Completion mode. Most respondents completed the web interview (69 percent; table 50). Slightly more than one half of the respondents from private, for-profit institutions completed the web student interview (51 percent). This was a significantly smaller percentage than the 70 percent of respondents from public institutions who completed the web interview ($z = 27.54$, $p < .01$). Graduate students were also found to have completed the web interview at a higher rate (78 percent) than undergraduate students (68 percent; $z = 0.65$, $p < .01$).

Spanish-language interviews. A Spanish-language abbreviated interview was made available toward the end of the NPSAS:08 student data collection. Approximately 1 percent of all completed interviews were conducted in Spanish (table 51). This interview was made available only through telephone interviews. The highest concentration of Spanish-language interviews came from the private, not-for-profit, less-than-4-year institutions (19 percent) and private, for-profit, less-than-2-year institutions (13 percent).

Table 50. Characteristics of interview respondent, by completion mode: 2008

Respondent characteristic	Total interviewed	Completion mode: Web Number	Completion mode: Web Percent of total	Completion mode: Telephone Number	Completion mode: Telephone Percent of total
Total	95,360	65,920	69.1	29,440	30.9
Age					
Under 24	53,440	38,510	72.1	14,930	27.9
24 through 29	19,580	13,590	69.4	5,990	30.6
30 or older	22,350	13,830	61.9	8,530	38.1
Student status					
Undergraduate	82,750	56,060	67.7	26,690	32.3
Graduate	12,610	9,860	78.1	2,760	21.9
Institutional level					
Less-than-2-year	4,530	1,690	37.4	2,830	62.6
2-year	26,610	16,870	63.4	10,740	40.4
4-year non-doctorate-granting	27,680	19,720	71.2	7,960	28.8
4-year doctorate-granting	36,550	28,640	78.4	7,910	21.6
Institutional control					
Public	60,740	42,370	69.8	18,360	30.2
Private not-for-profit	24,490	18,400	75.1	6,090	24.9
Private for-profit	10,140	5,150	50.8	4,990	49.2
Race[1]					
White	57,550	44,480	77.3	13,070	22.7
Black	11,410	6,870	60.2	4,550	39.8
Asian	5,850	4,870	83.3	980	16.7
American Indian or Alaska Native	800	480	59.6	320	40.4
Native Hawaiian or Other Pacific Islander	770	430	55.1	350	44.9
Other[2]	3,570	2,090	58.4	1,350	37.8
Two or more races	3,120	2,110	67.7	1,010	32.3
Gender					
Male	38,240	26,250	68.6	11,990	31.4
Female	57,120	39,660	69.4	14,750	25.8

[1] Race was determined from the student interview and required a valid response to be included in the analysis.
[2] The “Other” race category is defined by respondents who selected only “Other Specify” on the Race form in the student interview.
NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

Table 51. Characteristics of interview respondent, by language of interview administration 2008

Institutional characteristic and student type	Total interviewed	Language of interview administration			
		English		Spanish	
		Number	Percent	Number	Percent
Total	95,360	94,020	98.6	1,350	1.4
Institutional level					
Less-than-2-year	4,530	3,920	86.6	610	13.4
2-year	26,610	26,320	98.9	290	1.1
4-year non-doctorate-granting	27,680	27,330	98.7	350	1.3
4-year doctorate-granting	36,550	36,450	99.7	100	0.3
Institutional control					
Public	60,740	60,400	99.4	340	0.6
Private not-for-profit	24,490	24,010	98.0	480	2.0
Private for-profit	10,140	9,610	94.8	530	5.2
Type of institution					
Public less-than-2-year	900	880	97.1	30	2.9
Public 2-year	24,380	24,210	99.3	170	0.7
Public 4-year non-doctorate-granting	12,040	11,930	99.1	110	0.9
Public 4-year doctorate-granting	23,410	23,380	99.8	40	0.2
Private not-for-profit less-than-4-year	1,120	910	80.8	220	19.2
Private not-for-profit 4-year non-doctorate-granting	10,810	10,600	98.1	200	1.9
Private not-for-profit 4-year doctorate-granting	12,560	12,500	99.5	60	0.5
Private for-profit less-than-2-year	3,370	2,920	86.7	450	13.3
Private for-profit 2-year-or-more	6,770	6,690	98.8	80	1.2
Student type					
Total undergraduate	81,880	81,880	98.4	1,310	1.6
Potential B&B student[1]	17,900	17,900	99.4	100	0.6
Other undergraduate	63,980	63,980	98.1	1,210	1.9
Graduate/first-professional	12,140	12,140	99.7	30	0.3

[1] Approximately 10 students completed the interview questions needed to determine B&B eligibility but did not meet the minimum requirements to be classified as an interview respondent.
NOTE: Detail may not sum to totals because of rounding. Race was not included as a characteristic in this table because only the abbreviated interview was translated into Spanish. The abbreviated interview included only section A of the student interview and the race item was included in section E. B&B = Baccalaureate and Beyond.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

4.4 Identification of the Baccalaureate and Beyond Cohort

As noted earlier, NPSAS:08 serves as the base year of the B&B. In prior NPSAS studies that derived a B&B cohort, lists of potential baccalaureate recipients were collected with the student list of all enrolled undergraduates and graduate/first-professional students. Unfortunately, these baccalaureate lists often could not be provided until late in the spring or in the summer, after baccalaureate recipients could be positively identified, which negatively affected the data collection schedule. To help facilitate earlier receipt of lists, we requested that the enrollment lists for 4-year institutions include an indicator (B&B flag) of students who had received or would potentially

receive a baccalaureate degree during the NPSAS year (between July 1, 2007, and June 30, 2008).[28] Institutions were instructed to make this identification before spring graduation so as not to delay providing the lists because of this requirement. Four-year institutions were also asked to include an indicator of class level for undergraduates (first year, second year, third year, fourth year, or fifth year). From NPSAS:2000, it was estimated that about 55 percent of the fourth- and fifth-year students would be baccalaureate recipients during the NPSAS year and that about 7 percent of third-year students would also be baccalaureate recipients. This class-level indicator was used when the B&B flag was not provided for any students. These two indicators were used instead of requesting a separate baccalaureate recipient list.

Because most enrollment lists were received before June 30, and many were received before April, some sample students identified by the institution as baccalaureate candidates were determined during the interview not to be baccalaureate recipients (false positives). Likewise, some sample students not identified by the institution as baccalaureate candidates were determined during the interview to have actually received baccalaureate degrees (false negatives) during the specified timeframe.

Table 52 shows that of the 22,880 students who were sampled as potential baccalaureate recipients, 6,610 were not baccalaureate recipients, which is a false-positive rate of about 29 percent.[29] Conversely, of the 62,710 students who were sampled as other undergraduates, about 1,690 were baccalaureate recipients, which is a false-negative rate of about 3 percent. Also, of the 9,830 students who were sampled as graduate/first-professional students, about 50 were determined to be baccalaureate recipients in 2007–08, which is a false-negative rate of less than 1 percent. Overall, the false-negative rate was about 2 percent.

Table 52. Baccalaureate determination, by student type: 2008

Student type	Students interviewed[1]	Confirmed B&B eligibility	
		Number	Percent
Total sample	95,420	18,010	18.9
Total undergraduate	85,590	17,950	21.0
Potential B&B student	22,880	16,270	71.1
Other undergraduate	62,710	1,690	2.7
Graduate/first-professional	9,830	50	0.5

[1] Includes all eligible sample members who completed the eligibility section of the student interview because confirmation of baccalaureate receipt status required contact with the sample members.
NOTE: Detail may not sum to totals because of rounding. B&B = Baccalaureate and Beyond.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

[28] The B&B flag had values of "yes," "no," and "don't know."

[29] Eligibility for the B&B study could be determined without completing the interview. In order to confirm B&B eligibility, the respondent had to have completed at least section A of the student interview.

This page intentionally left blank.

Chapter 5. Variable Construction and File Development

The data files for the 2007–08 National Postsecondary Student Aid Study (NPSAS:08) contain student-level and institution-level data collected from institutional records, student interviews, governmental databases, and administrative databases. These files are fully documented and are available as a set of restricted research files by an electronic codebook (ECB) and as a public release Data Analysis System (DAS) that also comes fully documented. This chapter describes each file and details the editing and documentation process.

5.1 Overview of the NPSAS:08 Data Files

The primary analysis file, from which the study DASs were constructed, contains data for approximately 127,700 study respondents. The primary analysis file contains more than 600 variables, developed from multiple sources (for information on the completeness of data available for study respondents, see table 29). Throughout the data collection period, data were processed and examined for quality. Editing of student data began shortly after the start of web-interview data collection, when procedures and programs for this purpose were first developed. Similarly, editing of the institutional record data began shortly after computer-assisted data entry (CADE) data collection was initiated. Anomalous values were investigated and resolved, where appropriate, with the use of data corrections and logical recodes. Throughout data collection, interim files were delivered to the National Center for Education Statistics (NCES) for review.

After completion of all study data collection, separate DAS files were created for undergraduate and graduate/first-professional students. The first DASs, both undergraduate and graduate/first-professional, were adjudicated and approved for public release in April 2009. The ECB and DAS are both fully documented software products available from NCES. The DAS is available online at http://nces.ed.gov/das/.

Complete data for NPSAS:08 are located in the restricted-access files and are documented by the ECB. The restricted files and the ECB are available to researchers who have applied for and received authorization from NCES to access the restricted data use file. Researchers may obtain authorization by contacting the NCES Data Security Office. The restricted-use NPSAS:08 ECB contains the following files:[30]

- NPSAS analysis file. Contains analytic variables derived from all NPSAS:08 data sources, as well as selected direct student interview variables.
- Student base data file. Contains data collected from institutional records and the student interviews of the study respondents.
- Student interview school data file. Contains institutional data obtained from the student interviews for all study respondents. (A student can have more than one record in the file; a separate record exists for each student for each postsecondary institution the student attended during the study year, for a maximum of five institutions.)

[30] The SAT data files were used only for the creation of derived variables in combination with similar data from other sources and are not included in raw form on the ECB.

- Institutional file. Contains selected institution-level variables for the sampled institutions and can be linked to the Student base data file by the Integrated Postsecondary Education Data System (IPEDS) UNITID number.
- CPS 2007–08 data file. Contains data received from the Central Processing System (CPS) for the study respondents who matched to the 2007–08 financial aid application files.
- CPS 2008–09 data file. Contains data received from CPS for the study respondents who matched to the 2008–09 financial aid application files.
- NSLDS federal Pell Grant data file. Contains grant-level data received from the National Student Loan Data System (NSLDS) for the study respondents who received Pell Grants during the NPSAS year or prior years. (This is a history file, with separate records for each transaction.)
- NSLDS federal ACG/SMART Grant data file. Contains grant-level data received from NSLDS for the study respondents who received Academic Competitiveness Grants (ACGs) or National Science and Mathematics Access to Retain Talent (SMART) Grants during the NPSAS year or prior years. (This is a history file, with separate records for each transaction.)
- NSLDS federal loans data file. Contains loan-level data received from NSLDS for the study respondents who received loans during the NPSAS year or prior years. (This is a history file, with separate records for each transaction.)
- ACT data file. Contains ACT data received for the study respondents who took the ACT exam during academic years 2001–02 through 2006–07.
- Weights file. Contains all the sampling and analysis weights created for NPSAS:08 (contains a separate record for each study respondent).
- Weight history file. Contains all intermediate weight adjustment factors, as well as the final institutional and student weights created for NPSAS:08 (contains a separate record for each study respondent).

5.2 Online Coding and Verification

5.2.1 Online Coding

The web instrument included an online coding system used for the collection of data on students' major fields of study. It also included a coding module used to obtain information for all postsecondary institutions that the student attended during the NPSAS year in addition to the institution from which he or she was sampled (i.e., all institutions attended besides the NPSAS institution required coding). These online coding systems greatly reduced the project staff's coding efforts and the amount of file merging necessary after data collection. They provide the data file user with useful, familiar codes for analysis while ensuring that most codes were assigned during data collection rather than during the data editing phase.

The student's major or field of study was first entered as a text string; then an automated assisted coder was used. According to the text string entered, a list of specific major fields of study was displayed. Respondents chose one of these major categories, or they chose "none of the above," at which point they were presented with a set of general and specific dropdown lists of major

categories. If respondents decided not to provide a text string at the outset, they were sent to the dropdown list.

All postsecondary institutions in which the student had been enrolled between July 1, 2007, and the time of the interview (through June 2008) were selected from a list, in accordance with the respondent's report or the interviewer's entry of one or more of the following: institutional name, city, and state. After this selection, the official name of the institution, as well as selected IPEDS variables (institutional level and control), were inserted into the database.

5.2.2 Range and Consistency Checks

NPSAS:08 included two major web-based data collection systems: student record abstraction and the student interview. Both systems included edit checks to ensure that data collected were within valid ranges. To the extent feasible, both systems incorporated cross-item consistency edits. Although more extensive consistency checks would have been technically possible, use of such edits was limited to prevent excessive respondent burden. Examples of some of the general online edit checks include the following:

- Range checks were applied to all numerical entries so that only valid numeric responses could be entered.
- Consistency checks were used for cross-item comparisons. For example, if a respondent indicated that he or she was 23 years of age (as of January 1, 2008) but graduated from high school in 1988, the respondent was asked to verify this information.
- Enrollment dates were checked to verify that they were within the study period (July 1, 2007, through June 30, 2008).

Data collected from CPS were preloaded into CADE for data checking and to reduce burden on the user. Examples of these CPS items are date of birth and citizenship status.

5.3 Data Editing

The NPSAS:08 data were edited with the use of procedures developed and implemented for previous studies sponsored by NCES, including NPSAS:04. After data collection, the information collected in both CADE and the student instrument was subjected to various quality control checks and examinations. These checks were conducted to confirm that the collected data reflected appropriate item routing (*skip patterns*). Another evaluation involved examination of all variables with missing data and substitution of specific values to indicate the reason for the missing data. For example, an item may not have been applicable to particular students, a respondent may not have known the answer to the question, or a respondent may have skipped the item entirely (table 53).

Table 53. Description of codes for missing data: 2008

Missing data code	Description
−1	Don't know (student interview)
−1	Data not available (computer-assisted data entry)
−3	Not applicable
−6	Value out of range
−8	Item was not reached due to an error
−9	Data missing, reason unknown

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

Skip-pattern relationships in the database were examined by methodically cross-tabulating gate items and their associated nested items. In many instances, gate-nest relationships spanned multiple levels within the instrument: Items nested within a gate question may themselves have been gate items for additional items. Consequently, validating the gate-nest relationships often required several iterations and many multiway cross-tabulations to ensure the proper data were captured.

The data cleaning and editing process for the NPSAS:08 student interview and CADE data files involved a multistage process that consisted of the following:

1. Blank or missing data were replaced with −9 for all variables in the instrument database. There as one exception however – for student records submitted via data-CADE. In cases where a student record was missing from the financial aid file, it was assumed that the student was not a financial aid recipient and the aid amounts were logically recoded to $0 dollars. A one-way frequency distribution of every variable was reviewed to confirm that no missing or blank values remained. These same one-way frequencies revealed any out-of-range, or outlier, values, which were examined to make sure they fell into a reasonable range (e.g., hourly wages of $0.10, rather than $10.00). Creating SAS formats from expected values and the associated value labels also revealed any categorical outliers (SAS Institute Inc. 2006). Descriptive statistics were produced for all continuous variables. All values that were less than zero were temporarily recoded to *missing,* and the minimum, median, maximum, and mean values were examined to assess reasonableness of responses; anomalous data patterns were investigated and corrected as necessary.
2. Legitimate skips were identified with use of instrument source code and flowcharts. Gate-nest relationships were defined to replace −9 codes (*data missing, reason unknown*) with −3 codes (*not applicable*) as appropriate. Two-way cross-tabulations between each gate-nest combination were evaluated; high numbers of nonreplaced −9 codes were investigated to ensure skip-pattern integrity. Nested values were further checked to reveal instances in which the legitimate skip code overwrote valid data, which typically occurred if a respondent answered a gate question and the appropriate nested items but then reverted to change the value of the gate to one that opened on an alternate path of nested items. Because responses to the first nested items remained in the database, they required editing.
3. Variables were formatted (e.g., dates were formatted as YYYYMM), and time units were standardized for items that collected amounts of time in multiple units. In addition, any new codes assigned by expert coders for institutions and for majors from the student interview (including those institutions and majors unable to be coded during the interview) were merged back with the interview data files. At this stage, logical recodes were performed when the value of missing items could be determined from answers to previous questions or preloaded values. For example, if students said they had no children, then the number of children was coded to zero rather than to −3 or −9.
4. One-way frequency distributions for all categorical variables and descriptive statistics for all continuous variables were examined. Out-of-range, or outlier, values were replaced with the value −6 (i.e., out-of-range data).
5. One-way frequencies on all categorical variables were regenerated and examined. Variables with high counts of −9 values were investigated. Because web respondents

could skip items without providing an answer, −9 did remain a valid value, especially for sensitive items, such as those asking for income information.

Concurrently with data cleaning, documentation was developed to detail question text, response options, logical recoding, and the "applies to" text for each delivered variable (for documentation information, see the student instrument facsimile and CADE facsimile in appendix C).

5.4 Data Perturbation

To protect the confidentiality of information about specific individuals, NPSAS:08 data were subject to perturbation procedures to minimize disclosure risk. Perturbation procedures, which have been reviewed and approved by the NCES Disclosure Review Board, preserve central tendency estimates but may result in slight increases in nonsampling errors.

In a study like NPSAS, there are multiple sources of data for some variables (CPS, CADE, student interview, etc.), and reporting differences can occur in each. Data swapping and other forms of perturbation, implemented to protect respondent confidentiality, can lead to inconsistencies as well.

5.5 Statistical Imputations

All variables with missing data were imputed, in accordance with mass imputation procedures described by Krotki et al. (2005). The imputation procedures employed a three-step process. In the first step, missing values were logically imputed. In the second step, the criteria used to match variables into imputation classes in order to stratify the dataset were identified so that all imputations would be processed independently within each class. The third step implemented the weighted sequential hot-deck process, by which missing data were replaced with valid data from donor records that matched the recipients according to the criteria for matching.

Depending on patterns of missing data, some variables that were related substantively and required imputation were grouped into blocks, and the variables within a block were imputed simultaneously (vector imputation). Multiple-response items were imputed simultaneously, wherever possible; however, most of the variables were imputed sequentially. Basic demographic variables were imputed first, with the use of variables with nonmissing data. Then variables with increasing levels of missing data were imputed with the use of previously imputed variables based on the optimal matching criteria. The order in which variables were imputed was also determined to some extent by the substantive nature of the variables. For example, basic demographic variables (such as age) were imputed first and were used to impute education variables (such as attendance status and enrollment intensity). These variables were used to impute financial aid variables (such as aid and loan amounts).

In some cases, logical imputations were used to ensure data consistency. That is, if the relationship between multiple variables specified that the value of a known variable required a specific value of an imputed variable, then the imputed variable was assigned the appropriate value on the basis of the known-variable value. After the logical imputations were implemented, an integrated system using nonparametric classification trees and weighted sequential hot deck was used to impute missing data. The weight used was the latest intermediate version of the student analysis weight. This methodology and an evaluation of a few different imputation methods are described in Creel and Krotki (2006).

For all variables, a statistical process called *classification and regression trees* (CART) was used to identify the matching criteria most closely related to the variable being imputed (Breiman et al. 1984). This step produced a number of imputation classes, each of which contained a set of donors used to impute recipients belonging to that class. Imputation classes were formed on the basis of a CART analysis of likely candidates for variables related to those being imputed. Once a variable was imputed, it was made available to CART for subsequent classification trees. The resulting imputation classes varied for each variable or block of variables input to CART. Creating a nonparametric classification tree for each variable that was to be imputed would have been ideal but resource-intensive. Efforts were made to define groups of related variables to be imputed for which a single nonparametric classification tree could be created; then the related variables to be imputed were imputed as a vector.

Next, the imputation classes were used as input to a SAS macro that implemented the weighted sequential hot-deck procedure. Additionally, data were sorted within each imputation class to increase the chance of obtaining a close match between donor and recipient. The hot-deck process is sequential in that the search for donors occurs sequentially, starting with the recipient and progressing up and down the sorted file to find the set of eligible donors from which a random selection of one is made. The process is weighted in that it incorporates the sample weight of each record in the search and selection routine (Cox 1980; Iannacchione 1982, February). The values that were imputed have been flagged.

After the imputation process was completed, a set of checks was implemented to ensure the quality of the imputation process. For example, the number of times a donor was used, counts for each level of the imputed variable before and after imputation, and unweighted and weighted distributions before and after imputation within imputation classes were reviewed. In some cases, further intervention was required to ensure accuracy and consistency of imputation as determined by preexisting edit rules. For example, to impute the family size for a dependent student when it was known that the parents' marital status was *married,* the potential pool of donors was limited to those with married parents to prevent imputation of *family size* as a size of 2. After a value for a gateway or skip question was imputed, legitimate skip values for the subsequent questions were assigned.

Because of the number of variables and the complexity of the relationships among them, identifying and eliminating all inconsistencies was virtually impossible. The objective was to reduce inconsistencies as much as possible, especially for key analytic variables. The objective of the imputation program was to efficiently impute all missing data; the aim was to replace missing data with data that were valid in cases, with only a few relatively minor and unimportant exceptions.

Some results of the imputation process appear in appendix H, which presents the percentage missing for each variable subject to imputation for all students, undergraduate students, and graduate/first-professional students, and pre- and post-imputation distributions for eight key variables.

5.6 Composite and Derived Variable Construction

Analysts created the analytic variables by examining the data available for each student from the various data sources, prioritizing the data sources on an item-by-item basis, and reconciling discrepancies within and between sources. In some cases, the derived or composite variables were created by simple assignment of a value from the available source with the highest priority. In other cases, interview items were recoded or otherwise summarized to create a derived variable (for a

listing of the set of analysis variables derived for NPSAS:08, see appendix I). Details about the creation of each variable appear in the variable descriptions contained in the ECB and the DAS.

This page intentionally left blank.

Chapter 6. Unit Nonresponse Bias Analyses, Weighting, and Variance Estimation

Statistical analysis weights were computed for *study respondents* (defined in section 3.2) so that the study respondents would represent the target population described in section 2.1. The statistical analysis weights compensated for the unequal probability of selection of institutions and students in the 2007–08 National Postsecondary Student Aid Study (NPSAS:08) sample. The weights also adjusted for multiplicity at the institutional and student levels, unknown student eligibility, nonresponse, and poststratification. The institutional weight was computed and then used as a component of the student weight. Weights were computed for study respondents as the product of the following 10 weight components:

- institutional sampling weight (WT1);
- institutional multiplicity adjustment (WT2);
- institutional poststratification adjustment (WT3);
- institutional nonresponse adjustment (WT4);
- student sampling weight (WT5);
- student multiplicity adjustment (WT6);
- student unknown eligibility adjustment (WT7);
- student not located adjustment (WT8);
- student other nonresponse adjustment (WT9); and
- student poststratification adjustment (WT10).

Each weight component, described in the following sections, represents either a probability of selection or a weight adjustment. All nonresponse, extreme-weight, and poststratification adjustments were computed with the use of RTI's proprietary generalized exponential models (GEM) (Folsom and Singh 2000), which are similar to logistic models using bounds for adjustment factors and bounds on variance inflation. The GEM approach is a general version of weighting adjustments based on Deville and Särndal's logit model (1992). GEM is not a competing method of weighting class adjustment; rather, it is a unified approach to the nonresponse, poststratification, and extreme-weight adjustments. GEM controls at the margins as opposed to controlling at the cell level, as with weighting class adjustments. This locus of control allows consideration of greater numbers of variables. GEM is designed so that the sum of the unadjusted weights for all eligible units equals the sum of the adjusted weights for the respondents. GEM also constrains the nonresponse adjustment factors to be greater than or equal to 1.

To prevent the variance from becoming too large, the bounds on adjustment factors were loosened, where necessary. The unequal weighting effects (UWEs) and maximum adjustment factors were monitored to ensure reasonable values.

Weight adjustments produced by the GEM approach were calculated using PROC WTADJUST, a new procedure in SUDAAN Release 10 (RTI International 2008). A key feature and advantage of the GEM software is that the nonresponse adjustment and weight trimming and smoothing are accomplished in one step. Lower and upper bounds are set on the weight adjustment factors. The bounds on the weight adjustment factors can vary, depending on whether the weight falls inside or outside a range, such as the one defined by the bounds used to identify extreme weights (median weight ± 3 times the interquartile range). This variability in adjustment factors allows different bounds to be set for adjustments for weights that are considered high extreme (weight = median + 3 times the interquartile range), low extreme (weight = median − 3 times the interquartile range), or nonextreme. In this way, the extreme weights can be controlled, and the design effect due to unequal weighting can be reduced. (For details on the GEM Procedure, see appendix J.)

The bias in an estimated mean based on respondents, $\bar{y}_R$, is the difference between this mean and the target parameter, π (i.e., the mean that would be estimated if a complete census of the target population were conducted and everyone responded). This bias can be expressed as follows:

$$B(\bar{y}_R) = \bar{y}_R - \pi.$$

The estimated mean based on nonrespondents, $\bar{y}_{NR}$, can be computed if data for the particular variable are available for most of the nonrespondents. The true target parameter, π, can be estimated for these variables as follows:

$$\hat{\pi} = (1 - \eta)\bar{y}_R + \eta\bar{y}_{NR},$$

where η is the weighted unit (or item) nonresponse rate. For the variables that are from the frame, rather than from the sample, π can be estimated without sampling error. The bias can then be estimated as follows:

$$\hat{B}(\bar{y}_R) = \bar{y}_R - \hat{\pi}$$

or, equivalently,

$$\hat{B}(\bar{y}_R) = \eta(\bar{y}_R - \bar{y}_{NR}).$$

This formula shows that the estimate of the nonresponse bias is the difference between the mean for respondents and that for nonrespondents, multiplied by the weighted nonresponse rate.

Nonresponse bias analysis was conducted whenever the response rate at any level (institutions, students, items) was less than 85 percent.[31] Institutional and student nonresponse bias analyses were performed and are described in sections 6.1 and 6.2, respectively. An item nonresponse bias analysis also was performed and is described in section 6.3. Section 6.4 discusses variance estimation, including Taylor series, bootstrap replicate weights, and variance approximation using design effects.

[31] See *NCES Statistical Standards* for a discussion of nonresponse bias analysis (U.S. Department of Education 2003).

6.1 Institutional Nonresponse Bias Analysis and Weighting

6.1.1 Initial Institutional Weight Components

There were two initial institutional weight components.

Institutional Sampling Weight (WT1). The sampling weight for each sample institution was the reciprocal of its probability of selection. As described in appendix B, the probability of selection for institution *i* was

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

where

n_r = the sample size in stratum r,

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

and

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

Therefore, the institutional sampling weight was assigned as follows: WT1 = $1/\pi_r(i)$.

Institutional Multiplicity Adjustment (WT2). Each institution on the sampling frame initially had one chance of selection and an associated probability of selection; however, during institutional contacting and enrollment list collection some institutions were identified as having multiple chance of selection. That is, for about 10 sample institutions, student enrollment lists were provided that represented more than one institution without clearly identifying which institution or campus each student attended. The sample of students was selected from the one list. These institutions were treated as having multiple chances of being selected into the sample because each institution had an initial probability of selection, but the additional institutions represented on the list also had probabilities of selection. Therefore, the weight of the sample institution, which is based on the initial probability of selection, needed to be adjusted to account for the actual probability of selection for the group of institutions represented by the list. The number of chances of the institution's being selected was based on the number of institutions that were represented on the enrollment list.

This issue of some sample institutions having multiple chances of selection is because the lists for some sample institutions came from a system office or a main campus, and these lists contained students from the sample institution as well as one or more additional institutions. Some of these lists clearly identified the campus that each student attended, and each campus was treated as a separate institution. Hence, no adjustment was necessary in that case because each institution had its own probability of selection. However, other lists did not clearly identify the campus that each student attended. Also, some sample institutions had merged with another institution, and the lists for these institutions contained students from the original institution and the institution with which it merged.

When an institution had two chances of selection, a multiplicity adjustment was performed by first estimating, as if the selections were independent, the probability that either record could be selected:

$$p(A \text{ or } B) = p(A) + p(B) - p(A)p(B).$$

Then, the new sampling weight was calculated as the reciprocal of this probability:

$$\text{NEW_WT1} = 1/p(A \text{ or } B).$$

When an institution had three chances of selection, a multiplicity adjustment was performed by first estimating the probability that any record could be selected:

$$p(A \text{ or } B \text{ or } C) = (p(A) + p(B) + p(C)) - (p(A)p(B) + p(A)p(C) + p(B)p(C)) + p(A)p(B)p(C).$$

Then, the new sampling weight was calculated as the reciprocal of this probability:

$$\text{NEW_WT1} = 1/p(A \text{ or } B \text{ or } C).$$

When an institution had four or more chances of selection, a multiplicity adjustment was performed by first estimating the probability that any record could be selected:

$$p(A \text{ or } B \text{ or } C \text{ or } D \ldots) \approx 1 - (1 - p(A)) * (1 - p(B)) * (1 - p(C)) * (1 - p(D)) * \ldots.$$

Then, the new sampling weight was calculated as the reciprocal of this probability:

$$\text{NEW_WT1} = 1/p(A \text{ or } B \text{ or } C \text{ or } D \ldots).$$

Finally, the multiplicity adjustment factor was derived by dividing the new sampling weight by the old sampling weight, WT2 = NEW_WT1/WT1, for the institutions with positive multiplicity, and by setting it to unity (1.00) for all other institutions. Consequently, the product of WT1 and WT2 equals NEW_WT1 for the institutions with positive multiplicity, and it equals WT1 for all other institutions.

6.1.2 Assessing Institution Nonresponse Bias

As shown in table 9, there were 1,730 respondents from among the 1,960 eligible sample institutions (89 percent unweighted and 90 percent weighted). The institution weighted response rate is less than 85 percent for two of the nine institution types: (1) private, not-for-profit, less-than-4-year institutions; and (2) private, for-profit, less-than-2-year institutions. The weighted response rates, by type of institution, range from 81 percent for private, for-profit, less-than-2-year institutions to 95 percent for public, 4-year, non-doctorate-granting institutions.

As shown in table 10, the institution weighted response rate is less than 85 percent for five state-level sectors. The weighted response rates, by state-level sector, range from 74 percent for Texas private, not-for-profit, 4-year institutions to 100 percent for Georgia and Minnesota public, 2-year and 4-year institutions.

A nonresponse bias analysis was conducted for the two national-level sectors and five state-level sectors with a weighted response rate less than 85 percent. The nonresponse bias was estimated for variables known—that is, nonmissing—for most respondents and nonrespondents. Extensive data are available for all institutions from the Integrated Postsecondary Education Data System (IPEDS); the following variables were used:

- Carnegie classification;
- degree of urbanization;

- Office of Business Economics (OBE) region;
- Historically Black College or University indicator;
- Hispanic serving institution indicator;
- percentage of students receiving federal grant aid;
- percentage of students receiving state/local grant aid;
- percentage of students receiving institutional grant aid;
- percentage of students receiving student loan aid;
- percentage of students enrolled: Hispanic;
- percentage of students enrolled: Asian or Pacific Islander;
- percentage of students enrolled: Black, non-Hispanic;
- total undergraduate enrollment;
- male undergraduate enrollment;
- female undergraduate enrollment;
- total graduate/first-professional enrollment;
- male graduate/first-professional enrollment; and
- female graduate/first-professional enrollment.[32]

First, for the institution-level variables listed, the nonresponse bias was estimated and tested to determine whether the bias was significant at the 5 percent level. Second, nonresponse adjustments were computed, and the variables listed were included in the nonresponse models. The nonresponse adjustments (see section 6.1.3) were designed to significantly reduce or eliminate nonresponse bias for variables included in the models. Third, after the weights were computed, any remaining bias was estimated for the variables listed above, and statistical tests were performed to check the remaining significant nonresponse bias.

As shown in table 54, the institution weighting adjustments eliminated some, but not all, significant bias. For private, not-for-profit, less-than-4-year institutions and private, for-profit, less-than-2-year institutions, before weighting, 15.5 percent and 6.3 percent, respectively, of the variable categories were significantly biased. (For detailed results of the nonresponse bias analysis and for state-level results, see appendix K. For results of the nonresponse bias analysis after weight adjustments, see section 6.1.4.)

[32] For the continuous variables, categories were formed on the basis of quartiles.

Table 54. Summary of institution nonresponse bias analysis, by type of institution: 2008

Nonresponse bias statistics	Private not-for-profit less-than-4-year	Private for-profit less-than-2-year
Before weight adjustments		
Mean estimated relative bias	0.10	0.07
Median estimated relative bias	0.06	0.05
Percent significant bias	15.52	6.25
After weight adjustments		
Mean estimated relative bias	0.17	0.13
Median estimated relative bias	0.16	0.09
Percent significant bias	#	1.56

Rounds to zero.
NOTE: Nonresponse bias analysis was conducted for the two types of institutions with a weighted response rate less than 85 percent.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

6.1.3 Adjusting Institution Weights

There were two additional institution weight components.

Institution Poststratification Adjustment (WT3). To ensure population coverage, the institution sampling weight adjusted for multiplicity was adjusted, with the use of GEM, to control totals for enrollment by institution type, state (six states plus an "Other states" category), and size (*small* vs. *large*). The enrollment totals came from the 12-month unduplicated headcount from the IPEDS:08.enrollment file

Table 55 presents the variables associated with the control totals and the average weight adjustment factors, by these variables. The weight adjustment factors from GEM met the following constraints:

- minimum: 0.24;
- median: 1.00; and
- maximum: 2.17.

Table 55. Weight adjustment factors for institution poststratification: 2008

Model predictor variable[1]	Control total[2]	Average weight adjustment factor (WT3)
Total	25,399,655	†
Public less-than-2-year	84,644	1.11
Public 2-year, small	1,335,278	0.91
Public 2-year, large	8,860,428	1.02
Public 4-year non-doctorate-granting, small	445,149	0.95
Public 4-year non-doctorate-granting, large	1,883,388	0.99
Public 4-year doctorate-granting, small	1,306,405	1.03
Public 4-year doctorate-granting, large	4,570,558	1.00
Private not-for-profit less-than-4-year, large	120,270	1.48
Private not-for-profit 4-year non-doctorate-granting, small	276,468	0.94
Private not-for-profit 4-year non-doctorate-granting, large	1,614,297	1.00
Private not-for-profit 4-year doctorate-granting, small	154,711	0.95
Private not-for-profit 4-year doctorate-granting, large	2,432,291	0.99
Private for-profit less-than-2-year	475,423	1.01
Private for-profit 2-year-or-more, small	163,092	0.66
Private for-profit 2-year-or-more, large	1,677,253	1.02
California, small	137,739	1.25
California, large	3,280,085	0.97
Georgia, small	75,230	0.99
Georgia, large	511,715	0.97
Illinois, small	87,981	0.84
Illinois, large	1,199,682	1.00
Minnesota, small	41,520	0.77
Minnesota, large	456,749	0.97
New York, small	97,855	0.97
New York, large	1,412,585	0.98
Texas, small	133,735	1.15
Texas, large	1,575,929	0.95
Other states, small	1,174,428	1.05
Other states, large	14,415,615	0.98
All states, small	100,746	1.39
All states, large	698,061	1.06

† Not applicable.
[1] Size for poststratification weighting classes was based on the median enrollment within sector for the institutions on the sampling frame.
[2] Control totals are the sum of enrollment across institutions based on IPEDS:08 enrollment data.
NOTE: Detail may not sum to totals because of rounding. IPEDS = Integrated Postsecondary Education Data System.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

Institution Nonresponse Adjustment (WT4). The definition of *institutional respondent* is provided in section 3.1. A weighting adjustment using GEM was performed to compensate for nonresponding institutions. The nonresponse adjustments were designed to significantly reduce or eliminate nonresponse bias for variables included in the models. Predictor variables were chosen that were thought to predict response status and were nonmissing for most respondents and nonrespondents. The candidate predictor variables were those used in the nonresponse bias analysis already described, with the addition of sector and state. These variables were known for most respondents and nonrespondents. Any missing data were minimal and were imputed so that the variables input into GEM had no missing values.

Predictors used in the nonresponse modeling included all the candidate predictor variables identified, as well as certain potentially important interactions. To identify these interactions, the chi-square automatic interaction detection (CHAID) algorithm (Kass 1980) was used. CHAID is a hierarchical clustering algorithm that successively partitions individuals according to categorical predictors for a categorical dependent variable. The algorithm begins with all study individuals as a whole and cycles over each predictor, finding for each predictor an optimal partition of the individuals according to its levels. The most significant optimal partition is then retained, and the CHAID algorithm is again applied to the members of that partition to find further partitions, using the remaining predictors. The algorithm is stopped after a specified number of partitioning steps or if none of the partitions at a given step is found to be significant.

Application of the CHAID algorithm provided interaction terms for the nonresponse adjustment models. CHAID was run for up to three segments, resulting in identification of two-way and three-way interactions.

Some of the predictor variables/response categories ("unavailable or unknown" for state/local grant aid, institutional grant aid, or student loan; percent enrolled: Asian/Pacific Islander; and percent enrolled: Black, non-Hispanic) were dropped from the adjustment model because of singularity, which prevents the model from running properly. Singularity occurs when a combination of variables can be used to determine the values of another variable (e.g., institutions that had unavailable or unknown percentages of federal grant aid were the same institutions that had unavailable or unknown information for other aid variables).

Table 56 presents the response rates and the resulting adjustment factors, by the model variables. The weight adjustment factors from GEM met the following constraints:

- minimum: 1.00;
- median: 1.07; and
- maximum: 1.96.

Table 56. Weight adjustment factors for institution nonresponse adjustment: 2008

Model predictor variable	Number of respondents	Weighted response rate[1]	Average weight adjustment factor (WT4)[2]
Total	1,730	90.1	1.12
Institution strata			
Public less-than-2-year	20	93.3	1.07
Public 2-year	410	91.2	1.08
Public 4-year non-doctorate-granting	180	95.1	1.05
Public 4-year doctorate-granting	270	89.6	1.13
Private not-for-profit less-than-4-year	10	86.5	1.17
Private not-for-profit 4-year non-doctorate-granting	300	87.5	1.16
Private not-for-profit 4-year doctorate-granting	250	86.3	1.17
Private for-profit less-than-2-year	70	79.1	1.27
Private for-profit 2-year-or-more	220	89.8	1.11
Sector and state classification			
Public 4-year, public 2-year, private not-for-profit 4-year, and private for-profit degree-granting institutions			
California	150	84.1	1.20
Georgia	120	97.7	1.03
Illinois	90	92.2	1.10
Minnesota	100	96.6	1.04
New York	150	94.9	1.05
Texas	120	90.4	1.14
Other states	910	90.7	1.13
Public less-than-2-year, private not-for-profit less-than-4-year, and private for-profit non-degree institutions			
All states	100	81.0	1.23
Carnegie classification code			
Associate's	520	90.9	1.09
Research and doctoral	230	87.8	1.15
Master's	430	94.1	1.07
Baccalaureate	270	90.8	1.11
Special focus and other	150	78.8	1.28
Unavailable or unknown	130	81.5	1.21
Bureau of Economic Analysis Code (Office of Business Economics) region[3]			
New England	90	85.2	1.21
Mid East	310	91.8	1.10
Great Lakes	240	90.0	1.13
Plains	200	92.5	1.09
Southeast	400	92.1	1.09
Southwest	180	89.5	1.15
Rocky Mountains	40	95.1	1.06
Far West	230	85.9	1.18
Outlying areas	30	91.4	1.11

See notes at end of table.

Table 56. Weight adjustment factors for institution nonresponse adjustment: 2008—Continued

Model predictor variable	Number of respondents	Weighted response rate[1]	Average weight adjustment factor (WT4)[2]
Percent receiving federal grant aid			
0 < percent ≤ 21	430	89.8	1.12
21 < percent ≤ 33	420	88.6	1.13
33 < percent ≤ 51	410	92.4	1.09
51 < percent	390	89.7	1.14
Unavailable or unknown	90	88.4	1.18
Percent receiving state/local grant aid			
0 < percent ≤ 12, or unavailable/unknown	480	87.3	1.19
12 < percent ≤ 29	410	90.3	1.12
29 < percent ≤ 51	410	89.6	1.11
51 < percent	420	93.6	1.06
Percent receiving institutional grant aid			
0 < percent ≤ 4, or unavailable/unknown	500	89.3	1.13
4 < percent ≤ 26	410	91.8	1.09
26 < percent ≤ 65	410	90.2	1.12
65 < percent	410	87.3	1.15
Percent receiving student loan aid			
0 < percent < 25	420	90.7	1.10
25 < percent ≤ 51	420	91.7	1.09
51 < percent ≤ 72, or unavailable/unknown	510	88.8	1.14
72 < percent	380	85.8	1.16
Percent enrolled: Hispanic			
0 < percent ≤ 2, or unavailable/unknown	620	91.4	1.11
2 < percent ≤ 4	320	89.2	1.13
4 < percent ≤ 12	390	90.4	1.12
12 < percent	400	89.0	1.14
Percent enrolled: Asian or Pacific Islander			
0 < percent ≤ 1, or unavailable/unknown	630	92.5	1.10
1 < percent ≤ 2	280	88.7	1.15
2 < percent ≤ 5	400	90.2	1.12
5 < percent	420	88.6	1.14
Percent enrolled: Black, non-Hispanic			
0 < percent ≤ 3	440	91.6	1.11
3 < percent ≤ 7	440	87.5	1.15
7 < percent ≤ 18, or unavailable/unknown	430	90.4	1.12
18 < percent	420	91.3	1.10
Total undergraduate enrollment			
0 < percent ≤ 1,383	400	84.4	1.17
1,383 < percent ≤ 3,770	430	90.1	1.12
3,770 < percent ≤ 10,926	430	90.8	1.10
10,926 < percent	430	90.7	1.10
Unavailable or unknown	40	85.3	1.18

See notes at end of table.

Table 56. Weight adjustment factors for institution nonresponse adjustment: 2008—Continued

Model predictor variable	Number of respondents	Weighted response rate[1]	Average weight adjustment factor (WT4)[2]
Total male undergraduate enrollment			
0 < number ≤ 539	400	83.5	1.18
539 < number ≤ 1,559, or unavailable/unknown	470	91.5	1.10
1,559 < ≤ 4,687	430	90.7	1.11
4,687 < number	430	90.5	1.10
Total female undergraduate enrollment			
0 < number ≤ 819	400	82.7	1.19
819 < number ≤ 2,229, or unavailable/unknown	470	91.7	1.10
2,229 < number ≤ 6,266	430	90.7	1.10
6,266 < number	430	90.6	1.10
Total graduate/first-professional enrollment			
0 < number ≤ 489	230	86.9	1.18
489 < number ≤ 1,431	250	94.5	1.09
1,431 < number ≤ 3,568	240	91.4	1.11
3,568 < number	230	88.8	1.14
Unavailable or unknown	780	90.1	1.11
Total male graduate/first-professional enrollment			
0 < number ≤ 168, or unavailable/unknown	1,010	89.9	1.13
168 < number ≤ 511.5	240	93.1	1.10
511.5 < number ≤ 1,353	250	92.7	1.10
1,353 < number	230	88.5	1.15
Total female graduate/first-professional enrollment			
0 < number ≤ 281, or unavailable/unknown	1,010	89.7	1.13
281 < number ≤ 871	240	93.1	1.09
871 < number ≤ 2,150	250	92.7	1.09
2,150 < number	230	88.8	1.14
Degree of urbanization			
Large city	410	89.0	1.15
Mid-sized city	220	89.8	1.14
Small city	260	91.4	1.10
Large suburb	330	88.3	1.15
Mid-sized suburb	30	81.1	1.22
Small suburb	40	93.2	1.07
Urban area on fringe of town	50	88.0	1.13
Urban area distant from town	100	96.1	1.04
Urban area remote from town	90	93.7	1.06
Rural area on fringe of town	140	92.6	1.08
Rural area distant from town	30	87.4	1.12
Rural area remote from town	10	90.7	1.10
Historically Black College or University			
Yes	40	84.8	1.16
No	1,680	90.2	1.12
Unavailable or unknown	10	87.9	1.21

See notes at end of table.

Table 56. Weight adjustment factors for institution nonresponse adjustment: 2008—Continued

Model predictor variable	Number of respondents	Weighted response rate[1]	Average weight adjustment factor (WT4)[2]
Hispanic Serving Institution			
Yes	160	90.5	1.11
No	1,570	90.0	1.12
CHAID segments			
In California and associate's or research and doctoral	90	81.7	1.26
In California and not associate's or research and doctoral	60	93.6	1.12
In Georgia and public 2-year or public 4-year non-doctorate-granting	70	100.0	1.01
In Georgia and public 4-year doctorate-granting, private not-for-profit 4-year, or private for-profit degree-granting	40	88.0	1.05
In Illinois, Minnesota, or New York, percent enrolled Hispanic ≤ 12 or unavailable/unknown, and percent receiving state/local grant aid ≤ 29 or unavailable/unknown	70	85.9	1.16
In Illinois, Minnesota, or New York, percent enrolled Hispanic ≤ 12 or unavailable/unknown, and percent receiving state/local grant aid > 29	200	95.6	1.04
In Illinois, Minnesota, or New York, percent enrolled Hispanic > 12	60	99.3	1.01
In Texas or not in one of the six states but public 2-year, public 4-year, private not-for-profit 4-year, or private for-profit degree-granting and total male undergraduate enrollment ≤ 539	170	82.9	1.26
In Texas or not in one of the six states but public 2-year, public 4-year, private not-for-profit 4-year, or private for-profit degree-granting, total male undergraduate enrollment > 539 or unavailable/unknown, and total female graduate/first-professional enrollment ≤ 281 or unavailable/unknown	110	86.7	1.19
In Texas or not in one of the six states but public 2-year, public 4-year, private not-for-profit 4-year, or private for-profit degree-granting, total male undergraduate enrollment > 539 or unavailable/unknown, and total female graduate/first-professional enrollment > 281 and ≤ 871	150	96.0	1.05
In Texas or not in one of the six states but public 2-year, public 4-year, private not-for-profit 4-year, or private for-profit degree-granting, total male undergraduate enrollment > 539 or unavailable/unknown, and total female graduate/first-professional enrollment > 871	590	91.0	1.10
Public less-than-2-year, private not-for-profit less-than-4-year, or private for-profit non-degree-granting	100	81.0	1.23

[1] The response rate is expressed as a percentage.
[2] The average weight adjustment factor is expressed as a number.
[3] New England = Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont; Mid East = Delaware, District of Columbia, Maryland, New Jersey, New York, Pennsylvania; Great Lakes = Illinois, Indiana, Michigan, Ohio, Wisconsin; Plains = Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota; Southeast = Alabama, Arkansas, Florida, Georgia, Kentucky, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, West Virginia; Southwest = Arizona, New Mexico, Oklahoma, Texas; Rocky Mountains = Colorado, Idaho, Montana, Utah, Wyoming; Far West = California, Nevada, Oregon, Washington; Outlying Areas = Alaska, Hawaii, and Puerto Rico. Alaska and Hawaii were reclassified from the West to the outlying areas for the purposes of NPSAS.
NOTE: Numbers may not sum to totals due to rounding. Categories were formed from continuous variables based on quartiles. CHAID = chi-square automatic interaction detection.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

6.1.4 Institution Weighting Adjustment Performance

As shown in table 54, the institution weighting adjustments eliminated some, but not all, significant bias. After weighting, no significant bias remained for the variables analyzed for private, not-for-profit, less-than-4-year institutions; 2 percent significant bias remained for the variables analyzed for private, for-profit, less-than-2-year institutions. Significant bias was reduced for the variables known for study respondents, which are considered to be some of the more analytically important variables and are correlated with many of the other variables. Appendix K contains detailed tables showing the estimated bias before and after weight adjustments for each domain for which nonresponse bias analysis was conducted.

Table 57 summarizes the institution weight distributions and the variance inflation due to unequal weighting (i.e., UWE, by institution type). The median institution weights range from 1.1 for public, 4-year, doctorate-granting institutions to 18.2 for private, not-for-profit, less-than-4-year institutions. The mean institution weight ranges from 1.2 for public, 4-year, doctorate-granting institutions to 23.6 for private, not-for-profit, less-than-4-year institutions. The UWE is 11.2 overall and ranges from 1.0 for public, 4-year, doctorate-granting institutions to 13.1 for private, for-profit, 2-year–or-more institutions.

Table 57. Institution weight distribution and unequal weighting effects, by type of institution: 2008

Type of institution	Minimum	First quartile	Median	Third quartile	Maximum	Mean	Unequal weighting effect[1]
Total	0.52	1.04	1.23	2.62	307.84	3.76	11.18
Public less-than-2-year	1.23	1.83	5.09	8.63	38.25	7.84	2.24
Public 2-year	0.83	1.05	1.35	2.92	77.75	2.61	3.95
Public 4-year non-doctorate-granting	0.92	1.01	1.12	2.03	18.03	1.90	2.29
Public 4-year doctorate-granting	0.95	1.04	1.10	1.20	3.67	1.16	1.04
Private not-for-profit less-than-4-year	4.98	7.07	18.18	31.56	66.88	23.62	1.64
Private not-for-profit 4-year non-doctorate-granting	0.76	1.04	1.45	3.32	73.01	3.15	4.07
Private not-for-profit 4-year doctorate-granting	0.79	1.06	1.21	1.71	21.43	2.00	2.59
Private for-profit less-than-2-year	1.13	3.89	8.95	23.27	159.92	18.31	3.01
Private for-profit 2-year-or-more	0.52	0.95	1.35	4.98	307.84	7.53	13.14

[1] Unequal weighting effect calculated as follows: sample size × (sum of the weights^2) / (sum of the weights)^2.
NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

To assess the overall predictive ability of the nonresponse model, a Receiver Operating Characteristic (ROC) curve was used (Hanley and McNeil 1982). The ROC 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.[33] The ROC curve was developed in the following manner. For any specified probability, *c*, two proportions were calculated:

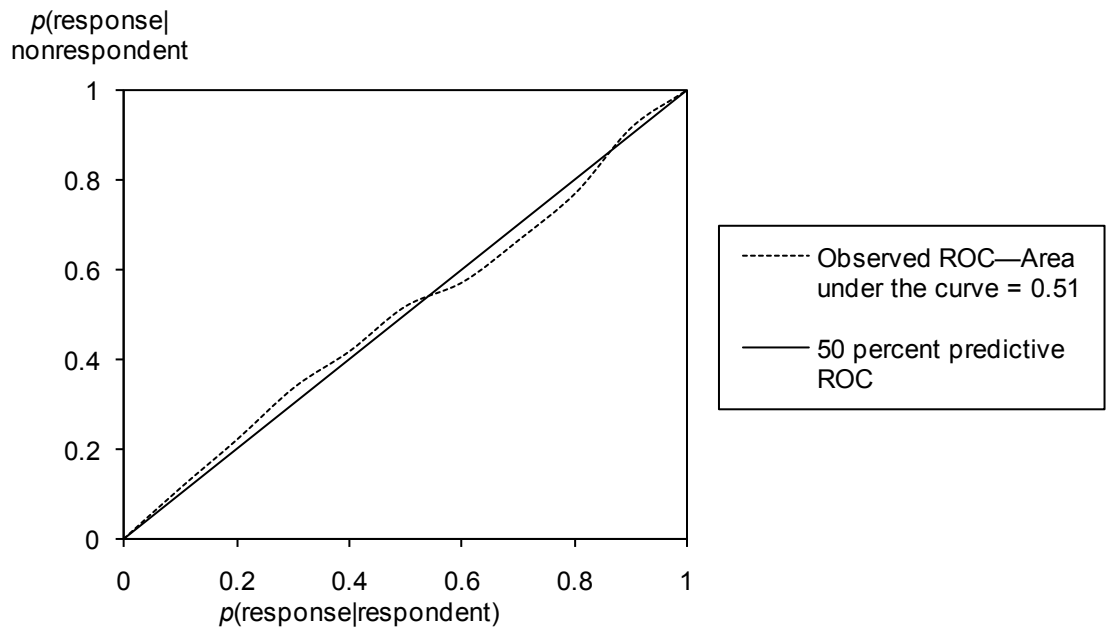
- the proportion of respondents with a predicted probability of response greater than *c,* and

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

- the proportion of nonrespondents with a predicted probability of response greater than *c*.

The plot of the first probability against the second, for *c* from zero to 1, resulted in the ROC curve shown in figure 19. The area under the curve equals the probability that the fitted model correctly classifies two randomly chosen individuals—one of whom is a true respondent, while the other is a true nonrespondent—where the individual 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, so the model always classifies the two individuals correctly. Figure 19 shows that the area under the ROC curve is 0.51, so the predicted probabilities give the correct classification 51 percent of the time (about one of every two pairings). Predictive probabilities from ROC curves can also be interpreted in terms of the nonparametric Wilcoxon test statistic, where the ROC area of 0.51 equals the value of the Wilcoxon test statistic. Viewed in this way, the Wilcoxon test does not reject the null hypothesis of no predictive ability. This result can be interpreted to mean that the variables used in the model are not definitive predictors of a sample institution's overall response propensity.

Figure 19. Receiver Operating Characteristic curve for overall institution response propensity: 2008



NOTE: ROC = Receiver Operating Characteristic.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

6.2 Student Nonresponse Bias Analysis and Weighting

6.2.1 Initial Student Weight Components

There were four initial student weight components.

Student Sampling Weight (WT5). The overall student sampling strata were defined by crossing the institution sampling strata with the student strata within institutions. (For the overall

sampling rates for these sampling strata, see appendix B.) The sample 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 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 sampling rate and $\pi_r(i)$ = the institution's probability of selection.

As discussed in appendix B, the institution-specific rates were designed to obtain the desired sample sizes and achieve nearly equal weights within the overall student strata.

If the institution's enrollment list was larger than expected on the basis of the IPEDS data, the preloaded student sampling rates would yield larger-than-expected sample sizes. Likewise, if the enrollment list was smaller than expected, the sampling rates would yield smaller-than-expected sample sizes. To maintain control on the sample sizes, the sampling rates were adjusted, when necessary, so that the number of students selected within an institution did not exceed 300. A minimum sample size constraint of 10 students was also imposed to ensure sufficient yield for variance estimation.

The student sampling weight was calculated as the reciprocal of the institution-specific student sampling rates, or

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

Student Multiplicity Adjustment (WT6). Students who attended more than one eligible institution during the 2007–08 academic year had multiple chances of being selected; that is, they could have been selected from any of the institutions they attended. These students therefore had a higher probability of being selected than was represented in their sampling weight.

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

$$\text{WT6} = 1/M,$$

where M is the multiplicity, or number of eligible institutions attended. The multiplicity was determined from the student interview, the Pell Grant payment file, and the National Student Loan Data System. If student multiplicity was missing, the average number of institutions attended, based on students with nonmissing data, was used. Averages were computed within type of institution and federal aid receipt status.

The weight adjustment factors met the following constraints:

- minimum: 0.20;
- median: 1.00; and
- maximum: 1.00.

Student Unknown Eligibility Adjustment (WT7). Final eligibility status could not be determined for nonresponding students who were never contacted. These students were treated as

eligible, and their weights were adjusted to compensate for the small portion of students who were actually ineligible (as described below).

Weighting classes were defined by the intersection of institution type with the students' matching status to financial aid files (Central Processing System [CPS], Pell Grant, and Stafford Loan). Table 58 presents the weight adjustment factors applied to the students with unknown eligibility. These weight adjustment factors were based on the estimated rate of eligibility among students with known eligibility status. For the known-eligible students, the weight adjustment factor was set equal to 1.

Table 58. Weight adjustment factors for unknown student eligibility status: 2008

Weighting class (institution level, by student type, by matching status to financial aid files)		Number adjusted for unknown eligibility	Weight adjustment factor (WT7)
Total		4,660	†
Public less-than-2-year	Matched Pell or Stafford file	#	1.00
	Matched CPS file only	#	1.00
	No matches	140	0.66
Public 2-year	Matched Pell or Stafford file	70	1.00
	Matched CPS file only	10	1.00
	No matches	2,410	0.86
Public 4-year non-degree-granting, undergraduate	Matched Pell or Stafford file	20	1.00
	Matched CPS file only	#	1.00
	No matches	280	0.94
Public 4-year non-degree-granting, graduate	Matched Pell or Stafford file	#	1.00
	Matched CPS file only	#	1.00
	No matches	#	0.95
Public 4-year degree-granting, undergraduate	Matched Pell or Stafford file	80	1.00
	Matched CPS file only	#	1.00
	No matches	610	0.96
Public 4-year degree-granting, graduate	Matched Pell or Stafford file	10	1.00
	Matched CPS file only	#	1.00
	No matches	70	0.96
Private not-for-profit less-than-4-year	Matched Pell or Stafford file	#	1.00
	Matched CPS file only	#	1.00
	No matches	20	0.60

See notes at end of table.

Table 58. Weight adjustment factors for unknown student eligibility status: 2008—Continued

Weighting class (institution level, by student type, by matching status to financial aid files)		Number adjusted for unknown eligibility	Weight adjustment factor (WT7)
Private not-for-profit 4-year non-degree-granting, undergraduate	Matched Pell or Stafford file	20	1.00
	Matched CPS file only	#	1.00
	No matches	240	0.91
Private not-for-profit 4-year non-degree-granting, graduate	Matched Pell or Stafford file	10	1.00
	Matched CPS file only	#	1.00
	No matches	40	0.92
Private not-for-profit 4-year degree-granting, undergraduate	Matched Pell or Stafford file	#	1.00
	Matched CPS file only	#	1.00
	No matches	220	0.96
Private not-for-profit 4-year degree-granting, graduate	Matched Pell or Stafford file	#	1.00
	Matched CPS file only	#	1.00
	No matches	70	0.94
Private for-profit less-than-2-year	Matched Pell or Stafford file	50	1.00
	Matched CPS file only	#	1.00
	No matches	150	0.49
Private for-profit 2-year	Matched Pell or Stafford file	#	1.00
	Matched CPS file only	#	1.00
	No matches	40	0.51
Private for-profit 4-year, undergraduate	Matched Pell or Stafford file	10	1.00
	Matched CPS file only	#	1.00
	No matches	80	0.83
Private for-profit 4-year, graduate	Matched Pell or Stafford file	#	1.00
	Matched CPS file only/No matches combined	10	0.94

† Not applicable.
# Rounds to zero.
NOTE: Numbers may not sum to totals due to rounding. CPS = Central Processing System.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

The weight adjustment factors met the following constraints:

- minimum: 0.49;
- median: 1.00; and
- maximum: 1.00.

6.2.2 Assessing Student Nonresponse Bias

As mentioned in section 3.2, a *study respondent* is defined as any student sample member who is determined eligible for the study and has valid data from any source for a selected set of key analytical variables. While these were the minimal data requirements, the vast majority of study respondents had considerably more complete data.

As shown in table 11, of the 132,800 eligible students, the unweighted and weighted response rate was 96 percent. The weighted study response rate for students is greater than 85 percent overall and for all institution types. The weighted study response rates by type of institution range from 89 percent for students in public, less-than-2-year institutions to 99 percent for students in private, for-profit, 2-year-or-more institutions.

As shown in table 12, the student weighted response rate is less than 85 percent for one state-level sector. The weighted response rates, by state-level sector, range from 84 percent for California public, 2-year institutions to about 100 percent for Illinois for-profit, degree-granting institutions.

In accordance with these results, a student-level nonresponse bias analysis was not necessary for any national-level sectors but was necessary for the one state-level sector with a weighted response rate less than 85 percent. The nonresponse bias was estimated for variables known—that is, nonmissing—for most respondents and nonrespondents. The following variables were used to assess student-level nonresponse bias:

- For all sample members
  - institution total enrollment;
  - CPS match (yes/no);
  - federal aid recipient (yes/no);
  - Pell Grant recipient (yes/no); and
  - Stafford Loan recipient (yes/no).
- For federally aided students
  - Pell Grant amount; and
  - Stafford Loan amount. [34]

Additionally, it was determined that percent part-time fall enrollment and in-state tuition were important variables to include in the nonresponse bias analysis for students in public, 2-year institutions. These variables are not known for both respondents and nonrespondents; however, institution-level data available from IPEDS were used to conduct the analyses.

First, for the variables listed, the nonresponse bias was estimated and tested to determine whether the bias was significant at the 5 percent level. Second, nonresponse adjustments were computed.[35] Third, after the weights were computed, any remaining bias was estimated for the variables listed, and statistical tests were performed to check the remaining significant nonresponse bias.

[34] For the continuous variables, categories were formed on the basis of quartiles.

[35] The nonresponse adjustments (see section 6.2.3) were not designed to significantly reduce or eliminate nonresponse bias for all variables tested for nonresponse bias. The two variables *percent part-time fall enrollment* and *in-state tuition* were not included, because they are relevant only for public, 2-year institutions and are a concern only for California.

For detailed state-level results, see appendix K. Results of the nonresponse bias analysis after weight adjustments are discussed in section 6.2.4.

6.2.3 Adjusting Student Weights

- There were three additional student weight components. The student weights were further adjusted for nonresponse, in two stages (inability to locate the student and other nonresponse), because the predictors of response propensity were potentially different for each of these nonresponse outcomes. Performing an additional adjustment for interview refusal was considered, but the number of study nonrespondents was not sufficient to support three nonresponse adjustments.

Using these two stages of nonresponse adjustment achieved greater reduction in nonresponse bias to the extent that different variables were significant predictors of response propensity at each stage.

Student Not Located Adjustment (WT8). The first type of adjustment for student nonresponse was an adjustment for the inability to locate the student. This weight adjustment was made to compensate for the potential study nonresponse bias. Predictor variables were chosen that were thought to predict response status and were nonmissing for both study respondents and nonrespondents. The candidate predictor variables included:

- institution type;
- state sector;
- being in one of six states with a state-representative sample of undergraduates (yes/no);
- region;
- institution enrollment from IPEDS file (categorical);
- student type;
- baccalaureate status;
- Pell Grant receipt (yes/no);
- Pell Grant amount (categorical);
- Stafford Loan receipt (yes/no);
- Stafford Loan amount (categorical);
- Parent Loan for Undergraduate Students (PLUS) amount (categorical);
- federal aid receipt (yes/no);
- institutional aid receipt (yes/no);
- state aid receipt (yes/no);
- any aid receipt (yes/no);
- CPS record indicator (yes/no);
- computer-assisted data entry record indicator (yes/no);

- computer-assisted telephone interviewing record indicator (yes/no);
- age group (three levels);
- education level (three levels);
- Social Security number indicator (yes/no);
- phone number count;
- e-mail address count; and
- mailing address count.

Predictors used in the nonresponse modeling included all the candidate predictor variables identified, as well as certain potentially important interactions. CHAID was used to identify these interactions (see the description in section 6.1.3). Application of the CHAID algorithm provided interaction terms for each of the nonresponse adjustment models. For each model, CHAID was run for up to three segments, resulting in identification of two-way and three-way interactions. Segments were retained if they were both statistically and practically significant.

The weight adjustments were computed with the use of GEM.[36] Table 59 presents the final predictor variables used in GEM to adjust the weights and the average weight adjustment factors resulting from these variables. The weight adjustment factors met the following constraints:

- minimum: 1.00;
- median: 1.00; and
- maximum: 5.65.

[36] See the description of GEM procedures at the beginning of this chapter.

Table 59. Weight adjustment factors for student location nonresponse adjustment: 2008

Model predictor variable	Number of located respondents	Weighted response rate	Average weight adjustment factor (WT8)
Total	129,640	97.4	1.02
Type of institution			
Public less-than-2-year	1,390	93.7	1.06
Public 2-year	34,350	95.0	1.04
Public 4-year non-doctorate-granting	16,130	98.6	1.01
Public 4-year doctorate-granting	29,860	98.4	1.02
Private not-for-profit less-than-4-year	1,640	99.2	1.01
Private not-for-profit 4-year non-doctorate-granting	13,290	98.2	1.02
Private not-for-profit 4-year doctorate-granting	16,610	98.8	1.01
Private for-profit less-than-2-year	5,700	98.6	1.02
Private for-profit 2-year-or-more	10,660	99.2	1.01
State-level institutional sectors			
All other states	73,210	98.1	1.02
California public 2-year	4,380	89.8	1.11
California public 4-year	3,760	98.2	1.01
California private not-for-profit 4-year	1,870	96.3	1.04
California private for-profit degree-granting	1,030	99.6	1.00
Georgia public 2-year	3,920	96.7	1.03
Georgia public 4-year	2,080	99.6	1.00
Georgia private not-for-profit 4-year	1,480	98.1	1.02
Georgia private for-profit degree-granting	1,610	99.7	1.00
Illinois public 2-year	2,300	94.2	1.06
Illinois public 4-year	1,800	98.4	1.02
Illinois private not-for-profit 4-year	2,160	99.3	1.01
Illinois private for-profit degree-granting	1,410	99.6	1.00
Minnesota public 2-year	4,260	97.2	1.03
Minnesota public 4-year	2,220	97.8	1.02
Minnesota private not-for-profit 4-year	1,280	99.3	1.00
Minnesota private for-profit degree-granting	1,270	96.2	1.04
New York public 2-year	4,000	98.3	1.02
New York public 4-year	3,110	99.0	1.01
New York private not-for-profit 4-year	3,360	98.7	1.01
New York private for-profit degree-granting	1,420	97.7	1.02
Texas public 2-year	2,150	95.2	1.05
Texas public 4-year	3,060	97.9	1.02
Texas private not-for-profit 4-year	1,240	99.4	1.01
Texas private for-profit degree-granting	1,280	99.5	1.01

See notes at end of table.

Table 59. Weight adjustment factors for student location nonresponse adjustment: 2008—Continued

Model predictor variable	Number of located respondents	Weighted response rate	Average weight adjustment factor (WT8)
Region[1]			
New England	5,230	98.1	1.02
Mid East	22,730	97.8	1.02
Great Lakes	19,630	98.2	1.02
Plains	15,840	98.5	1.02
Southeast	30,140	98.2	1.02
Southwest	12,460	97.1	1.03
Rocky Mountains	4,230	98.5	1.01
Far West	16,450	93.9	1.05
Outlying Areas	2,940	98.6	1.01
Institution enrollment size			
0 < enrollment total ≤ 3,856	32,390	98.9	1.01
3,856 < enrollment total ≤ 10,698	32,360	97.7	1.02
10,698 < enrollment total ≤ 24,146	32,320	96.5	1.03
24,146 < enrollment total	32,570	97.0	1.03
Education level (based on sampling list)			
Undergraduate	115,870	97.1	1.02
Graduate, master's, doctoral	11,740	98.9	1.01
First-professional	2,030	99.3	1.01
Age group			
15 to 23	72,430	97.6	1.02
24 to 29	26,630	97.3	1.02
30 or older	30,580	96.8	1.03
Student type (based on data collection)			
Undergraduate	115,330	97.1	1.02
Graduate, master's, doctoral	12,710	99.0	1.01
First-professional	1,600	98.9	1.01
Baccalaureate status (from sampling list)			
Yes	28,860	98.4	1.01
No	82,690	97.5	1.02
Don't know	18,090	96.3	1.03
Any aid recipient			
Yes	98,610	99.6	1.00
No	31,030	92.4	1.08
Federal aid recipient			
Yes	77,360	99.8	1.00
No	52,280	94.8	1.05
Institution aid recipient			
Yes	38,920	99.6	1.00
No	90,710	96.5	1.03
State aid recipient			
Yes	36,840	99.9	1.00
No	92,800	96.7	1.03

See notes at end of table.

Table 59. Weight adjustment factors for student location nonresponse adjustment: 2008—Continued

Model predictor variable	Number of located respondents	Weighted response rate	Average weight adjustment factor (WT8)
Pell Grant amount received			
Pell Grant amount = 0	83,960	96.5	1.03
0 < Pell Grant amount ≤ 1,616	11,530	99.8	1.00
1,616 < Pell Grant amount ≤ 2,840	11,330	99.7	1.00
2,840 < Pell Grant amount ≤ 4,309	10,420	99.6	1.00
Pell Grant amount = 4,310	12,390	99.7	1.00
Stafford Loan amount received			
Stafford Loan amount = 0	69,960	95.8	1.04
0 < Stafford Loan amount ≤ 3,500	16,180	99.8	1.00
3,500 < Stafford Loan amount ≤ 5,499	10,510	99.7	1.00
5,499 < Stafford Loan amount ≤ 7,510	18,070	99.8	1.00
7,510 < Stafford Loan amount	14,920	99.7	1.00
PLUS amount received			
PLUS amount = 0	122,850	97.2	1.02
0 < PLUS amount ≤ 5,000	1,710	99.7	1.00
5,000 < PLUS amount ≤ 9,799	1,680	99.9	1.00
9,799 < PLUS amount ≤ 15,000	1,710	99.9	1.00
15,000 < PLUS amount	1,690	100.0	1.00
In CPS			
Yes	81,550	100.0	1.00
No	48,090	94.1	1.06
In CADE			
Yes	124,320	98.0	1.02
No	5,310	85.6	1.15
In CATI			
Yes	95,360	100.0	1.00
No	34,270	90.9	1.08
Count of phone numbers			
None	2,940	89.6	1.10
1	64,960	96.7	1.03
2	47,740	98.5	1.01
More than 2	14,000	99.3	1.01
Count of e-mail addresses			
None	8,490	89.9	1.09
1	54,210	96.3	1.03
2	55,750	99.0	1.01
More than 2	11,180	99.9	1.00
Count of mailing addresses			
None	1,420	85.2	1.15
1	65,800	96.5	1.03
2	44,900	98.6	1.01
More than 2	17,530	99.4	1.01

See notes at end of table.

Table 59. Weight adjustment factors for student location nonresponse adjustment: 2008—Continued

Model predictor variable	Number of located respondents	Weighted response rate	Average weight adjustment factor (WT8)
Preloaded Social Security number			
Yes	116,630	97.7	1.02
No	13,010	94.1	1.06
CHAID segments			
In CADE but not in CATI or CPS	13,630	86.4	1.15
Not in CADE, CATI, or CPS	370	30.8	3.17
In CPS but not in CATI, public 2-year	3,200	100.0	1.00
In CPS but not in CATI, not [public 2-year; private not-for-profit 4-year-or-more; or private not-for-profit 2-year]	11,340	100.0	1.00
In CPS but not in CATI and [private not-for-profit 4-year-or-more; or private not-for-profit 2-year]	5,730	99.9	1.00
In CATI	95,360	100.0	1.00

[1] New England = Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont; Mid East = Delaware, District of Columbia, Maryland, New Jersey, New York, Pennsylvania; Great Lakes = Illinois, Indiana, Michigan, Ohio, Wisconsin; Plains = Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota; Southeast = Alabama, Arkansas, Florida, Georgia, Kentucky, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, West Virginia; Southwest = Arizona, New Mexico, Oklahoma, Texas; Rocky Mountains = Colorado, Idaho, Montana, Utah, Wyoming; Far West = California, Nevada, Oregon, Washington; Outlying Areas = Alaska, Hawaii, and Puerto Rico. Alaska and Hawaii were reclassified from the West to the outlying areas for the purposes of NPSAS.

NOTE: Numbers may not sum to totals due to rounding. Categories were formed from continuous variables based on quartiles. CADE = computer-assisted data entry; CATI = computer-assisted telephone interviewing; CHAID = chi-square automatic interaction detection; CPS = Central Processing System; PLUS = Parent Loan for Undergraduate Students.

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

Student Other Nonresponse Adjustment (WT9). The second stage of adjustment for student nonresponse was an adjustment for other study nonresponse, given that the student was located. This additional type of student nonresponse adjustment was made to further compensate for the potential student nonresponse bias. The same GEM procedure was used as in the adjustment for not locating students (WT8). Candidate predictor variables were the same as those used in the student location adjustments. As in the other nonresponse adjustment, a CHAID analysis was performed on the predictor variables to detect important interactions. The resulting segment interactions and all the main effect variables were then included in GEM.

Table 60 presents the final predictor variables used in GEM to adjust the student weights and the average weight adjustment factor resulting from these variables. The weight adjustment factors met the following constraints:

- minimum: 1.00;
- median: 1.00; and
- maximum: 5.31.

Table 60. Weight adjustment factors for student refusal and other nonresponse adjustment: 2008

Model predictor variable	Number of respondents	Weighted response rate	Average weight adjustment factor (WT9)
Total	127,700	97.3	1.02
Type of institution			
Public less-than-2-year	1,330	90.5	1.10
Public 2-year	33,430	95.0	1.03
Public 4-year non-doctorate-granting	15,990	98.9	1.01
Public 4-year doctorate-granting	29,530	98.4	1.01
Private not-for-profit less-than-4-year	1,600	98.0	1.02
Private not-for-profit 4-year non-doctorate-granting	13,130	97.8	1.02
Private not-for-profit 4-year doctorate-granting	16,460	98.7	1.01
Private for-profit less-than-2-year	5,620	98.4	1.01
Private for-profit 2-year-or-more	10,600	99.3	1.01
Institution state			
All other states	72,230	98.2	1.02
California public 2-year	4,070	87.7	1.10
California public 4-year	3,690	97.7	1.02
California private not-for-profit 4-year	1,810	94.4	1.05
California private for-profit degree-granting	1,020	99.7	1.00
Georgia public 2-year	3,880	98.2	1.01
Georgia public 4-year	2,080	99.7	1.00
Georgia private not-for-profit 4-year	1,470	99.6	1.00
Georgia private for-profit degree-granting	1,600	99.7	1.00
Illinois public 2-year	2,190	95.4	1.03
Illinois public 4-year	1,780	99.3	1.01
Illinois private not-for-profit 4-year	2,150	99.7	1.00
Illinois private for-profit degree-granting	1,410	100.0	1.00
Minnesota public 2-year	4,190	98.3	1.01
Minnesota public 4-year	2,200	98.3	1.02
Minnesota private not-for-profit 4-year	1,280	99.7	1.00
Minnesota private for-profit degree-granting	1,260	95.6	1.03
New York public 2-year	3,960	98.9	1.01
New York public 4-year	3,090	99.4	1.01
New York private not-for-profit 4-year	3,330	99.3	1.01
New York private for-profit degree-granting	1,410	97.0	1.02
Texas public 2-year	2,100	97.2	1.02
Texas public 4-year	3,010	97.7	1.02
Texas private not-for-profit 4-year	1,240	99.5	1.00
Texas private for-profit degree-granting	1,270	99.5	1.00

See notes at end of table.

Table 60. Weight adjustment factors for student refusal and other nonresponse adjustment: 2008—Continued

Model predictor variable	Number of respondents	Weighted response rate	Average weight adjustment factor (WT9)
Region[1]			
New England	5,160	98.3	1.01
Mid East	22,380	97.8	1.02
Great Lakes	19,390	98.3	1.01
Plains	15,680	98.6	1.01
Southeast	29,900	98.7	1.01
Southwest	12,240	97.6	1.02
Rocky Mountains	4,190	98.7	1.01
Far West	15,840	92.6	1.05
Outlying Areas	2,920	99.2	1.01
Institution enrollment size			
0 < enrollment total ≤ 3,856	32,110	98.9	1.01
3,856 < enrollment total ≤ 10,698	31,930	97.5	1.02
10,698 < enrollment total ≤ 24,146	31,690	96.5	1.02
24,146 < enrollment total	31,980	97.1	1.02
Education level (based on sampling list)			
Undergraduate	114,070	97.1	1.02
Graduate, master's, doctoral	11,620	98.6	1.01
First-professional	2,010	98.7	1.01
Age group			
15 to 23	71,620	98.1	1.01
24 to 29	26,250	97.6	1.02
30 or older	29,830	95.7	1.03
Student type (based on data collection)			
Undergraduate	113,540	97.1	1.02
Graduate, master's, doctoral	12,590	98.7	1.01
First-professional	1,580	98.4	1.02
Baccalaureate status (from sampling list)			
Yes	28,550	98.5	1.01
No	81,460	97.4	1.02
Don't know	17,700	96.5	1.02
Any aid recipient			
Yes	98,390	99.7	1.00
No	29,320	92.1	1.07
Federal aid recipient			
Yes	77,240	99.8	1.00
No	50,470	94.8	1.04
Institution aid recipient			
Yes	38,870	99.8	1.00
No	88,830	96.4	1.03
State aid recipient			
Yes	36,820	99.9	1.00
No	90,880	96.6	1.02

See notes at end of table.

Table 60. Weight adjustment factors for student refusal and other nonresponse adjustment: 2008—Continued

Model predictor variable	Number of respondents	Weighted response rate	Average weight adjustment factor (WT9)
Pell Grant amount received			
Pell Grant amount = 0	82,090	96.4	1.03
0 < Pell Grant amount ≤ 1,616	11,520	99.9	1.00
1,616 < Pell Grant amount ≤ 2,840	11,310	99.8	1.00
2,840 < Pell Grant amount ≤ 4,309	10,410	99.8	1.00
Pell Grant amount = 4,310	12,380	99.9	1.00
Stafford Loan amount received			
Stafford Loan amount = 0	68,130	95.8	1.03
0 < Stafford Loan amount ≤ 3,500	16,140	99.8	1.00
3,500 < Stafford Loan amount ≤ 5,499	10,490	99.8	1.00
5,499 < Stafford Loan amount ≤ 7,510	18,040	99.9	1.00
7,510 < Stafford Loan amount	14,900	99.7	1.00
PLUS amount received			
PLUS amount = 0	120,930	97.2	1.02
0 < PLUS amount ≤ 5,000	1,700	99.5	1.00
5,000 < PLUS amount ≤ 9,799	1,680	99.9	1.00
9,799 < PLUS amount ≤ 15,000	1,700	99.9	1.00
15,000 < PLUS amount	1,690	99.9	1.00
In CPS			
Yes	81,550	100.0	1.00
No	48,090	94.1	1.06
In CADE			
Yes	124,320	98.0	1.02
No	5,310	85.6	1.15
In CATI			
Yes	95,360	100.0	1.00
No	34,270	90.9	1.08
Count of phone numbers			
None	2,940	89.6	1.10
1	64,960	96.7	1.03
2	47,740	98.5	1.01
More than 2	14,000	99.3	1.01
Count of e-mail addresses			
None	8,110	90.3	1.06
1	53,010	96.2	1.03
2	55,420	99.1	1.01
More than 2	11,160	99.7	1.00
Count of mailing addresses			
None	1,370	93.0	1.06
1	64,410	96.4	1.03
2	44,490	98.5	1.01
More than 2	17,430	99.2	1.01

See notes at end of table.

Table 60. Weight adjustment factors for student refusal and other nonresponse adjustment: 2008—Continued

Model predictor variable	Number of respondents	Weighted response rate	Average weight adjustment factor (WT9)
Preloaded SSN			
Yes	115,200	97.7	1.01
No	12,500	94.6	1.05
CHAID segments			
Received aid, SSN not preloaded, and has 0 e-mail addresses	430	95.7	1.03
Received aid, SSN not preloaded, and has 1 e-mail address	4,320	97.9	1.02
Received aid, SSN not preloaded, and has 2 e-mail addresses	2,470	99.5	1.00
Received aid, SSN not preloaded, and has 3 or more e-mail addresses	240	100.0	1.00
Received aid, SSN preloaded, and Pell Grant amount = 0	47,440	99.8	1.00
Received aid, SSN preloaded, and 0 < Pell Grant amount ≤ 4,000	31,660	100.0	1.00
Received aid, SSN preloaded, and Pell Grant amount > 4,000	11,810	100.0	1.00
Received no aid, Southeast or Rocky Mountains region or Outlying Areas, and public 4-year-or-more	190	98.9	1.01
Received no aid, Southeast or Rocky Mountains region or Outlying Areas, and private not-for-profit 4-year-or-more, private for-profit 4-year-or-more, public 2-year, or private not-for-profit 2-year	5,780	96.1	1.04
Received no aid, Southeast or Rocky Mountains region or Outlying Areas, and private for-profit 2-year	410	82.3	1.16
Received no aid, Southeast or Rocky Mountains region or Outlying Areas, and public less-than-2-year	530	96.7	1.04
Received no aid, Southeast or Rocky Mountains region or Outlying Areas, and private not-for-profit less-than-2-year	130	91.1	1.09
Received no aid, Southeast or Rocky Mountains region or Outlying Areas, and private for-profit less-than-2-year	210	99.9	1.00
Received no aid, New England, Mid East, or Southwest region, and private not-for-profit 4-year-or-more	3,970	92.6	1.06
Received no aid, New England, Mid East, or Southwest region, and private for-profit 4-year-or-more	1,230	95.8	1.04
Received no aid, New England, Mid East, or Southwest region, and public 2-year	1,930	93.1	1.06
Received no aid, New England, Mid East, or Southwest region, and private not-for-profit 2-year	60	78.6	1.25
Received no aid, New England, Mid East, or Southwest region, and private for-profit 2-year or public less-than-2-year	1,790	94.8	1.05

See notes at end of table.

Table 60. Weight adjustment factors for student refusal and other nonresponse adjustment: 2008—Continued

Model predictor variable	Number of respondents	Weighted response rate	Average weight adjustment factor (WT9)
CHAID segments—Continued			
Received no aid, New England, Mid East, or Southwest region, and private not-for-profit less-than-2-year	140	85.0	1.15
Received no aid, New England, Mid East, or Southwest region, and private for-profit less-than-2-year	300	98.2	1.06
Received no aid, Far West region, and public 4-year-or-more	180	57.6	1.74
Received no aid, Far West region, and private not-for-profit 4-year-or-more	2,410	81.0	1.19
Received no aid, Far West region, and private for-profit 4-year-or-more	740	97.0	1.03
Received no aid, Far West region, and public 2-year	980	92.3	1.07
Received no aid, Far West region, and private not-for-profit 2-year or private for-profit 2-year	180	74.4	1.30
Received no aid, Far West region, and public less-than-2-year, private not-for-profit less-than-2-year, or private for-profit less-than-2-year	580	94.4	1.06
Received no aid, Mid East or Great Lakes region, and public 4-year-or-more	60	100.0	1.00
Received no aid, Mid East or Great Lakes region, and private not-for-profit 4-year-or-more	3,390	94.3	1.05
Received no aid, Mid East or Great Lakes region, and private for-profit 4-year-or-more, public 2 year, or private not-for-profit 2-year	2,510	98.2	1.02
Received no aid, Mid East or Great Lakes region, and private for-profit 2-year	670	89.3	1.13
Received no aid, Mid East or Great Lakes region, and public less-than-2-year	640	93.7	1.08
Received no aid, Mid East or Great Lakes region, and private not-for-profit less-than-2-year or private for-profit less-than-2-year	310	82.6	1.25

[1] New England = Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont; Mid East = Delaware, District of Columbia, Maryland, New Jersey, New York, Pennsylvania; Great Lakes = Illinois, Indiana, Michigan, Ohio, Wisconsin; Plains = Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota; Southeast = Alabama, Arkansas, Florida, Georgia, Kentucky, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, West Virginia; Southwest = Arizona, New Mexico, Oklahoma, Texas; Rocky Mountains = Colorado, Idaho, Montana, Utah, Wyoming; Far West = California, Nevada, Oregon, Washington; Outlying Areas = Alaska, Hawaii, and Puerto Rico. Alaska and Hawaii were reclassified from the West to the outlying areas for the purposes of NPSAS.

NOTE: Numbers may not sum to totals due to rounding. Categories were formed from continuous variables based on quartiles. CADE = computer-assisted data entry; CATI = computer-assisted telephone interviewing; CHAID = chi-square automatic interaction detection; CPS = Central Processing System; PLUS = Parent Loan for Undergraduate Students; SSN = Social Security number.

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

Student Poststratification Adjustment (WT10). To ensure population coverage, the student weights were further adjusted, with the use of GEM, to known population control totals (control totals) for key variables. The random sample of students may have had a distribution that differed from the population distribution; poststratification is a method to reduce the standard errors by adjusting estimates to external data. Control totals were established for the following:

- number of undergraduate subsidized and unsubsidized Stafford Loan recipients, by institution type within six representative states;
- number of undergraduate subsidized and unsubsidized Stafford Loan recipients, by institution type within remaining states;
- number of graduate/first-professional subsidized and unsubsidized Stafford Loan recipients, by institution type;
- total Stafford Loan amounts disbursed to undergraduate students;
- total Stafford Loan amounts disbursed to graduate/first-professional students;
- Pell Grant amounts awarded, by institution type within six representative states;
- Pell Grant amounts awarded, by institution type within remaining states;
- PLUS amounts disbursed to graduate/first-professional students, by institution type;
- state grant amounts awarded to undergraduate students, by institution type within six representative states;
- number of Science and Mathematics Access to Retain Talent (SMART) Grant recipients;
- number of Academic Competitiveness Grant (ACG) recipients;
- fall undergraduate student enrollment, by institution type within six representative states;
- fall undergraduate student enrollment, by institution type within remaining states;
- fall graduate/first-professional student enrollment, by institution type;
- non–fall undergraduate student enrollment, by institution type within six representative states;
- non–fall undergraduate student enrollment, by institution type within remaining states; and
- non–fall graduate/first-professional student enrollment, by institution type.

The PLUS, Stafford Loan, Pell Grant, SMART Grant, and ACG control totals were obtained from the U.S. Department of Education. The fall enrollment counts were obtained from the 2008 IPEDS Fall Enrollment Survey, and the non–fall enrollment counts were derived from the same source (using the 12 month unduplicated headcount). The state grant control totals were obtained from state websites. In the same step, high-extreme weights were poststratified to the control totals, truncated, and smoothed by GEM, while the other weights were poststratified to the control totals.

Stafford loans, for which we have several control totals, is the largest single loan program—in terms of the number of students affected as well as the dollars involved. Therefore, having accurate data on Stafford loans by loan type, institution type, and graduate/undergraduate level is crucial for weighting Stafford borrowing in the survey. Prior to NPSAS:08, the dollars used for poststratifying student weights were the gross loan commitments–the amounts that schools and lenders expect to award to students based on their loan applications–collected by the Department of Education (ED). For NPSAS:08, net disbursements–the amounts that the students actually receive–

were used for poststratification instead because they more accurately reflect the amount of money students are actually borrowing. For 2007-2008 the overall difference between these two measures, gross and net, is about 18 percent, or $10 billion.

When NPSAS began using Stafford loan totals for poststratification in 1995-1996, a decision was made to use gross commitment data (rather than net disbursement data) for poststratification, based on the view that gross commitment data were (a) timely; (b) widely used; and (c) an accurate measure of what students actually borrow and apply to schooling costs.

Gross commitment dollars are available near the beginning of the lending process and are a very important early gauge of how much capital lenders will need and how much money is expected to flow from lenders to schools. In the process of distributing loan funds, there is some drop-off in amounts as students change their minds about whether or how much to borrow, decide not to enroll or drop out, forget to pick up their checks, or a number of other reasons. Once final loan amounts have been distributed, the amount that has been disbursed is less than originally posted.

Net disbursement amounts are available only after accounts between schools, lenders, and ED have all been reconciled. In 1995-1996 this number was not available to ED (and NPSAS) until an annual reconciliation of accounts took place. Gross commitment data were used because they permitted NPSAS data to reach policymakers only seven months after the June 30th close of the federal student aid year. To ensure that NPSAS is relevant to federal policymaking, NCES Postsecondary Studies has assigned high priority to releasing NPSAS on a timely basis, targeting its quadrennial release to support the work of each incoming (or, returning) administration. In addition, gross commitments were the standard for reporting on lending totals: the student loan industry and schools typically used the gross commitment numbers to indicate loan volume. Moreover, gross commitments and disbursements were not widely different from one another.

A new Federal loan data collection system was implemented in 1994. In the years following disbursement data gradually became more swiftly available and usable. ED now collects these data weekly through electronic means, and the net disbursement data are more readily available after the award period ends. Lenders and others now publish both gross commitments and net disbursements, and disbursement totals are now understood to be the best measure of actual student borrowing (e.g. In 2008, The College Board shifted from reporting commitments to disbursements in its annual publication, *Trends in Student Aid* (2008). Most importantly, the gap between commitments and disbursements has gradually been growing, reaching 17.6 percent for 2007-2008. In light of these changes, the NPSAS:08 poststratification control totals are based on Stafford loan disbursement totals.

To ensure consistency in methodology and comparability of findings, which is especially important for trend analysis, data from prior cycles since NPSAS:96 have been reweighted, using Stafford loan disbursement totals for the poststratification process.

After poststratification, weighted estimates for key variables were compared with other estimates, such as financial aid estimates published by the College Board (2008) and estimates from NPSAS:04. These comparisons showed that the NPSAS:08 estimates were reasonable.

Table 61 presents the variables associated with the control totals and the average weight adjustment factors by these variables. The weight adjustment factors from GEM are summarized below and met the following constraints:

- minimum: 0.01;

- median: 1.21; and
- maximum: 161.12.

After this last weight adjustment was performed, the final student weight (study weight or WTA000) was computed as the product of the 10 weight components described in this section and in section 6.1.

Table 61. Weight adjustment factors for student poststratification: 2008

Model predictor variable	Control total	Average weight adjustment factor (WT10)
Undergraduate fall enrollment, by institution type for representative states		
Undergraduate fall enrollment, California, public 2-year	1,494,349	1.26
Undergraduate fall enrollment, California, 4-year non-doctorate-granting	185,028	1.04
Undergraduate fall enrollment, California, 4-year doctorate-granting	344,275	1.53
Undergraduate fall enrollment, California, private not-for-profit 4-year non-doctorate-granting	50,756	1.82
Undergraduate fall enrollment California, private not-for-profit 4-year doctorate-granting	88,940	1.43
Undergraduate fall enrollment, California private for-profit 2-year-or-more	103,013	0.90
Undergraduate fall enrollment, Georgia, public 2-year	137,872	1.00
Undergraduate fall enrollment, Georgia, 4-year non-doctorate-granting	74,753	1.17
Undergraduate fall enrollment, Georgia, 4-year doctorate-granting	110,080	1.11
Undergraduate fall enrollment, Georgia, private not-for-profit 4-year non-doctorate-granting	32,059	1.05
Undergraduate fall enrollment, Georgia, private not-for-profit 4-year doctorate-granting	15,595	1.48
Undergraduate fall enrollment, Georgia, private for-profit 2-year-or-more	22,588	1.22
Undergraduate fall enrollment, Illinois, public 2-year	347,277	1.21
Undergraduate fall enrollment, Illinois, 4-year non-doctorate-granting	20,695	1.04
Undergraduate fall enrollment, Illinois, 4-year doctorate-granting	132,043	1.12
Undergraduate fall enrollment, Illinois, private not-for-profit 4-year non-doctorate-granting	61,752	1.07
Undergraduate fall enrollment, Illinois, private not-for-profit 4-year doctorate-granting	73,786	0.95
Undergraduate fall enrollment, Illinois, private for-profit 2-year-or-more	54,962	1.06
Undergraduate fall enrollment, Minnesota, public 2-year	119,569	1.17

See notes at end of table.

Table 61. Weight adjustment factors for student poststratification: 2008—Continued

Model predictor variable	Control total	Average weight adjustment factor (WT10)
Undergraduate fall enrollment, by institution type for representative states—Continued		
Undergraduate fall enrollment, Minnesota, 4-year non-doctorate-granting	30,203	0.46
Undergraduate fall enrollment, Minnesota, 4-year doctorate-granting	75,008	2.53
Undergraduate fall enrollment, Minnesota, private not-for-profit 4-year non-doctorate-granting	28,040	1.33
Undergraduate fall enrollment, Minnesota, private not-for-profit 4-year doctorate-granting	22,913	1.49
Undergraduate fall enrollment, Minnesota private for-profit 2-year-or-more	22,682	1.24
Undergraduate fall enrollment, New York, public 2-year	281,498	1.13
Undergraduate fall enrollment, New York, 4-year non-doctorate-granting	241,599	1.24
Undergraduate fall enrollment, New York, 4-year doctorate-granting	61,611	0.90
Undergraduate fall enrollment, New York, private not-for-profit 4-year non-doctorate-granting	142,460	1.16
Undergraduate fall enrollment, New York, private not-for-profit 4-year doctorate-granting	197,694	1.03
Undergraduate fall enrollment, New York, private for-profit 2-year-or-more	41,594	1.10
Undergraduate fall enrollment, Texas, public 2-year	556,197	1.20
Undergraduate fall enrollment, Texas, 4-year non-doctorate-granting	66,040	0.90
Undergraduate fall enrollment, Texas, 4-year doctorate-granting	360,373	1.22
Undergraduate fall enrollment, Texas, private not-for-profit 4-year non-doctorate-granting	35,412	1.32
Undergraduate fall enrollment, Texas, private, not-for-profit 4-year doctorate-granting	57,547	1.19
Undergraduate fall enrollment, Texas, private for-profit 2-year-or-more	30,620	0.81
Undergraduate fall enrollment, by institution type for nonrepresentative states		
Undergraduate fall enrollment, public less-than-2-year	68,807	2.49
Undergraduate fall enrollment, public 2-year	3,449,758	1.15
Undergraduate fall enrollment, 4-year non-doctorate-granting	1,323,317	1.28
Undergraduate fall enrollment, 4-year doctorate-granting	2,818,355	1.14
Undergraduate fall enrollment, private not-for-profit less-than-4-year	66,948	0.80
Undergraduate fall enrollment, private not-for-profit 4-year non-doctorate-granting	988,236	1.07
Undergraduate fall enrollment, private not-for-profit 4-year doctorate-granting	784,200	1.24
Undergraduate fall enrollment, private for-profit less-than-2-year	256,955	1.18
Undergraduate fall enrollment, private for-profit 2-year-or-more	800,377	1.05
Graduate/first-professional fall enrollment, by institution type for all states		
Graduate/first-professional fall enrollment, 4-year non-doctorate-granting	211,226	1.08
Graduate/first-professional fall enrollment, 4-year doctorate-granting	1,147,729	1.18
Graduate/first-professional fall enrollment, private not-for-profit 4-year non-doctorate-granting	211,061	1.06
Graduate/first-professional fall enrollment, private not-for-profit 4-year doctorate-granting	923,469	1.19
Graduate/first-professional fall enrollment, private for-profit 2-year-or-more	192,340	0.98

See notes at end of table.

Table 61. Weight adjustment factors for student poststratification: 2008—Continued

Model predictor variable	Control total	Average weight adjustment factor (WT10)
Undergraduate non–fall enrollment, by institution type for representative states		
Undergraduate non–fall enrollment, California, public 2-year	544,376	1.84
Undergraduate non–fall enrollment, California, 4-year non-doctorate-granting	11,031	1.48
Undergraduate non–fall enrollment, California, 4-year doctorate-granting	24,565	3.89
Undergraduate non–fall enrollment, California, private not-for-profit 4-year non-doctorate-granting	7,991	3.03
Undergraduate non–fall enrollment, California, private not-for-profit 4-year doctorate-granting	8,257	2.60
Undergraduate non–fall enrollment, California, private for-profit 2-year-or-more	49,579	2.05
Undergraduate non–fall enrollment, Georgia, public 2-year	67,090	4.93
Undergraduate non–fall enrollment, Georgia, 4-year non-doctorate-granting	14,146	3.75
Undergraduate non–fall enrollment, Georgia, 4-year doctorate-granting	13,870	4.36
Undergraduate non–fall enrollment, Georgia, private not-for-profit 4-year non-doctorate-granting	2,564	1.39
Undergraduate non–fall enrollment, Georgia, private not-for-profit 4-year doctorate-granting	1,503	3.46
Undergraduate non–fall enrollment, Georgia, private for-profit 2-year-or-more	11,229	3.20
Undergraduate non–fall enrollment, Illinois, public 2-year	265,591	3.67
Undergraduate non–fall enrollment, Illinois, 4-year non-doctorate-granting	1,439	1.87
Undergraduate non–fall enrollment, Illinois, 4-year doctorate-granting	11,067	1.62
Undergraduate non–fall enrollment, Illinois, private not-for-profit 4-year non-doctorate-granting	6,798	1.35
Undergraduate non–fall enrollment, Illinois, private not-for-profit 4-year doctorate-granting	8,909	2.93
Undergraduate non–fall enrollment, Illinois, private for-profit 2-year-or-more	29,662	1.15
Undergraduate non–fall enrollment, Minnesota, public 2-year	35,996	1.51
Undergraduate non–fall enrollment, Minnesota, 4-year non-doctorate-granting	5,419	0.94
Undergraduate non–fall enrollment, Minnesota, private not-for-profit 4-year doctorate-granting	1,595	1.66
Undergraduate non–fall enrollment, Minnesota, private for-profit 2-year-or-more	8,534	2.41
Undergraduate non–fall enrollment, New York, public 2-year	99,901	2.56
Undergraduate non–fall enrollment, New York, 4-year non-doctorate-granting	44,003	2.06
Undergraduate non–fall enrollment, New York, 4-year doctorate-granting	5,693	2.64
Undergraduate non–fall enrollment, New York, private not-for-profit 4-year non-doctorate-granting	17,289	4.31
Undergraduate non–fall enrollment, New York, private not-for-profit 4-year doctorate-granting	33,150	5.47
Undergraduate non–fall enrollment, New York, private for-profit 2-year-or-more	20,831	2.13
Undergraduate non–fall enrollment, Texas, public 2-year	235,540	2.33
Undergraduate non–fall enrollment, Texas, 4-year non-doctorate-granting	13,522	1.08
Undergraduate non–fall enrollment, Texas, 4-year doctorate-granting	42,694	2.12
Undergraduate non–fall enrollment, Texas, private not-for-profit 4-year non-doctorate-granting	7,859	2.19
Undergraduate non–fall enrollment, Texas, private not-for-profit 4-year doctorate-granting	6,194	2.71
Undergraduate non–fall enrollment, Texas, private for-profit 2-year-or-more	17,374	2.80

See notes at end of table.

Table 61. Weight adjustment factors for student poststratification: 2008—Continued

Model predictor variable	Control total	Average weight adjustment factor (WT10)
Undergraduate non–fall enrollment, by institution type for nonrepresentative states		
Undergraduate non–fall enrollment, public less-than-2-year	23,721	1.94
Undergraduate non–fall enrollment, public 2-year	1,476,769	2.27
Undergraduate non–fall enrollment, public 4-year non-doctorate-granting	311,949	2.41
Undergraduate non–fall enrollment, public 4-year doctorate-granting	346,291	2.07
Undergraduate non–fall enrollment, private not-for-profit less than 4-year	23,155	1.13
Undergraduate non–fall enrollment, private not-for-profit 4-year non-doctorate-granting	173,000	2.12
Undergraduate non–fall enrollment, private not-for-profit 4-year doctorate-granting	93,733	2.52
Undergraduate non–fall enrollment, private for-profit less-than-2-year	188,057	2.12
Undergraduate non–fall enrollment, private for-profit 2-year-or-more	337,267	2.16
Graduate/first-professional non–fall enrollment, by institution type for all states		
Graduate/first-professional non–fall enrollment, public 4-year non-doctorate-granting	113,271	1.72
Graduate/first-professional non–fall enrollment, public 4-year doctorate-granting	277,312	1.77
Graduate/first-professional non–fall enrollment, private not-for-profit 4-year non-doctorate-granting	89,638	1.53
Graduate/first-professional non–fall enrollment, private not-for-profit 4-year doctorate-granting	229,769	2.00
Graduate/first-professional non–fall enrollment, private for-profit 2-year-or-more	60,440	4.30
Amount of Pell Grants awarded, by institution type for representative states		
Aggregate Pell dollars California, public 2-year	671,000,000	0.97
Aggregate Pell dollars California, public 4-year, University of California	145,200,000	1.46
Aggregate Pell dollars California, public 4-year, California state universities and colleges	390,200,000	1.23
Aggregate Pell dollars California, private not-for-profit 4-year	88,000,000	1.74
Aggregate Pell dollars California, private for-profit degree-granting	238,100,000	1.21
Aggregate Pell dollars Georgia, public 2-year, University System of Georgia	48,500,000	1.06
Aggregate Pell dollars Georgia, public 2-year, Technical College System of Georgia and other 2-year colleges	71,800,000	1.16
Aggregate Pell dollars Georgia, public, 4-year, University System of Georgia	162,300,000	1.18
Aggregate Pell dollars Georgia, private not-for-profit 4-year	44,900,000	1.32
Aggregate Pell dollars Georgia, private for-profit degree-granting	102,600,000	2.30
Aggregate Pell dollars Illinois, public 2-year	190,300,000	1.15
Aggregate Pell dollars Illinois, public 4-year	119,700,000	1.06
Aggregate Pell dollars Illinois, private 4-year not-for-profit	102,600,000	0.97
Aggregate Pell dollars Illinois, private for-profit degree-granting	74,900,000	0.93
Aggregate Pell dollars Minnesota, public 2-year	92,200,000	1.05
Aggregate Pell dollars Minnesota, public 4-year	59,100,000	0.54
Aggregate Pell dollars Minnesota, private 4-year not-for-profit	29,200,000	1.26
Aggregate Pell dollars Minnesota, private for-profit degree-granting	26,900,000	1.42

See notes at end of table.

Table 61. Weight adjustment factors for student poststratification: 2008—Continued

Model predictor variable	Control total	Average weight adjustment factor (WT10)
Amount of Pell Grants awarded, by institution type for representative states—Continued		
Aggregate Pell dollars New York, public 2-year, SUNY	160,100,000	1.09
Aggregate Pell dollars New York, public 2-year, CUNY	112,000,000	1.15
Aggregate Pell dollars New York, public 4-year, SUNY	159,200,000	1.29
Aggregate Pell dollars New York, public 4-year, CUNY	169,300,000	0.80
Aggregate Pell dollars New York, private 4-year not-for-profit	250,800,000	1.05
Aggregate Pell dollars New York, private for-profit degree-granting	124,800,000	1.27
Aggregate Pell dollars Texas, public 2-year	409,300,000	1.05
Aggregate Pell dollars Texas, public 4-year	430,200,000	1.10
Aggregate Pell dollars Texas, private 4-year not-for-profit	71,400,000	1.19
Aggregate Pell dollars Texas, private for-profit degree-granting	70,300,000	0.85
Amount of Pell Grants awarded, by institution type for nonrepresentative states		
Aggregate Pell dollars, public less than 2-year	79,900,000	1.98
Aggregate Pell dollars, public 2-year	2,646,100,000	1.00
Aggregate Pell dollars, public 4-year	3,179,200,000	1.10
Aggregate Pell dollars, private not-for-profit less than 4-year	109,400,000	0.67
Aggregate Pell dollars, private not-for-profit 4-year	1,565,400,000	1.12
Aggregate Pell dollars, private for-profit less-than-2-year	751,400,000	1.30
Aggregate Pell dollars, private for-profit 2-year-or-more	1,702,300,000	1.42
Amount of Stafford Loans awarded		
Undergraduate Stafford Loan amount	35,790,000,000	1.19
Graduate/first-professional Stafford Loan amount	21,004,000,000	1.17
Number of undergraduate students awarded Stafford Loans, by institution type for representative states		
Undergraduate subsidized Stafford Loan recipients, California, public 2-year	36,120	0.64
Undergraduate subsidized Stafford Loan recipients, California, public 4-year	172,249	1.32
Undergraduate subsidized Stafford Loan recipients, California, private not-for-profit 4-year	66,166	1.71
Undergraduate unsubsidized Stafford Loan recipients, California, public 2-year	17,719	0.47
Undergraduate unsubsidized Stafford Loan recipients, California, public 4-year	84,018	1.25
Undergraduate unsubsidized Stafford Loan recipients, California, private not-for-profit 4-year	40,459	1.76
Undergraduate subsidized Stafford Loan recipients, Georgia, public 2-year	13,825	1.28
Undergraduate subsidized Stafford Loan recipients, Georgia, public 4-year	72,812	1.26
Undergraduate subsidized Stafford Loan recipients, Georgia, private not-for-profit 4-year	27,941	1.12
Undergraduate unsubsidized Stafford Loan recipients, Georgia, public 2-year	9,159	0.97
Undergraduate unsubsidized Stafford Loan recipients, Georgia, public 4-year	56,358	1.18

See notes at end of table.

Table 61. Weight adjustment factors for student poststratification: 2008—Continued

Model predictor variable	Control total	Average weight adjustment factor (WT10)
Number of undergraduate students awarded Stafford Loans, by institution type for representative states—Continued		
Undergraduate unsubsidized Stafford Loan recipients, Georgia, private not-for-profit 4-year	19,481	1.30
Undergraduate subsidized Stafford Loan recipients, Illinois, public 2-year	20,363	0.66
Undergraduate subsidized Stafford Loan recipients, Illinois, public 4-year	64,763	1.18
Undergraduate subsidized Stafford Loan recipients, Illinois, private not-for-profit 4-year	77,136	1.00
Undergraduate unsubsidized Stafford Loan recipients, Illinois, public 2-year	13,113	0.77
Undergraduate unsubsidized Stafford Loan recipients, Illinois, public 4-year	39,171	1.13
Undergraduate unsubsidized Stafford Loan recipients, Illinois, private not-for-profit 4-year	44,628	1.08
Undergraduate subsidized Stafford Loan recipients, Minnesota, public 2-year	47,378	1.07
Undergraduate subsidized Stafford Loan recipients, Minnesota, public 4-year	43,062	0.76
Undergraduate subsidized Stafford Loan recipients, Minnesota, private not-for-profit 4-year	27,506	1.21
Undergraduate unsubsidized Stafford Loan recipients, Minnesota, public 2-year	33,602	1.08
Undergraduate unsubsidized Stafford Loan recipients, Minnesota, public 4-year	33,027	0.57
Undergraduate unsubsidized Stafford Loan recipients, Minnesota, private not-for-profit 4-year	16,505	1.34
Undergraduate subsidized Stafford Loan recipients, New York, public 2-year, SUNY	42,479	0.97
Undergraduate subsidized Stafford Loan recipients, New York, public 2-year, CUNY	4,720	0.75
Undergraduate subsidized Stafford Loan recipients, New York, public 4-year, SUNY	81,019	1.18
Undergraduate subsidized Stafford Loan recipients, New York, public 4-year, CUNY	14,298	0.68
Undergraduate subsidized Stafford Loan recipients, New York, private not-for-profit 4-year	151,322	0.94
Undergraduate unsubsidized Stafford Loan recipients, New York, public 2-year, SUNY	32,896	0.89
Undergraduate unsubsidized Stafford Loan recipients, New York, public 2-year, CUNY	3,655	0.95
Undergraduate unsubsidized Stafford Loan recipients, New York, public 4-year, SUNY	55,905	1.63
Undergraduate unsubsidized Stafford Loan recipients, New York, public 4-year, CUNY	9,866	1.07
Undergraduate unsubsidized Stafford Loan recipients, New York, private not-for-profit 4-year	75,430	1.13
Undergraduate subsidized Stafford Loan recipients, Texas, public 2-year	65,651	0.93
Undergraduate subsidized Stafford Loan recipients, Texas, public 4-year	175,569	1.21
Undergraduate subsidized Stafford Loan recipients, Texas, private not-for-profit 4-year	44,575	1.16
Undergraduate unsubsidized Stafford Loan recipients, Texas, public 2-year	41,960	0.97

See notes at end of table.

Table 61. Weight adjustment factors for student poststratification: 2008—Continued

Model predictor variable	Control total	Average weight adjustment factor (WT10)
Number of undergraduate students awarded Stafford Loans, by institution type for representative states—Continued		
Undergraduate unsubsidized Stafford Loan recipients, Texas, public 4-year	126,160	1.23
Undergraduate unsubsidized Stafford Loan recipients, Texas, private not-for-profit 4-year	34,161	1.47
Number of undergraduate students awarded Stafford Loans, by institution type for nonrepresentative states		
Undergraduate subsidized Stafford Loan recipients, public 2-year	607,118	0.92
Undergraduate subsidized Stafford Loan recipients, public 4-year	1,623,233	1.16
Undergraduate subsidized Stafford Loan recipients, private not-for-profit 4-year	958,852	1.11
Undergraduate unsubsidized Stafford Loan recipients, public 2-year	432,720	0.80
Undergraduate unsubsidized Stafford Loan recipients, public 4-year	1,177,656	1.14
Undergraduate unsubsidized Stafford Loan recipients, private not-for-profit 4-year	598,367	1.26
Number of undergraduate students awarded Stafford Loans, by institution type for all states		
Undergraduate subsidized Stafford Loan recipients, public less-than-2-year	12,803	1.96
Undergraduate subsidized Stafford Loan recipients, private not-for-profit less-than-4-year	34,686	0.78
Undergraduate subsidized Stafford Loan recipients, private for-profit less-than-2-year	286,218	1.59
Undergraduate subsidized Stafford Loan recipients, private for-profit 2-year-or-more	1,444,223	1.08
Undergraduate unsubsidized Stafford Loan recipients, public less-than-2-year	10,650	1.12
Undergraduate unsubsidized Stafford Loan recipients, private not-for-profit less-than-4-year	27,470	1.10
Undergraduate unsubsidized Stafford Loan recipients, private for-profit less-than-2-year	247,287	1.40
Undergraduate unsubsidized Stafford Loan recipients, private for-profit 2-year-or-more	1,312,244	1.57
Number of graduate/first-professional students awarded Stafford Loans, by institution type for all states		
Graduate/first-professional subsidized Stafford Loan recipients, all states, public 4-year	496,855	1.21
Graduate/first-professional subsidized Stafford Loan recipients, all states, private not-for-profit 4-year	557,862	1.20
Graduate/first-professional subsidized Stafford Loan recipients, all states, private for-profit 2-year-or-more	172,732	1.60
Graduate/first-professional unsubsidized Stafford Loan recipients, all states, public 4-year	435,947	1.06
Graduate/first-professional unsubsidized Stafford Loan recipients, all states, private not-for-profit 4-year	543,238	1.19
Graduate/first-professional unsubsidized Stafford Loan recipients, all states, private for-profit 2-year-or-more	182,880	1.74

See notes at end of table.

Table 61. Weight adjustment factors for student poststratification: 2008—Continued

Model predictor variable	Control total	Average weight adjustment factor (WT10)
Amount of PLUS awarded to graduate/first-professionals, by institution type for all states		
Graduate/first-professional PLUS Amount, all states, public 4-year non-doctorate-granting	5,400,000	0.55
Graduate/first-professional PLUS Amount, all states, public 4-year doctorate-granting	530,900,000	0.94
Graduate/first-professional PLUS Amount, all states, private not-for-profit 4-year non-doctorate-granting	105,200,000	1.08
Graduate/first-professional PLUS Amount, all states, private not-for-profit 4-year doctorate-granting	1,885,800,000	1.00
Graduate/first-professional PLUS Amount, all states, private for-profit 2-year-or-more	82,100,000	0.98
Amount of state grants awarded to undergraduates, by institution type for representative states		
Undergraduate state grant amount, California, public, community colleges	103,000,000	1.08
Undergraduate state grant amount, California, public 4-year, University of California	337,000,000	1.68
Undergraduate state grant amount, California, public 4-year, California state universities and colleges	223,000,000	1.32
Undergraduate state grant amount, California, private not-for-profit, independent colleges and universities	235,000,000	1.83
Undergraduate state grant amount, Minnesota, public 2-year	36,000,000	1.09
Undergraduate state grant amount, Minnesota, public 4-year	61,000,000	0.60
Undergraduate state grant amount, Minnesota, private not-for-profit 4-year	42,000,000	1.26
Undergraduate state grant amount, New York, public 4-year, SUNY	187,000,000	1.21
Undergraduate state grant amount, New York, public 2-year, SUNY	112,000,000	1.21
Undergraduate state grant amount, New York, public 4-year, CUNY	148,000,000	0.96
Undergraduate state grant amount, New York, public 2-year, CUNY	50,000,000	1.43
Undergraduate state grant amount, New York, independent colleges and universities	291,000,000	1.01
Undergraduate state grant amount, Texas, public 2-year	51,000,000	1.01
Undergraduate state grant amount, Texas, public 4-year	260,000,000	1.18
Undergraduate state grant amount, Texas, private not-for-profit 4-year	111,000,000	1.60
Number of students awarded SMART Grants for all states		
SMART Grant recipients	71,000	0.98
Number of students awarded ACGs for all states		
ACG recipients	400,000	0.99

NOTE: Detail may not sum to totals because of rounding. ACG = Academic Competitiveness Grant; CUNY = City University of New York; PLUS = Parent Loan for Undergraduate Students; SMART = Science and Mathematics Access to Retain Talent; SUNY = State University of New York.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

6.2.4 Student Weighting Adjustment Performance

As discussed in section 2.2, the response rates for the nation-level sectors were greater than 85 percent, so a nonresponse bias analysis was not necessary at the national level. Appendix K contains a detailed table showing the estimated bias before and after weight adjustments for California public, 2-year institutions—the one state-level sector that required nonresponse bias analysis.

Table 62 summarizes the institution weight distributions and the variance inflation due to UWE, by student type and type of institution. The median student weight ranges from 30 for students in private, for profit, 2-year-or-more institutions to 201 for students in public, 2-year institutions. The mean student weight ranges from 56 for students in private, not-for-profit, less-than-4-year institutions to 273 for students in public, 2-year institutions. The UWE is 2.5 overall and ranges from 1.3 for first-professional students to 7.0 for students in private, for-profit, 2-year-or-more institutions.

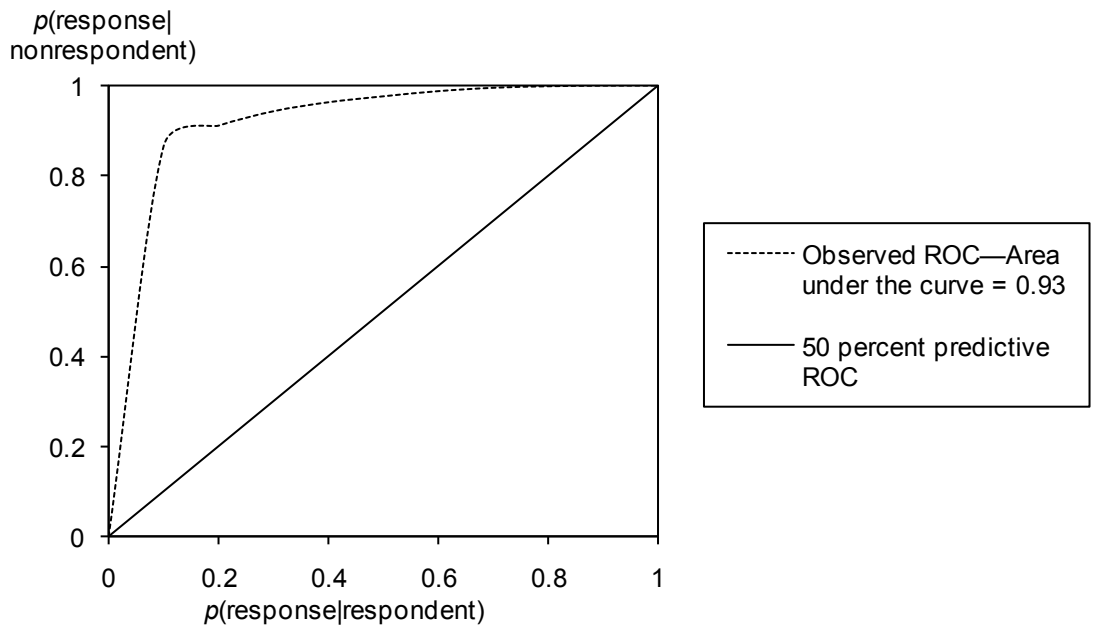
Table 62. Student weight distribution and unequal weighting effects: 2008

Analysis domain	Minimum	First quartile	Median	Third quartile	Maximum	Mean	UWE[1]
Total	0.21	44.23	119.21	265.89	9275.14	190.95	2.50
Student type							
Undergraduate	0.21	41.11	111.64	261.31	4566.02	184.33	2.39
Graduate	0.56	96.53	173.45	303.83	9275.14	252.22	2.97
First-professional	2.04	112.75	157.40	225.95	1080.97	178.43	1.32
Type of institution							
Public less-than-2-year	5.67	31.58	56.38	98.84	215.24	69.36	1.42
Public 2-year	0.28	52.95	201.28	423.87	2487.37	272.54	1.91
Public 4-year non-doctorate-granting	0.27	46.56	137.83	233.05	6409.74	166.79	2.05
Public 4-year doctorate-granting	0.44	72.85	153.45	291.27	5541.46	195.48	1.95
Private not-for-profit less-than-4-year	0.97	30.62	48.46	71.83	1132.90	56.24	1.62
Private not-for-profit 4-year non-doctorate-granting	0.54	33.93	86.40	181.62	3582.81	141.26	2.83
Private not-for-profit 4-year doctorate-granting	0.46	58.68	118.52	211.41	3918.61	154.73	2.17
Private for-profit less-than-2-year	1.09	44.23	62.96	98.62	394.61	79.24	1.44
Private for-profit 2-year-or-more	0.21	8.48	30.26	131.59	9275.14	170.09	7.03

[1] UWE calculated as sample size multiplied by the sum of the squared weights, divided by the sum of the weights squared.
NOTE: Detail may not sum to totals due to rounding. UWE = unequal weighting effect.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

To assess the overall predictive ability of the nonresponse model, an ROC curve was used and developed as described in section 6.1.4 and shown in figure 20. Figure 20 shows that the area under the ROC curve is 0.93, so the predicted probabilities give the correct classification 93 percent of the time. Predictive probabilities from ROC curves can also be interpreted in terms of the nonparametric Wilcoxon test statistic, where the ROC area of 0.93 equals the value of the Wilcoxon test statistic. Viewed in this way, the Wilcoxon test does not reject the null hypothesis of no predictive ability. This result can be interpreted to mean that the variables used in the model are highly informative but not definitive predictors of a sample student's overall response propensity. The predicted probabilities of response (c) were obtained as the product of the predicted response probabilities obtained at each of the three GEM nonresponse adjustment steps. Note that for the last GEM step (other nonresponse adjustment), predicted probabilities were not directly available for students who had already been dropped from the model because, in the previous step, they were not located. For these students, their predicted probability was set equal to the mean of the predicted probabilities of students still in the model.

Figure 20. Receiver Operating Characteristic curve for overall student response propensity: 2008



NOTE: ROC = Receiver Operator Characteristic.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

6.3 Item Nonresponse Bias Analysis

When item response rates were less than 85 percent, a nonresponse bias analysis was conducted. Item response rates (RRI) are calculated as the ratio of the number of respondents for whom an in-scope response was obtained (I^x for item x) to the number of respondents who were asked to answer that item. The number asked to answer an item is the number of unit-level respondents (I) minus the number of respondents with a valid skip item for item x (V^x). When an abbreviated questionnaire is used to convert refusals, the eliminated questions are treated as item nonresponse (U.S. Department of Education 2003):

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

A student is defined to be an *item respondent* for an analytic variable if that student has data for that variable from any source, including logical imputation. As shown in table 63, the weighted item response rates for all study respondents ranged from about 37 percent to 100 percent.

While values for many variables were derived from multiple sources, including the student interview, student record data, and administrative data sources, some variables were obtained from only one source. Because the weighted response rate for the student interview was about 72 percent, items obtained solely from that source have 28 percent nonresponse, even when all interview respondents provided an answer. This issue is compounded for nested items following gate questions, especially those applicable to a small subset of the sample members, because follow-up items to unanswered gate items are also treated as nonresponse.

To give an example, the student interview included a set of items about community service and was the only source for these data. Students were first asked if they had performed any community service or volunteer work during the 2007–08 academic year. Those who had were then asked about the types of community service or volunteer work they had performed. If the first item in the set was not answered, the subsequent questions about the types of community service or volunteer work were treated as nonresponse. More specifically, the gate question (COMSERV) had a weighted response rate of about 62 percent and was therefore missing for about one-third of study respondents. Of those who responded to the gate, only about 44 percent reported that they had performed any community service or volunteer work. One of the follow-up items, COMSERVA, was not applicable (was skipped) for those who reported not having performed any community service or volunteer work. These inapplicable cases were excluded from the response rate calculation, so the denominator used in computing the response rate for COMSERVA included those cases who were nonrespondents to the gate item (COMSERV), as well as those with a value of yes for the gate item. Additionally, some students who responded to the gate did not provide a response to the follow-up item, so COMSERVA has item nonresponse for some cases where COMSERV is yes. Consequently, the low response rate for COMSERVA is driven both by the large amount of missing data for COMSERV and by the small number of cases where COMSERVA was applicable but has missing data.

A nonresponse bias analysis was conducted for all items with a weighted response rate less than 85 percent for all students or for students in a particular sector. The possibility of estimating the degree of bias depends on having some variables that reflect key characteristics of respondents and for which there is little or no missing data. The variables listed above in the student-level bias analysis section, as well as sector, region, and student type, were used to compare the item respondents and nonrespondents. Additionally, gender and age group were used because they were known for all study respondents. Also, institution strata were used in analyses of items for all students. These variables are important to the study and are related to many of the items being analyzed for low item response rates. For these items, the nonresponse bias prior to imputation was estimated as described in the beginning of this chapter for each of these variables known for most respondents and nonrespondents and tested to determine if the bias was significant at the 5 percent level.

Table 63. Summary of weighted item response rates for all students and by type of institution: 2008

		Weighted response rate									
Variable	Variable label	All students	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting	Public 4-year doctorate-granting	Private not-for-profit less-than-4-year	Private not-for-profit 4-year non-doctorate-granting	Private not-for-profit 4-year doctorate-granting	Private for-profit less-than-2-year	Private for-profit 2-year-or-more
AGE	Age as of 12/31/07	100.0	99.8	100.0	100.0	100.0	100.0	99.8	99.9	100.0	100.0
AGEGROUP	Age groups as of 12/31/07	100.0	99.8	100.0	100.0	100.0	100.0	99.8	99.9	100.0	100.0
ATTEND	Attendance intensity in fall	99.4	98.9	99.4	99.8	99.8	98.7	99.4	99.8	98.0	98.1
ATTENDA	Reason enrolled: complete associate's degree	50.2	†	50.0	†	†	58.3	†	†	†	52.1
ATTENDB	Reason enrolled: complete undergraduate certificate or degree	62.1	62.2	63.7	22.0	26.5	61.5	17.0	18.7	55.3	56.4
ATTENDC	Reason enrolled: earn course credits at a different school	62.1	62.2	63.7	22.0	26.5	61.5	17.0	18.7	55.3	56.4
ATTENDD	Reason enrolled: gain job or occupational skills	62.1	62.2	63.7	22.0	26.5	61.5	17.0	18.7	55.3	56.4
ATTENDE	Reason enrolled: prepare for job certification or license	62.1	62.2	63.7	22.0	26.5	61.5	17.0	18.7	55.3	56.4
ATTENDF	Reason enrolled: prepare to transfer	61.7	62.2	63.7	3.6	†	61.5	#	†	55.3	56.4
ATTENDG	Reason enrolled: solely for personal interest	62.1	62.2	63.7	22.0	26.5	61.5	17.0	18.7	55.3	56.4
ATTENDMR	Main reason for attending NPSAS	62.0	62.2	63.6	22.0	26.5	61.5	17.0	18.7	55.1	56.3
BAYEAR	Year received bachelor's degree	80.6	†	†	75.5	81.9	†	83.1	81.7	†	72.2
CITIZEN2	Citizenship	97.5	83.8	96.1	98.5	99.0	97.5	97.5	98.5	98.1	97.6
COMHOUR	Community service: number of hours volunteered per month	36.4	17.6	26.6	39.9	48.2	14.6	45.8	50.8	8.3	21.0
COMHSRQ	Community service: required in high school	27.4	19.6	21.4	34.3	46.5	12.7	43.8	50.2	9.8	14.4
COMHSVL	Community service: voluntary in high school	27.4	19.6	21.4	34.3	46.5	12.7	44.0	50.2	9.8	14.6
COMONE	Community service: one time event	42.6	21.3	31.3	46.5	55.5	17.9	51.8	57.5	10.8	24.8
COMREQ	Community service: required or part of program	39.5	21.7	30.5	44.6	53.4	17.3	51.3	55.9	10.9	21.8
COMSERV	Community service: volunteered in last 12 months	62.3	49.8	57.3	65.2	70.3	40.1	65.1	69.4	36.6	53.6

See notes at end of table.

Table 63. Summary of weighted item response rates for all students and by type of institution: 2008—Continued

		Weighted response rate									
Variable	Variable label	All students	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting	Public 4-year doctorate-granting	Private not-for-profit less-than-4-year	Private not-for-profit 4-year non-doctorate-granting	Private not-for-profit 4-year doctorate-granting	Private for-profit less-than-2-year	Private for-profit 2-year-or-more
COMSERVA	Community service type: fundraising	42.9	21.7	31.7	46.7	55.9	18.0	52.3	57.9	10.9	25.0
COMSERVB	Community service type: health services	42.9	21.7	31.7	46.7	55.9	18.0	52.3	57.9	10.9	25.0
COMSERVC	Community service type: homeless shelter or soup kitchen	42.9	21.7	31.7	46.7	55.9	18.0	52.3	57.9	10.9	25.0
COMSERVD	Community service type: neighborhood improvement	42.9	21.7	31.7	46.7	55.9	18.0	52.3	57.9	10.9	25.0
COMSERVE	Community service type: non-education-related work with kids	42.9	21.7	31.7	46.7	55.9	18.0	52.3	57.9	10.9	25.0
COMSERVF	Community service type: service to church	42.9	21.7	31.7	46.7	55.9	18.0	52.3	57.9	10.9	25.0
COMSERVG	Community service type: tutoring or education-related	42.9	21.7	31.7	46.7	55.9	18.0	52.3	57.9	10.9	25.0
COMSERVX	Community service type: other	42.9	21.7	31.7	46.7	55.9	18.0	52.3	57.9	10.9	25.0
CRBALDUE	Credit cards: balance due on all credit cards	26.0	17.4	23.8	28.8	31.1	10.8	25.3	22.9	14.3	22.4
CRDAP04	Advanced Placement credit	63.0	53.0	57.2	65.9	70.1	33.6	69.6	69.1	37.5	47.8
CRDCL04	Earned college credit during high school	63.2	53.3	57.4	66.2	70.3	33.8	69.9	69.1	37.7	47.8
CRNUMCRD	Credit cards: number of credit cards in own name	62.8	53.0	55.9	65.6	69.4	31.7	69.4	68.3	39.1	47.2
CRPARPAY	Credit cards: parents help pay credit card bills	47.4	30.7	40.2	50.5	55.7	21.8	51.9	53.4	24.1	32.6
CRPAYOFF	Credit cards: pay off or carry balance	47.4	30.7	40.2	50.5	55.7	21.8	51.8	53.2	24.3	32.6
CSTBKS	Cost of books and supplies (student reported)	61.8	48.2	57.1	65.0	70.2	39.3	64.8	69.3	33.7	49.8
CSTCMPTR	Cost of special equipment (student reported)	53.6	44.7	48.4	56.4	61.8	35.6	54.9	60.6	30.6	45.4
DEGPR	Prior degree earned since high school	69.9	61.9	64.8	72.3	76.4	62.0	74.2	76.0	54.4	63.6

See notes at end of table.

Table 63. Summary of weighted item response rates for all students and by type of institution: 2008—Continued

		Weighted response rate									
Variable	Variable label	All students	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting	Public 4-year doctorate-granting	Private not-for-profit less-than-4-year	Private not-for-profit 4-year non-doctorate-granting	Private not-for-profit 4-year doctorate-granting	Private for-profit less-than-2-year	Private for-profit 2-year-or-more
DEGPRAA	Prior degree: associate's degree	43.6	34.0	32.0	43.6	54.2	36.7	48.7	62.1	19.9	40.0
DEGPRBA4	Prior degree: 4-year bachelor's degree	43.6	34.0	32.0	43.6	54.2	36.7	48.7	62.1	19.9	40.0
DEGPRBA5	Prior degree: 5-year bachelor's degree	43.6	34.0	32.0	43.6	54.2	36.7	48.7	62.1	19.9	40.0
DEGPRCRT	Prior degree: undergraduate certificate/diploma	43.6	34.0	32.0	43.6	54.2	36.7	48.7	62.1	19.9	40.0
DEGPRDOC	Prior degree: doctoral degree	43.6	34.0	32.0	43.6	54.2	36.7	48.7	62.1	19.9	40.0
DEGPRFP	Prior degree: first professional degree	43.6	34.0	32.0	43.6	54.2	36.7	48.7	62.1	19.9	40.0
DEGPRMS	Prior degree: master's degree	43.6	34.0	32.0	43.6	54.2	36.7	48.7	62.1	19.9	40.0
DEGPRPTB	Prior degree: post-BA certificate	43.6	34.0	32.0	43.6	54.2	36.7	48.7	62.1	19.9	40.0
DEGPRPTM	Prior degree: post-MA certificate	43.6	34.0	32.0	43.6	54.2	36.7	48.7	62.1	19.9	40.0
DEPANY	Dependents: has any dependents	69.4	62.8	60.5	68.5	70.3	82.2	75.1	76.3	85.7	90.3
DEPCARE	Dependents: children in daycare	55.6	42.7	54.1	59.3	66.2	39.7	57.5	63.8	30.7	51.1
DEPCHILD	Dependents: has dependent children	52.1	51.4	45.9	49.4	43.1	73.7	60.8	51.9	80.3	85.0
DEPCOST	Dependents: children in daycare—monthly daycare costs	34.4	21.1	32.3	38.7	48.1	22.2	36.7	43.3	13.1	29.0
DEPEND	Dependency status	91.8	90.9	87.7	92.4	92.2	96.2	95.6	95.0	97.2	99.6
DEPINC	Dependent parent income (cont)	55.2	38.6	43.0	62.1	55.6	73.9	73.2	63.1	78.6	87.4
DEPINCX	Dependent parent income (cat)	36.5	19.8	31.1	44.9	46.4	25.6	45.1	34.6	29.5	16.8
DEPNUM	Dependents: has any dependents (number)	51.9	51.4	45.6	49.3	42.8	73.7	60.7	51.7	80.3	85.0
DEPNUMCH	Dependents: has dependent children (number)	50.3	50.5	43.8	47.4	40.8	72.6	59.6	49.8	79.7	84.5
DEPOLD	Dependents: children—age of oldest	36.0	31.3	32.9	36.8	34.3	31.1	41.9	39.8	27.5	47.6
DEPOTHER	Dependents: has dependent(s) other than children	52.1	51.4	45.8	49.4	43.2	73.7	60.7	51.8	80.3	85.0
DEPTYPE	Dependents: types of dependents	51.9	51.4	45.6	49.3	42.8	73.7	60.7	51.7	80.3	85.0
DEPYNG	Dependents: children—age of youngest	36.0	31.3	32.9	36.8	34.3	31.1	41.9	39.8	27.5	47.6

See notes at end of table.

Table 63. Summary of weighted item response rates for all students and by type of institution: 2008—Continued

		Weighted response rate									
Variable	Variable label	All students	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting	Public 4-year doctorate-granting	Private not-for-profit less-than-4-year	Private not-for-profit 4-year non-doctorate-granting	Private not-for-profit 4-year doctorate-granting	Private for-profit less-than-2-year	Private for-profit 2-year-or-more
DISMOBIL	Disability: mobility impairment	62.3	49.7	57.2	65.1	70.3	40.4	65.2	69.3	36.5	53.6
DISOTHER	Disability: other long-lasting condition	62.3	49.7	57.2	65.1	70.3	40.3	65.1	69.4	36.5	53.6
DISSENSR	Disability: sensory impairment	62.3	49.7	57.2	65.0	70.3	40.6	65.1	69.5	36.5	53.5
DISTALL	Distance education: entire program	26.0	5.1	25.1	28.6	30.7	10.6	24.4	24.4	2.5	26.9
DISTEDUC	Distance education: took courses in 2007–08	63.3	50.2	58.2	66.0	71.2	41.5	66.0	70.4	37.6	54.6
DISTHOME	Distance from NPSAS school to home	58.7	49.4	55.4	62.1	67.2	37.0	62.4	64.6	35.8	49.0
DISTLOC	Distance education: location of courses	25.9	5.2	25.1	28.6	30.7	10.6	24.3	24.4	2.5	26.9
DISTWK	Job: distance from NPSAS school to work	42.0	38.4	39.2	44.8	48.7	25.8	43.7	48.2	25.1	29.2
DISTYPES	Disability: main type of condition or impairment	14.1	17.4	13.6	14.4	16.3	8.1	14.1	14.7	7.3	13.3
DSTUINC	Dependent student income (cont)	25.5	11.9	19.3	33.6	31.1	23.3	36.9	25.7	26.8	15.9
DSTUINCX	Dependent student income (cat)	37.0	20.5	31.8	45.4	46.9	25.7	45.4	35.1	29.6	16.8
EFCCPS	EFC (CPS)	50.4	42.8	39.4	54.4	48.2	71.1	63.1	51.7	80.8	83.7
EMPLYAM1	Employer aid (student)	78.7	63.3	75.9	82.5	86.8	60.7	80.1	81.9	56.6	63.0
EMPLYAM2	Employer aid (parents)	79.5	85.8	80.6	77.9	76.1	78.9	79.6	75.1	79.2	88.5
ENR01	Monthly enrollment status 2007/07	99.6	99.0	99.7	99.9	99.9	99.2	99.7	99.8	98.3	98.1
ENR02	Monthly enrollment status 2007/08	99.5	99.0	99.5	99.8	99.8	98.7	99.5	99.8	98.1	98.1
ENR03	Monthly enrollment status 2007/09	99.5	99.0	99.4	99.8	99.8	98.7	99.4	99.8	98.0	98.2
ENR04	Monthly enrollment status 2007/10	99.5	99.0	99.4	99.8	99.8	98.7	99.4	99.8	98.0	98.1
ENR05	Monthly enrollment status 2007/11	99.5	98.9	99.5	99.8	99.8	98.7	99.4	99.8	97.9	98.1
ENR06	Monthly enrollment status 2007/12	99.5	98.9	99.5	99.8	99.8	98.7	99.6	99.8	98.0	98.3

See notes at end of table.

Table 63. Summary of weighted item response rates for all students and by type of institution: 2008—Continued

Variable	Variable label	Weighted response rate									
		All students	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting	Public 4-year doctorate-granting	Private not-for-profit less-than-4-year	Private not-for-profit 4-year non-doctorate-granting	Private not-for-profit 4-year doctorate-granting	Private for-profit less-than-2-year	Private for-profit 2-year-or-more
ENR07	Monthly enrollment status 2008/01	99.5	98.9	99.4	99.8	99.8	98.2	99.5	99.8	97.9	98.4
ENR08	Monthly enrollment status 2008/02	99.5	98.9	99.4	99.8	99.8	98.1	99.5	99.8	97.8	98.4
ENR09	Monthly enrollment status 2008/03	99.5	98.9	99.4	99.8	99.8	98.1	99.5	99.8	97.8	98.4
ENR10	Monthly enrollment status 2008/04	99.5	98.8	99.4	99.8	99.8	98.2	99.5	99.8	97.8	98.4
ENR11	Monthly enrollment status 2008/05	99.5	98.9	99.4	99.8	99.8	98.2	99.5	99.8	97.9	98.4
ENR12	Monthly enrollment status 2008/06	99.6	98.9	99.7	99.8	99.9	99.2	99.8	99.8	98.0	98.5
EVER2PUB	Ever attended community college	83.5	61.9	100.0	70.5	74.2	61.4	73.1	73.1	54.3	62.5
EVER4YR	Ever attended 4-year institution	84.2	61.4	63.1	100.0	100.0	60.8	100.0	100.0	54.4	87.8
FEDBEN	Received federal benefit: any	50.6	42.9	39.5	54.6	48.4	71.1	63.6	52.3	80.8	83.8
FEDBENA	Received federal benefit: Food Stamp benefit	79.8	69.5	72.1	81.9	83.3	85.2	85.8	83.8	88.1	90.9
FEDBENB	Received federal benefit: Free/Reduced Price School Lunch benefits	50.6	43.1	39.6	54.6	48.4	71.1	63.6	52.4	80.9	83.9
FEDBENC	Received federal benefit: Supplemental Security Income benefits	79.8	69.4	72.0	81.9	83.3	85.2	85.8	83.8	88.1	90.9
FEDBEND	Received federal benefit: TANF benefits	79.8	69.4	72.0	81.9	83.3	85.2	85.8	83.8	88.1	90.9
FEDBENE	Received federal benefit: WIC benefits	50.6	43.1	39.5	54.6	48.4	71.1	63.6	52.3	80.9	83.8
FINAIDA	Financial aid decisions: compared lender options	62.3	49.8	57.3	65.1	70.0	40.7	64.8	69.4	37.3	53.9
FINAIDB	Financial aid decisions: discussed with family/friends	62.3	49.8	57.3	65.1	70.0	40.7	64.8	69.4	37.3	53.9
FINAIDC	Financial aid decisions: researched on Internet	62.3	49.8	57.3	65.1	70.0	40.7	64.8	69.4	37.3	53.9

See notes at end of table.

Table 63. Summary of weighted item response rates for all students and by type of institution: 2008—Continued

Variable	Variable label	Weighted response rate									
		All students	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting	Public 4-year doctorate-granting	Private not-for-profit less-than-4-year	Private not-for-profit 4-year non-doctorate-granting	Private not-for-profit 4-year doctorate-granting	Private for-profit less-than-2-year	Private for-profit 2-year-or-more
FINAIDD	Financial aid decisions: talked with staff	62.3	49.8	57.3	65.1	70.0	40.7	64.8	69.4	37.3	53.9
FINSPRT	Other financial support received	62.4	49.9	57.3	65.2	70.4	40.5	65.1	69.6	36.8	53.6
FINSPRTA	Other financial support: housing costs	14.1	10.9	11.3	15.1	18.9	8.3	17.4	19.4	7.5	9.8
FINSPRTB	Other financial support: living expenses	14.1	10.9	11.3	15.1	18.9	8.3	17.4	19.4	7.5	9.8
FINSPRTC	Other financial support: other education	14.1	10.9	11.3	15.1	18.9	8.3	17.4	19.4	7.5	9.8
FINSPRTD	Other financial support: tuition and fees	14.1	10.9	11.3	15.1	18.9	8.3	17.4	19.4	7.5	9.8
GAINSUR	Graduate assistantship: included health insurance	68.3	†	†	59.8	71.1	†	55.6	65.0	†	†
GENDER	Gender	100.0	100.0	100.0	99.9	100.0	100.0	100.0	100.0	99.9	100.0
GPA	Grade point average	89.6	47.7	85.7	93.5	96.7	68.8	92.7	93.4	72.5	88.5
GRADLVL	Graduate class level	98.4	†	†	98.3	97.9	†	99.1	98.4	†	99.9
GRADPYR	Year began graduate degree	76.6	†	†	77.3	78.6	†	74.1	76.0	†	70.2
GRADTAA	Teaching assistantship duties: answered student e-mails	65.9	†	†	51.5	67.8	†	58.2	64.5	†	†
GRADTAB	Teaching assistantship duties: assisted faculty with grading	65.9	†	†	51.5	67.8	†	58.2	64.5	†	†
GRADTAC	Teaching assistantship duties: assumed teaching responsibility	65.9	†	†	51.5	67.8	†	58.2	64.5	†	†
GRADTAD	Teaching assistantship duties: held office hours	65.9	†	†	51.5	67.8	†	58.2	64.5	†	†
GRADTAE	Teaching assistantship duties: led discussions	65.9	†	†	51.5	67.8	†	58.2	64.5	†	†
GRADTAF	Teaching assistantship duties: supervised lab sections	65.9	†	†	51.5	67.8	†	58.2	64.5	†	†
GRASTAMT	Total assistantships amount	71.1	†	†	70.9	74.4	†	67.0	70.5	†	60.3
GRGRDAMT	Other graduate assistantship amount	71.7	†	†	71.0	75.4	†	67.3	71.0	†	60.3

See notes at end of table.

Table 63. Summary of weighted item response rates for all students and by type of institution: 2008—Continued

Variable	Variable label	Weighted response rate									
		All students	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting	Public 4-year doctorate-granting	Private not-for-profit less-than-4-year	Private not-for-profit 4-year non-doctorate-granting	Private not-for-profit 4-year doctorate-granting	Private for-profit less-than-2-year	Private for-profit 2-year-or-more
GRRESAMT	Graduate research assistantship amount	71.6	†	†	70.9	75.4	†	67.3	70.8	†	60.3
GRTEAAMT	Graduate teaching assistantship amount	71.7	†	†	70.9	75.3	†	67.2	71.1	†	60.3
HCHONORS	Number of honors subjects	59.1	22.5	45.2	68.8	78.2	26.5	71.4	79.5	21.5	29.0
HCMATHHI	Highest level of math completed/planned	73.7	58.4	67.3	77.9	84.5	45.7	78.8	82.3	43.9	58.5
HCTKBIOL	Took/planned to take Biology	79.6	60.3	73.2	84.8	90.7	49.8	85.0	88.6	47.3	63.1
HCTKCHEM	Took/planned to take Chemistry	78.9	59.6	72.2	84.2	90.1	49.5	84.3	88.1	46.6	62.1
HCTKPHYS	Took/planned to take Physics	75.9	57.9	69.4	80.6	86.9	47.7	81.6	85.0	44.6	59.5
HCYSENGL	Years completed/planned English	55.1	22.6	41.9	64.3	73.7	25.1	67.1	72.1	19.7	27.8
HCYSLANG	Years completed/planned foreign languages	53.9	21.3	40.6	63.1	72.3	24.9	65.8	71.3	19.1	27.2
HCYSMATH	Years completed/planned math	80.4	60.4	73.8	85.8	91.5	51.2	86.3	89.2	47.8	63.2
HCYSSCIE	Years completed/planned science	79.8	60.2	73.1	85.3	90.8	50.4	85.8	88.7	47.3	63.0
HCYSSOCI	Years completed/planned social studies	54.9	22.2	41.6	64.2	73.6	25.1	66.9	72.1	19.5	27.3
HIGHLVEX	Highest level of education ever expected	68.1	57.8	62.4	70.2	75.7	59.6	72.4	75.3	51.9	60.3
HISPANIC	Hispanic or Latino origin	92.4	90.5	92.1	94.1	95.4	92.3	92.2	91.5	86.3	85.5
HISPTYPE	Hispanic type	38.1	30.3	40.5	43.6	48.6	15.7	30.3	32.3	19.0	27.1
HOMEPAR	Parents own home or pay mortgage	62.7	52.6	55.8	65.6	69.4	31.0	69.3	68.4	38.6	46.6
HOMESTUD	Student owns home or pays mortgage	52.5	39.8	49.6	54.8	59.5	38.5	53.8	60.5	23.8	45.4
HSCRDAP	Earned Advanced Placement credit in high school	61.4	51.7	56.5	64.5	69.1	37.7	67.0	67.8	37.3	50.6
HSCRDCOL	Earned college credits at a college in high school	61.5	52.0	56.7	64.8	69.2	37.9	67.3	67.8	37.3	50.7
HSDEG	High school degree type	92.2	81.1	92.6	93.2	93.1	90.9	91.2	88.4	90.6	90.7
HSGPA	Grade point average in high school	56.8	21.7	42.8	65.9	76.1	25.5	69.7	76.4	20.3	27.3

See notes at end of table.

Table 63. Summary of weighted item response rates for all students and by type of institution: 2008—Continued

Variable	Variable label	Weighted response rate									
		All students	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting	Public 4-year doctorate-granting	Private not-for-profit less-than-4-year	Private not-for-profit 4-year non-doctorate-granting	Private not-for-profit 4-year doctorate-granting	Private for-profit less-than-2-year	Private for-profit 2-year-or-more
HSGRADYY	High school graduation year	87.2	69.0	88.2	89.2	91.0	76.2	84.7	83.0	73.3	80.1
HSTYPE	Type of high school attended	62.0	50.7	58.1	65.4	69.9	46.5	66.5	69.2	38.1	52.1
INCOMPAR	Earnings in 2007	60.6	48.2	53.4	63.5	67.6	30.8	67.3	66.4	36.1	45.2
INCOMST	Independent student and spouse income (cont)	61.8	48.3	56.6	64.7	69.9	40.0	64.6	69.0	36.5	53.4
INDEPINC	Independent student and spouse income (cont)	25.7	32.4	21.4	21.6	17.8	48.0	26.9	26.9	54.2	67.9
INJURIS	Tuition jurisdiction (in/out of area)	99.8	100.0	99.9	99.7	99.9	†	†	†	†	†
INLNAMT	Institutional loans	98.2	†	†	99.5	98.5	†	93.5	98.4	†	100.0
ISTUINC	Independent student income (cont)	25.8	32.4	21.4	21.7	17.8	48.0	27.0	27.0	54.3	68.0
ISTUINCX	Independent student income (cat)	42.7	48.7	40.3	36.3	36.1	59.4	40.3	48.6	58.6	74.2
JOBAFFOR	Job: can afford school without working (student)	43.6	28.0	38.7	50.7	55.4	20.4	43.8	46.2	18.6	26.7
JOBCLASS	Job: related to coursework (non-degree)	45.4	13.1	46.6	42.7	50.3	35.1	36.5	42.7	4.9	48.6
JOBEARN	Job: earnings from work while enrolled (excluding work-study/assistantship)	51.1	37.8	48.1	54.2	58.4	29.9	51.7	56.0	24.5	43.4
JOBEFFA	Job affects school: helped with career preparation (student)	43.6	28.0	38.7	50.7	55.4	20.3	43.7	46.0	18.5	26.8
JOBEFFB	Job affects school: helped with class work (student)	43.6	28.0	38.7	50.7	55.4	20.3	43.7	46.0	18.5	26.8
JOBEFFC	Job affects school: limited access to campus facilities (student)	43.6	28.0	38.7	50.7	55.4	20.3	43.7	46.0	18.5	26.8
JOBEFFD	Job affects school: limited the class schedule (student)	43.6	28.0	38.7	50.7	55.4	20.3	43.7	46.0	18.5	26.8
JOBEFFE	Job affects school: limited the number of classes (student)	43.6	28.0	38.7	50.7	55.4	20.3	43.7	46.0	18.5	26.8
JOBEFFF	Job affects school: restricted choice of classes (student)	43.6	28.0	38.7	50.7	55.4	20.3	43.7	46.0	18.5	26.8
JOBEFFGR	Job: effect on grades (student)	43.6	28.0	38.7	50.6	55.4	20.3	43.8	46.0	18.5	26.8
JOBEMPL	Job: type of employer	54.2	39.5	51.1	57.8	61.1	32.6	55.8	59.1	25.2	46.3

See notes at end of table.

Table 63. Summary of weighted item response rates for all students and by type of institution: 2008—Continued

		Weighted response rate									
Variable	Variable label	All students	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting	Public 4-year doctorate-granting	Private not-for-profit less-than-4-year	Private not-for-profit 4-year non-doctorate-granting	Private not-for-profit 4-year doctorate-granting	Private for-profit less-than-2-year	Private for-profit 2-year-or-more
JOBHOUR	Job: hours worked per week (excluding work-study/assistantship)	53.8	40.2	51.5	56.9	60.2	32.5	54.6	57.9	25.9	46.3
JOBMAJOR	Job: related to major (degree)	50.0	42.2	40.3	56.4	59.8	29.1	57.0	59.5	26.8	43.6
JOBNUM	Job: number (excluding work-study/assistantship)	62.9	50.2	57.9	65.6	70.7	40.7	65.5	70.0	37.1	54.0
JOBONOFF	Job: located primarily on or off campus	55.5	40.8	52.7	59.2	62.3	33.3	57.0	60.1	26.9	47.4
JOBPRIOR	Job: have job prior to enrollment at NPSAS school	55.4	40.9	52.6	59.1	62.2	33.1	56.9	60.1	26.7	47.2
JOBROLE	Job: primary role as student or employee (excluding work-study/assistant)	55.2	40.9	52.4	58.9	62.0	33.2	56.8	59.8	26.5	47.1
JOBRSNA	Reason for working: earn spending money	43.7	28.0	38.7	50.7	55.5	20.4	43.8	46.2	18.6	26.8
JOBRSNB	Reason for working: gain job experience	43.7	28.0	38.7	50.7	55.5	20.4	43.8	46.2	18.6	26.8
JOBRSNC	Reason for working: minimize debt	43.7	28.0	38.7	50.7	55.5	20.4	43.8	46.2	18.6	26.8
JOBRSND	Reason for working: parents' expectations	51.8	33.5	46.5	56.4	59.6	20.1	54.6	53.0	26.0	34.3
JOBRSNE	Reason for working: pay educational expenses	43.7	28.0	38.7	50.7	55.5	20.4	43.8	46.2	18.6	26.8
JOBRSNF	Reason for working: pay living expenses	43.7	28.0	38.7	50.7	55.5	20.4	43.8	46.2	18.6	26.8
JOBRSNG	Reason for working: to send money home	43.7	28.0	38.7	50.7	55.5	20.4	43.8	46.2	18.6	26.8
JOBSAVE	Amount saved from summer 2007 earnings	45.9	27.8	40.0	51.0	54.3	17.3	50.1	46.7	19.9	27.1
JOBSUMMR	Worked during summer 2007	51.8	33.5	46.5	56.5	59.6	20.1	54.3	53.0	26.1	34.1
JOBTYPE	Job: type of job	55.4	40.9	52.6	59.1	62.1	33.2	57.0	60.0	26.8	47.4
JOBWEEK	Job: weeks worked while enrolled	55.0	40.5	52.1	58.7	61.8	32.8	56.5	59.8	26.4	47.0
LNCOSIGN	Cosign on student loans	61.7	61.4	60.4	66.2	68.6	48.9	67.3	69.7	35.8	51.5
LNREPAY	Expect help with repaying student loans	59.2	50.6	52.1	62.1	66.0	42.6	62.6	67.2	35.0	50.5

See notes at end of table.

Table 63. Summary of weighted item response rates for all students and by type of institution: 2008—Continued

		Weighted response rate									
Variable	Variable label	All students	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting	Public 4-year doctorate-granting	Private not-for-profit less-than-4-year	Private not-for-profit 4-year non-doctorate-granting	Private not-for-profit 4-year doctorate-granting	Private for-profit less-than-2-year	Private for-profit 2-year-or-more
LOCALRES	Residence while enrolled	76.2	68.1	67.5	80.3	83.5	83.6	84.9	84.7	78.7	83.3
MAJCHG	Majors: ever formally changed	53.8	†	44.0	59.9	64.2	44.4	62.7	63.6	†	46.3
MAJCHGFQ	Majors: frequency of formally changed	23.1	†	16.4	30.1	36.9	4.8	25.2	28.8	†	6.8
MAJORS	Field of study/major (detailed)	97.0	96.5	95.2	98.2	98.5	94.0	97.2	98.2	93.6	97.6
MILTYPE	Military type	98.3	94.6	97.2	99.0	99.3	98.9	99.0	97.7	99.3	99.9
MNTRENT	Monthly mortgage/rent amount	55.1	41.3	49.9	55.2	63.6	37.3	54.8	65.1	28.2	48.4
ORPHAN	Orphan or ward of court	81.8	72.3	74.1	84.7	85.0	83.0	90.1	86.2	88.6	91.5
OWNINVPA	Parents own investments, business or farm over $10,000	62.1	52.1	55.2	65.0	68.8	30.8	68.9	67.7	38.6	45.2
OWNINVST	Student owns investments, business or farm over $10,000	53.1	42.2	50.3	55.4	59.5	41.2	53.8	60.4	27.5	48.4
PAGI	Dependent student parents AGI	24.5	11.0	18.0	32.3	30.4	20.8	36.1	25.1	23.1	14.6
PARBORN	Born in the U.S. (parents)	62.5	50.0	57.4	65.3	70.5	40.2	65.1	69.5	36.6	53.6
PARCOLL	Parents taking college courses	63.4	52.5	56.2	66.4	70.2	31.1	70.1	69.8	38.8	46.7
PARHELPA	Help from parents: housing	63.7	53.0	56.5	66.6	70.5	31.7	70.2	70.2	39.4	47.2
PARHELPB	Help from parents: other educational expenses	63.7	53.0	56.5	66.6	70.5	31.7	70.2	70.2	39.4	47.2
PARHELPC	Help from parents: other living expenses	63.7	53.0	56.5	66.6	70.5	31.7	70.2	70.2	39.4	47.2
PARHELPD	Help from parents: tuition and fees	63.7	53.0	56.5	66.6	70.5	31.7	70.2	70.2	39.4	47.2
PDADED	Father's highest education level	62.3	49.8	57.3	65.2	70.3	40.2	65.1	69.3	36.5	53.5
PFAMNUM	Family size (dependent)	79.5	66.6	70.0	83.4	83.6	81.6	89.2	85.6	86.0	91.6
PFEDBEN	Dependent student parents number of federal benefits	54.1	35.2	40.8	61.3	54.8	73.2	72.9	62.6	77.9	87.3
PFEDTAX	Dependent student parents federal tax paid	24.6	10.4	17.9	32.4	30.3	23.1	36.3	25.2	26.4	15.7
PINCOL	Number of family members in college (dependent)	71.5	54.4	60.7	75.3	75.0	77.2	83.4	79.9	82.8	89.6
PMARITAL	Marital status (parents)	80.3	67.7	71.1	84.2	84.1	81.6	90.0	86.0	86.7	92.2
PMOMED	Mother's highest education level	62.0	49.5	56.9	64.8	70.1	40.0	64.6	69.1	36.3	53.2

See notes at end of table.

Table 63. Summary of weighted item response rates for all students and by type of institution: 2008—Continued

Variable	Variable label	Weighted response rate									
		All students	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting	Public 4-year doctorate-granting	Private not-for-profit less-than-4-year	Private not-for-profit 4-year non-doctorate-granting	Private not-for-profit 4-year doctorate-granting	Private for-profit less-than-2-year	Private for-profit 2-year-or-more
PRIMLANG	English is the primary language	62.4	49.7	57.4	65.3	70.5	40.3	65.1	69.5	36.6	53.5
PRIVAID	Private sources grants	99.0	98.3	98.9	99.4	99.5	96.7	98.2	99.1	95.6	99.5
PRIVLOAN	Private (alternative) loans	66.8	52.3	59.1	69.4	76.0	47.0	70.8	75.9	47.0	61.9
PROGSTAT	Completed degree program in 2007–08	96.5	96.2	94.7	98.7	96.6	97.5	98.0	97.2	96.4	98.9
PSECTYR	Year first enrolled in postsecondary education	65.6	58.6	61.3	69.1	72.6	59.0	71.2	71.3	51.5	59.9
PTAXFILE	Dependent student parents federal tax filed	25.0	10.7	18.3	33.2	30.6	23.1	36.8	25.5	26.6	15.9
RAASIAN	Race: Asian	92.0	93.4	92.9	92.6	95.2	69.6	91.4	91.1	74.0	83.9
RABLACK	Race: Black or African American	92.0	93.4	92.9	92.6	95.2	69.6	91.4	91.1	74.0	83.9
RAINDIAN	Race: American Indian or Alaska Native	92.0	93.4	92.9	92.6	95.2	69.6	91.4	91.1	74.0	83.9
RAINDTRB	Race: American Indian or Alaska Native recognized tribe	73.7	62.5	74.0	68.2	79.7	47.5	76.0	74.7	64.6	73.3
RAISLAND	Race: Native Hawaiian/Other Pacific Islander	92.0	93.4	92.9	92.6	95.2	69.6	91.4	91.1	74.0	83.9
RAOTHER	Race: Other	92.0	93.4	92.9	92.6	95.2	69.6	91.4	91.1	74.0	83.9
RAWHITE	Race: White	92.0	93.4	92.9	92.6	95.2	69.6	91.4	91.1	74.0	83.9
REANOAPA	Reason for not applying: did not want to take on the debt	49.7	35.8	47.2	50.7	57.5	52.3	46.4	52.6	29.5	30.7
REANOAPB	Reason for not applying: forms were too much work	49.7	35.8	47.2	50.7	57.5	52.3	46.4	52.6	29.5	30.7
REANOAPC	Reason for not applying: no information about how to apply	49.7	35.8	47.2	50.7	57.5	52.3	46.4	52.6	29.5	30.7
REANOAPD	Reason for not applying: no need	49.7	35.8	47.2	50.7	57.5	52.3	46.4	52.6	29.5	30.7
REANOAPE	Reason for not applying: thought ineligible	49.7	35.8	47.2	50.7	57.5	52.3	46.4	52.6	29.5	30.7
REMENUM	Remedial courses: number taken in 2007–08	15.3	8.4	16.3	19.7	19.6	3.8	15.6	13.3	2.7	8.0
REMETOOK	Remedial courses: took in 2007–08	27.8	23.2	29.8	30.9	29.5	14.3	27.3	23.7	12.5	18.2
REMEVER	Remedial courses: ever taken	60.6	50.4	56.8	64.2	68.8	40.5	64.9	67.0	37.6	53.0
SAGI	Independent student AGI	23.1	28.9	18.8	19.7	16.3	42.1	25.1	24.9	45.6	64.4

See notes at end of table.

Table 63. Summary of weighted item response rates for all students and by type of institution: 2008—Continued

Variable	Variable label	Weighted response rate									
		All students	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting	Public 4-year doctorate-granting	Private not-for-profit less-than-4-year	Private not-for-profit 4-year non-doctorate-granting	Private not-for-profit 4-year doctorate-granting	Private for-profit less-than-2-year	Private for-profit 2-year-or-more
SAMESTAT	Attend institution in state of legal residence	98.4	97.4	97.6	99.0	98.8	98.6	98.9	98.5	99.9	99.7
SCHWKHR	Hours per week on schoolwork outside of class	62.7	49.4	57.8	65.5	70.6	40.9	65.5	69.5	37.0	53.7
SEOGAMT	Federal Supplemental Educational Opportunity Grant (SEOG)	98.9	97.7	98.9	99.5	99.1	98.5	98.5	98.6	95.2	99.4
SFEDBEN	Independent student number of federal benefits	41.3	41.0	31.4	40.0	34.3	66.3	49.7	41.7	79.0	82.6
SFEDTAX	Independent student federal tax paid	25.0	31.3	20.4	20.7	17.4	47.7	26.4	26.5	54.0	66.2
SINCOL	Number of family members in college (independent)	55.3	55.7	49.1	54.4	46.6	74.4	63.3	54.7	85.5	87.3
SJASST	School job: assistantship	59.7	†	53.9	63.4	67.6	50.6	64.4	66.5	†	51.9
SJCOMSER	Work-study job: community service project	11.7	†	6.3	12.0	14.2	6.0	27.5	28.0	†	5.0
SJEARN	School job: total earnings	15.4	†	6.0	13.8	25.3	4.2	27.3	33.3	†	4.9
SJHOURS	School job: hours worked per week (work-study/assistantship)	15.8	†	6.5	14.6	25.7	5.6	27.9	32.8	†	5.3
SJMAJOR	School job: related to major or coursework	13.4	†	7.2	14.3	17.1	6.8	29.2	30.9	†	5.8
SJONOFF	School job: located primarily on or off campus	13.1	†	7.1	14.0	16.7	6.4	28.7	30.3	†	5.8
SJSCHOOL	School job: for NPSAS or another institution/organization	13.0	†	6.9	14.0	16.7	6.4	28.6	30.2	†	5.7
SJTUTOR	Work-study job: literacy education or tutoring	11.7	†	6.2	12.0	14.2	5.8	27.5	27.9	†	5.0
SJWEEK	School job: weeks worked	16.1	†	6.9	14.9	26.0	6.1	28.3	33.1	†	5.6
SJWKST	School job: work-study job	59.7	†	53.9	63.4	67.6	50.6	64.4	66.5	†	51.9
SMARITAL	Marital status (student)	88.2	81.6	81.8	89.2	90.2	97.8	95.3	92.6	97.6	96.4
SPINCOL	Spouse attending college	63.6	49.9	61.9	66.1	72.8	45.9	63.2	69.3	32.1	50.2
SPSINC	Independent student spouse income (cont)	11.8	19.4	9.8	9.5	6.8	27.4	14.2	12.3	24.8	47.0

See notes at end of table.

Table 63. Summary of weighted item response rates for all students and by type of institution: 2008—Continued

Variable	Variable label	Weighted response rate									
		All students	Public less-than-2-year	Public 2-year	Public 4-year non-doctorate-granting	Public 4-year doctorate-granting	Private not-for-profit less-than-4-year	Private not-for-profit 4-year non-doctorate-granting	Private not-for-profit 4-year doctorate-granting	Private for-profit less-than-2-year	Private for-profit 2-year-or-more
SPSINCX	Independent student spouse income (cat)	24.3	34.2	24.2	19.7	18.3	35.9	25.1	28.9	27.9	53.4
STAXFILE	Independent student federal tax filed	25.6	32.2	21.1	21.4	17.8	47.9	26.8	26.8	54.3	67.9
STUSTATE	State of legal residence (student)	98.5	97.4	97.6	99.1	98.9	98.6	99.0	98.7	100.0	100.0
TEACTDER	ACT derived composite score	75.0	33.6	60.9	86.1	93.6	34.4	84.8	91.8	27.7	37.4
TESATDER	SAT derived combined score	75.0	33.6	60.9	86.1	93.6	34.4	84.8	91.8	27.7	37.4
TESATMDE	SAT derived math score	71.6	33.6	58.0	81.5	88.3	33.2	82.6	90.1	27.7	37.0
TESATVDE	SAT derived verbal score	71.6	33.6	58.0	81.4	88.3	33.2	82.6	90.1	27.7	37.0
TETOOK	Took SAT or ACT exams	84.8	62.0	77.8	91.3	96.7	52.5	90.4	95.1	49.7	64.7
TFEDWRK	Federal work-study	99.4	99.3	99.5	99.7	99.7	97.6	98.6	99.1	97.8	99.5
TRANCRD	Ever transfer credits to NPSAS	53.1	†	36.6	51.2	54.3	69.1	55.1	57.4	†	27.5
UGDEGAA	Associate degree type	99.7	100.0	99.3	99.9	100.0	100.0	100.0	100.0	100.0	100.0
UNTAXBF	Received untaxed benefits in 2007	61.6	48.3	58.1	64.0	70.5	44.8	60.7	69.4	35.9	55.0
UNTAXBFA	Received untaxed benefits: child support	25.0	25.1	25.3	25.6	24.1	14.8	22.6	18.2	18.2	30.0
UNTAXBFB	Received untaxed benefits: disability payments	30.2	32.9	33.4	26.4	26.6	23.1	20.7	24.4	20.2	32.7
UNTAXBFC	Received untaxed benefits: worker's compensation	17.4	21.1	19.8	15.7	12.4	12.3	14.3	9.6	15.2	23.0
USBORN	Born in the U.S. (student)	62.4	49.7	57.4	65.2	70.5	40.3	65.1	69.6	36.7	53.1
VETBEN	Veterans benefits	99.6	99.7	99.5	99.4	99.9	99.9	99.1	99.9	99.9	99.5
VETERAN	Veteran status	98.6	95.6	97.6	99.1	99.4	98.9	99.1	98.0	99.4	99.9
VOTEEVER	Vote: ever voted	62.5	50.2	57.4	65.9	70.2	40.6	65.4	69.9	37.7	54.0
VOTEREG	Vote: registered to vote	62.6	50.1	57.4	65.8	70.3	40.7	65.4	70.0	37.8	54.0

† Not applicable.

NOTE: Nonresponse bias analysis was conducted only for each item with a weighted response rate less than 85 percent. Nonresponse bias analysis was based on the student-level variables known for both respondents and nonrespondents (described in section 6.2.3). While values for many variables are derived from multiple sources, including the student interview, student record data, and administrative data sources, some variables are obtained from only one source. Given that the weighted response rate to the student interview was about 72 percent, items obtained solely from the student interview have 28 percent nonresponse, even when all student interview respondents provided an answer. Computing the response rate is a complex process for nested items following gate questions. Response rates for nested items include those cases with an applicable reported or imputed value for the gate item. AGI = adjusted gross income; cat = categorical variables; cont = continuous variables; CPS = Central Processing System; EFC = expected family contribution; SAT = SAT Reasoning Test; TANF = Temporary Assistance for Needy Families; WIC = Special Supplemental Nutrition Program for Women, Infants, and Children.

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2007–08 National Postsecondary Student Aid Study (NPSAS:08).

Appendix K contains a table (table K-9) using one variable (DISTYPES) to illustrate the estimated bias before imputation for all students. Similar computations were done for about 220 additional variables with item response rates less than 85 percent for all students or for students in at least one sector. Table K-10 in appendix K summarizes these computations. This table also shows a large range for the percent of variable categories with significant bias across all items analyzed prior to imputation. A byproduct of the imputation (described in section 5.5) is the reduction or elimination of item-level nonresponse bias. Imputation reduces or eliminates nonresponse bias by replacing missing data with statistically plausible values. Missing data and the associated nonresponse bias for variables are usually not ignorable (i.e., the respondents' distribution patterns differ from those in the full population); therefore, replacing missing data with reasonable values produces imputed sample distributions that resemble full population distributions, thus reducing nonresponse bias. The use of carefully constructed imputation classes, donor-imputee matching criteria, and random hot-deck searches within imputation cells are all designed to ensure that imputed data are, in fact, plausible and that the nonresponse bias is ignorable within the imputation classes.

While item-level bias before imputation is measurable, such bias after imputation is not, so how well the imputation worked in reducing bias cannot be directly evaluated. The item estimates before and after imputations were compared, therefore, to determine whether the imputation changed the biased estimate, which would suggest a reduction in bias.

For continuous variables, the difference between the mean before imputation and the mean after imputation was estimated. For categorical variables, the estimated difference was computed for each of the categories as the percentage of students in that category before imputation minus the percentage of students in that category after imputation. The estimated change in means before and after imputation was then tested to determine if the difference was significant at the .05 level.

The relative item nonresponse bias in estimates before imputation is not comparable to the percent difference in means after imputation, because the ways the measures are calculated differ. Because imputation procedures are designed specifically to identify donors with characteristics similar to those with missing data, the imputation is assumed to reduce bias within imputation classes.

A significant difference in the item means after imputation implies a bias reduction due to imputation. Likewise, a nonsignificant difference suggests that imputation may not have reduced bias, that the sample size was too small for a significant difference to be detectable, or that there was little bias to be reduced.

As shown in table K-10 in appendix K, there are significant differences between estimates computed before and after imputation for about 61 percent of the variables (i.e., those with statistically significant [starred] percent differences in pre- and post-imputation means) analyzed for all students and for about 87 percent of the variables analyzed for at least one sector. These results indicate a likely reduction in bias for these variables. Approximately 10 percent of the variables that were analyzed were found to have no significant differences. While some of these variables may be biased, others have a small amount of bias prior to imputation or have small sample sizes if they are only applicable to graduate and first-professional students or to a subset of students. However, analysts should use the potentially significantly biased items with caution.

6.4 Variance Estimation

For probability-based sample surveys, most estimates are nonlinear statistics. For example, a mean or proportion, which is expressed as $\Sigma wy/\Sigma w$, is nonlinear because the denominator is a survey

estimate of the (unknown) population total. In this situation, the variances of the estimates cannot be expressed in closed form. Two procedures for estimating variances of survey statistics are the Taylor-series linearization procedure and the bootstrap replication procedure, which are both available for the NPSAS data files. The analysis strata and replicates created for the Taylor-series procedure are discussed in section 6.4.1, and section 6.4.2 discusses the replicate weights created for the bootstrap procedure. Section 6.4.3 discusses the computation and use of design effects to measure the effects that complex sample design features had on the variances of survey estimates.

6.4.1 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) presented the mathematical formulation of this procedure.

For stratified multistage surveys, the Taylor-series procedure requires analysis strata and analysis primary sampling units (PSUs), also called *replicates*, as defined from the sampling strata and PSUs used in the first stage of sampling. For NPSAS:08, analysis strata and analysis PSUs were defined separately for all students combined and can be used for analyses of any domain.

The first step was to identify the PSUs used at the first stage of sample selection. As discussed in chapter 2, the PSUs included the 930 noncertainty institutions. The 1,030 certainty institutions were also treated as PSUs due to institution nonresponse, even though the students represent the first stage of sampling. The next step was to sort the PSUs by the 46 institution strata, then by certainty versus noncertainty, and then by the selection order for the noncertainty institutions and by IPEDS ID for the certainty institutions. From this sorted list, the analysis PSUs were then defined by collapsing the PSUs as required so that each analysis PSU contained at least four respondents. This sample size requirement ensured stable variance estimates. Analysis PSUs were then paired to form analysis strata. This process resulted in 845 analysis strata. The names of the analysis strata and analysis PSU variables are ANALSTR and ANALPSU, respectively.

The procedure described above may overestimate the variance because it does not always account for the finite population correction (FPC) at the institution stage of sampling. Alternatively, the Taylor-series procedure can account for the FPC if the secondary sampling units (SSUs) and PSU counts are considered in addition to the analysis strata and analysis PSUs. These variable names are FANALSTR, FANALPSU, FANALSSU, and PSUCOUNT for the analysis strata, PSUs, and SSUs and the PSU counts, respectively. These variables were created as part of the process for creating the bootstrap replicate weights (described below). The analysis PSUs were paired to form analysis strata differently for this process.

6.4.2 Bootstrap Replicate Weights

The variance estimation strategy was chosen for NPSAS:08 to satisfy 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 for nonresponse and for poststratification of selected total estimates to known external totals;
5. satisfactory properties for estimating variances of nonlinear statistics and percentages, as well as for linear statistics;
6. ability to apply finite population corrections at the institution stage of sampling and reflect the reduction in variance due to the high sampling rates in some first-stage sampling strata; and
7. ability to test hypotheses about students on the basis of normal distribution theory by ignoring the finite population corrections at the student level of sampling.

Commonly applied bootstrap variance estimation techniques satisfy requirements 1 through 5. To meet requirements 6 and 7 as well, 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*.

m_{hi} = the number of second-stage units selected from institution *i* in stratum *h*.

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

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

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

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

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

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

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

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

2. Generate a uniform (0, 1) random number R_h for each stratum *h*.
3. If $R_h \le \text{FPC}_h$, 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. Steps1 through 5 should then be repeated 200 times to form 200 replicate samples.

This method uses random switching between PSU bootstrap sampling and SSU bootstrap sampling to represent the proper mix (in expectation) of the first- and second-stage variance components when an FPC is applied at the first stage of sampling. The general method described by Flyer (1987) for half-sample replication is extended 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 was added to the analysis file. These weights are zero for units not selected in a particular bootstrap sample; weights for other units are inflated for the bootstrap subsampling. The initial analytic weights for the complete sample are also included for the purpose of computing the desired estimates. The vector of replicate weights allows for computation of additional estimates for the sole purpose of estimating a variance. Assuming *B* sets of replicate weights, the variance of any estimate, $\hat{\theta}$, can be estimated by replicating the estimation procedure for each replicate and computing a simple variance of the replicate estimates, as follows:

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

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.

Once the replicate weights are provided, this estimate can be produced by most survey software packages (e.g., SUDAAN, RTI International [2008], computes this estimate by invoking the DESIGN = BRR option). (For an example of SUDAAN code, see appendix L)

The number of replicate weights was set to 200 on the basis of the requirements of the National Center for Education Statistics (NCES) Data Analysis System (DAS). For the 200 replicate weights included on the analysis file (WTA001–WTA200), the poststratification process was repeated so that the variance would account for the poststratification weight adjustment. For a few of the replicates, one or two of the control totals could not be met, because of model convergence problems (i.e., there was no solution to satisfy all model equations simultaneously).

6.4.3 Variance Approximation

The survey design effect for a statistic is defined as the ratio of the design-based variance estimate over the variance estimate that would have been obtained from a simple random sample of the same size (if that were practical). It is often used to measure the effects that sample design features have on the precision of survey estimates. For example, stratification tends to decrease the variance, but multistage sampling and unequal sampling rates usually increase the variance. In addition, weight adjustments for nonresponse (performed to reduce nonresponse bias) and poststratification increase the variance by increasing the weight variation. Because of these effects, most complex multistage sampling designs, like NPSAS:08, result in design effects greater than 1.0. That is, the design-based variance is larger than the simple random sample variance.

Specifically, the survey design effect for a given estimate, $\hat{\theta}$, is defined as

$$\text{Deff}(\hat{\theta}) = \frac{Var_{design}(\hat{\theta})}{Var_{srs}(\hat{\theta})}.$$

The square root of the design effect is another measure, which can also be expressed as the ratio of the standard errors, or

$$\text{Deft}(\hat{\theta}) = \frac{SE_{design}(\hat{\theta})}{SE_{srs}(\hat{\theta})}.$$

In appendix M, design effect estimates are presented for important survey domains and estimates among undergraduate students, graduate students, and first-professional students, in order to summarize the effects of stratification, multistage sampling, unequal probabilities of selection, and the weight adjustments. These design effects were estimated with the use of SUDAAN and the bootstrap variance estimation procedure described in section 6.4.2 and appendix L. If one must perform a quick analysis of NPSAS:08 data without using one of the software packages for analysis of complex survey data, the design effect tables in appendix N can be used to make approximate adjustments to the standard errors of survey statistics computed with the standard software packages that assume simple random sampling designs; however, one cannot be confident about the actual design-based standard errors without performing the analysis with one of the software packages specifically designed for analysis of data from complex sample surveys. (For details about the use of such software packages, see appendix L.)

Large design effects imply large standard errors and relatively poor precision. Small design effects imply small standard errors and good precision. In general terms, a design effect less than 2.0 is low, from 2.0 to 3.0 is moderate, and greater than 3.0 is high. Moderate and high design effects often occur in complex surveys such as NPSAS. Unequal weighting causes large design effects and is often due to nonresponse and poststratification adjustments; however, in NPSAS, the unequal weighting is also due to the sample design and different sampling rates between institution strata, as well as to the different sampling rates between student strata.

References

Breiman, L., Friedman, J., Olshen, R., and Stone, C. (1984). *Classification and Regression Trees*. New York: Chapman & Hall.

Chromy, J.R. (1979). Sequential Sample Selection Methods. In *Proceedings of the Section on Survey Research Methods, American Statistical Association* (pp. 401–406). Alexandria, VA: American Statistical Association.

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*. New York: Academic Press.

Cominole, M., Riccobono, J., Siegel, P., Caves, L., and Rosen, J. (2008). *2008 National Postsecondary Student Aid Study (NPSAS:08) Field Test Methodology Report* (NCES 2008-01). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Cox, B. (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.

Creel, D., and Krotki, K. (2006). Creating Imputation Classes Using Classification Tree Methodology. In *Proceedings of the Section on Survey Research Methods, American Statistical Association*. [CD-ROM]. Alexandria, VA: American Statistical Association.

Deville, J., and Särndal, C.-E. (1992). Calibration Estimators in Survey Sampling. *Journal of the American Statistical Association, 87*: 376–382.

Education Sciences Reform Act of 2002, 20 U.S.C. § 3419. (2007).

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

Folsom, R.E., and Singh, A.C. (2000). The Generalized Exponential Model for Sampling Weight Calibration for Extreme Values, Nonresponse, and Poststratification. In *Proceedings of the Section on Survey Research Methods, American Statistical Association* (pp. 598–603). Alexandria, VA: American Statistical Association.

General Education Provisions Act, 20 U.S.C. §§ 9541 to 9548. (2007).

Hanley, J.A., and McNeil, B.J. (1982). The Meaning and Use of the Area Under a Receiver Operating Characteristic Curve. *Diagnostic Radiology, 143*(1): 29–36.

Higher Education Act of 1965, *as amended by* the Higher Education Amendments of 1986, 20 U.S.C. § 1070 et seq. (2007).

Higher Education Act of 1965, *as amended by* the Higher Education Opportunity Act of 2008, 20 U.S.C. § 1015. (2008).

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

Iannacchione, V.G. (2003). Sequential Weight Adjustments for Location and Cooperation Propensity for the 1995 National Survey of Family Growth. *Journal of Official Statistics, 19*(1): 31–43.

Kass, G.V. (1980). An Exploratory Technique for Investigating Large Quantities of Categorical Data. *Applied Statistics, 29*(2): 119–127.

Kott, P.S. (1988). Model-Based Finite Population Correction for the Horvitz-Thompson Estimator. *Biometrika, 75*(4): 797–799.

Krotki, K., Black, S., and Creel, D. (2005). Mass Imputation. In *Proceedings of the Section on Survey Research Methods, American Statistical Association* [CD-ROM]. Alexandria, VA: American Statistical Association.

National Education Statistics Act of 1994, 20 U.S.C. §§ 9541 to 9547 and 9573. (2007).

RTI International. (2008). *SUDAAN User's Manual, Release 10.0*. Research Triangle Park, NC: Author.

SAS Institute Inc. (2006). *SAS 9.1.3 Output Delivery System: User's Manual* (Vols. 1–2). Cary, NC: Author.

The College Board. (2008). *Trends in Student Aid*. Trends in Higher Education Series. Retrieved December 18, 2009, from http://professionals.collegeboard.com/profdownload/trends-in-student-aid-2008.pdf.

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. (2003). NCES Statistical Standards, *NCES 2003-601*. Washington, DC: Author.

U.S. Department of Education, National Center for Education Statistics. (1999). NCES Restricted-Use Data Procedures Manual. Washington, DC: Author.

Woodruff, R.S. (1971). A Simple Method for Approximating the Variance of a Complicated Estimate. *Journal of the American Statistical Association, 66*(334): 411–414.