Chapter 6
Weighting and Variance Estimation

Statistical analysis weights were computed for two sets of respondents: CATI respondents and study respondents. (They were not computed separately for CADE respondents because it was expected that analysis of any items collected in CADE would be based on the larger set of study respondents.) The statistical analysis weights compensated for unequal sampling rates and differential propensities to respond. CATI, CADE, and study respondents were defined as follows:

CATI respondent: any sample member who

- completed at least Section A of the CATI interview or
- completed an abbreviated (telephone or paper copy) interview.

CADE respondent: any sample member for whom

- the CADE financial aid gate question was answered, AND
- the CADE enrollment section had some enrollment data provided, AND
- the CADE student characteristics section had at least one valid response for the set of items: date of birth; marital status; race; and sex. If the case was a CPS match, it was considered it to have successfully met this criterion.

Study respondent: any sample member who was

- a CATI respondent and/or
- a CADE respondent.

6.1 Study and CATI Weight Components

Weights were computed first for study respondents (STUDYWT) as the product of the following 13 weight components:

(1) Adjustment for Field Test Sampling (WT1)
(2) Institution Sampling Weight (WT2)
(3) Adjustment for Institution Multiplicity (WT3)
(4) Institution Poststratification Adjustment (WT4)
(5) Adjustment for Institution Nonresponse (WT5)

(6) Student Sampling Weight (WT6)
(7) Student Subsampling Weight (WT7)
(8) Adjustment for Students Never Sent to CATI (WT8)
(9) Adjustment for Student Multiplicity (WT9)
(10) Adjustment for Unknown Eligibility Status (WT10)
(11) Weight Trimming Adjustment (WT11)
(12) Adjustment for Study Nonresponse (WT12)
(13) Poststratification Adjustment for Study Respondents (WT13).

These study weights were used as the base for CATI weights. The CATI weights (CATIWT) were the product of the study weights and the following four additional weight components:

(14) Adjustment for Not Locating Students (WT14)
(15) Adjustment for CATI Refusals (WT15)
(16) Adjustment for Other CATI Nonresponse (WT16)
(17) Poststratification Adjustment for CATI Respondents (WT17)

The study weights and the CATI weights are the two statistical analysis weights on the analysis files. Each weight component is described below and represents either a probability of selection or a weight adjustment. The weight adjustments included nonresponse and poststratification adjustments to compensate for potential nonresponse bias and frame errors. All nonresponse adjustment and poststratification models were fit using RTI's proprietary generalized exponential models (GEMs),[1] which are similar to logistic models using bounds for adjustment factors. Also, multiplicity and trimming adjustments were performed. Each of these 17 weighting components is described in more detail below.

(1) Adjustment for Field Test Sampling (WT1)

The NPSAS field test sample was selected using stratified simple random sampling, so these sample institutions were deleted from the full-scale institution sampling frame without compromising population coverage. Each institution on the sampling frame received a first-stage sampling weight based on the probability that it was *not* selected for the field test.

The institutions in stratum r on the institution sampling frame were partitioned as follows. Let $j = 1, 2, \ldots, J_1(r)$ represent those institutions not on the frame from which the field test sample was selected (near certainty and new IPEDS 1998–99 institutions).

- Let $j=J_1(r)+1, J_1(r)+2, \ldots, J_2(r)$ represent those that were on the frame for the field test but were not selected.

- Let $j=J_2(r)+1, J_2(r)+2, \ldots, J(r)$ represent the institutions in the simple random sample of $n_f(r)$ institutions selected for the field test.

[1] R.E Folsom. and A.C. Singh (2000). "The Generalized Exponential Model for Sampling Weight Calibration for Extreme Values, Nonresponse, and Poststratification." *Proceedings of the Section on Survey Research Methods of the American Statistical Association,* pp. 598–603.

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

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

(2) Institution Sampling Weight (WT2)

The sampling weight for each sample institution was the reciprocal of its probability of selection. As noted earlier in chapter 2, the probability of selection for institution *i* was

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

Therefore, the institution sampling weight was assigned as follows:

$WT2 = 1 / \pi_r(i)$.

(3) Adjustment for Institution Multiplicity (WT3)

During institution recruitment, six sample schools that had two or three records listed on the IPEDS frame were found. In most cases, it was caused by schools that had recently merged. If two records were sampled, then one record was retained for tracking survey results and the other record was classified as ineligible.

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 or B) = P(A) + P(B) - P(A)P(B).

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

NEW_WT2 = 1 / P(A 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 or B 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:

NEW_WT2 = 1 / P(A or B or C).

Finally, the multiplicity adjustment factor was derived by dividing the new sampling weight by the old sampling weight,

WT3 = NEW_WT2 / WT2,

for the institutions with positive multiplicity, and setting it to unity (1.00) for all other institutions. Hence, the product of WT2 and WT3 equals NEW_WT2 for the institutions with positive multiplicity and equals WT2 for all other institutions.

(4) Institution Poststratification Adjustment (WT4)

To ensure population coverage, the sampling weights were adjusted to control totals for enrollment using a weighting class adjustment. Institution type and size were used to define the weighting classes. The weight adjustment factor was the ratio of the population enrollment to the sample total of the weight multiplied by the enrollment within weighting classes:

$$PS_c = \frac{\sum_{i \varepsilon Pop(c)} E_i}{\sum_{i \varepsilon Samp(c)} W_i \bullet E_i}$$

where

c = the weighting class,

W_i = the cumulative institution weight (WT1 • WT2 • WT3), and

E_i = the institution's enrollment from the sampling frame.

Table 6-1 presents the weight adjustment factors for each weighting class.

Table 6-1.—Weight adjustment factors for institution poststratification and nonresponse

Weighting class (institution sector and size[1])	Number of respondents	Weighted response rate	Post-stratification weight adjustment factor (WT4)	Nonresponse weight adjustment factor (WT5)
Total	1,082	94.0	†	†
Public less than 2-year	34	89.9	1.10	1.11
Public 2-year, small	99	97.9	1.08	1.02
Public 2-year, large	99	90.1	1.07	1.11
Public 4-year non-doctorate-granting, small	63	95.1	1.13	1.05
Public 4-year non-doctorate-granting, large	64	98.4	0.99	1.02
Public 4-year doctorate-granting, small	110	92.8	1.09	1.08
Public 4-year doctorate-granting, large	110	96.1	1.04	1.04
Private not-for-profit less-than-4-year	35	93.7	1.06	1.07
Private not-for-profit 4-year, non-doctorate-granting, small	86	89.4	1.04	1.12
Private not-for-profit 4-year, non-doctorate-granting, large	87	89.0	1.15	1.12
Private not-for-profit 4-year doctorate-granting, small	84	92.9	1.20	1.08
Private not-for-profit 4-year doctorate-granting, large	84	93.2	1.07	1.07
Private for-profit 2-year, small	38	91.7	1.26	1.09
Private for-profit 2-year, large	39	86.5	1.09	1.16
Private for-profit 2-year-or-more	50	95.8	1.03	1.04

†Not applicable.

[1] Size for poststratification weighting classes was based on the median enrollment within sector for the institutions on the sampling frame. Size for nonresponse weighting classes was based on the median enrollment within the sector for the sample institutions. Three of the sectors had too few responding institutions to split by size.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Postsecondary Student Aid Study, 1999–2000 (NPSAS:2000).

(5) Adjustment for Institution Nonresponse (WT5)

For weighting purposes, a school was considered a responding school if it provided an enrollment list and if at least one student from the institution was a study respondent. A weighting class adjustment was performed to compensate for nonresponding institutions, using institution type and size as the weighting classes. The calculated response rates were enhanced by multiplying the institution's weight by enrollment:

$$R_c = \frac{\sum_{i \varepsilon Resp(c)} E_i}{\sum_{i \varepsilon Elig(c)} W_i \bullet E_i}$$

where

c = the weighting class,

W_i = the cumulative institution weight (WT1 • WT2 • WT3 • WT4), and

E_i = the institution's enrollment.

The weight adjustment was then the reciprocal of this response rate. This enhancement forced the estimated total enrollment to be the same for the responding institutions as it was for the eligible institutions, and thus for the population since we poststratified to population totals. Table 6-1 presents the response rates and the resulting adjustment factors by institution type and size.

(6) Student Sampling Weight (WT6)

The overall student sampling strata were defined by crossing the institution sampling strata with the student strata within institutions. The overall sampling rates for these sampling strata can be found in appendix G. 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 sampling rate for student stratum *s* within institution *i* was calculated as the overall 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

π_r (i) = the institution's probability of selection.

As discussed in appendix G, 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 based on 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 did not exceed by more than 50 students the expected sample size of the institution based on the IPEDS data. A minimum sample size constraint of 40 students also was imposed so that at least 30 respondents from each participating institution could be expected.

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

$WT6 = 1 / f_{s|i}$.

(7) Student Subsampling Weight (WT7)

When schools provided hard-copy lists for student sampling, they often did not provide separate lists by strata (e.g., undergraduate and graduate students were on the same list). When that happened, the combined list was sampled at the highest of the sampling rates for the strata contained within the list. After the original sample was keyed, strata with the lower sampling rates were then subsampled to achieve the desired sampling rates. The student subsampling weight adjustment factor, WT7, was the reciprocal of this subsampling rate. This weight factor was unity (1.00) for most students because this subsampling was not necessary for most institutions.

(8) Adjustment for Students Never Sent to CATI (WT8)

To speed up data collection, some students were sent to CATI before CADE data were abstracted from the institution. This could be done when locating information or a Social Security number was available for the student from the enrollment file or from CPS. However, potentially eligible students were never sent to CATI if such information was unavailable or if the institution refused to provide CADE data before the decision to send the institution's students to CATI.[2] To adjust for students from responding institutions who were never sent to CATI, a weighting class adjustment was performed using the 22 institution strata as weighting classes. Table 6-2 presents the weight adjustment factors.

(9) Adjustment for Student Multiplicity (WT9)

Students who attended more than one eligible institution during the 1999–2000 academic year had multiple chances of being selected. That is, they could have been selected from any of the institutions they attended. Therefore, these students had a higher probability of being selected than was represented in their sampling weight. This multiplicity was adjusted 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

$$WT9 = 1 / M,$$

where M is the multiplicity, or number of institutions attended. The multiplicity was determined from the CATI interview, the Pell Grant payment file, and the National Student Loan Data System. Unless there was evidence to the contrary, the student multiplicity was presumed to be unity (1.00).

[2] If the institution had no study respondents, then the institution was considered a nonrespondent, which was handled through the institution nonresponse adjustment.

Table 6-2.—Weight adjustment factors for students never sent to CATI

Weighting class (institution stratum)	Number sent to CATI	Weight adjustment factor (WT8)
Total	69,595	†
Public less than 2-year	1,525	1.00
Public 2-year	10,663	1.00
Public 4-year non-doctorate-granting		
Bachelor's high education	302	1.00
Bachelor's low education	1,026	1.00
Master's high education	2,087	1.00
Master's low education	6,463	1.00
Public 4-year doctorate-granting		
Doctorate-granting high education	2,249	1.00
Doctorate-granting low education	5,631	1.00
First-professional high education	3,993	1.00
First-professional low education	9,653	1.02
Private not-for-profit less-than-2-year	563	1.02
Private not-for-profit 2-year	1,175	1.00
Private not-for-profit 4-year, non-doctorate-granting		
Bachelor's high education	889	1.00
Bachelor's low education	1,610	1.00
Master's high education	1,567	1.02
Master's low education	3,826	1.01
Private not-for-profit 4-year doctorate-granting		
Doctorate-granting high education	741	1.00
Doctorate-granting low education	1,386	1.00
First-professional high education	3,248	1.00
First-professional low education	4,010	1.01
Private for-profit less-than-2-year	4,399	1.02
Private for-profit 2-year or more	2,589	1.00

†Not applicable.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Postsecondary Student Aid Study, 1999–2000 (NPSAS:2000).

(10) Adjustment for Unknown Eligibility Status (WT10)

Some students were determined to be ineligible while the student record data were being abstracted using CADE. We did not attempt to interview these students, and they received a weight of zero. Students were sent to CATI if they were not classified as ineligible, and their final eligibility status was then determined from the CATI interviews. However, for the students

whom RTI staff were unable to contact, the final eligibility status could not be determined. These students were treated as eligible, their weights were adjusted to compensate for the small portion of students who were actually ineligible (as described below), and they were included in the analysis files.

Weighting classes were defined by the cross of institution type and the students' matching status to financial aid files (CPS, Pell, and loan). Table 6-3 presents the weight adjustment factors applied to the students with unknown eligibility. These weight adjustment factors were simply the eligibility rate estimated among students with known eligibility status. For the eligible students, the weight adjustment factor was set equal to one.

(11) Weight Trimming Adjustment (WT11)

Some of the student sampling weights were initially large because student sampling rates were fixed and sometimes very small. Also, the cumulative effect of the adjustment factors could cause these large weights to increase further. These very large weights could cause excessive weight variation, which results in inflated sampling variances and mean square errors.

The mean square error of an estimate, $\hat{\theta}$, is defined as the expected value of the squared total error, or

$$MSE(\hat{\theta}) = E(\theta - \hat{\theta})^2 .$$

This can be rewritten as

$$MSE(\hat{\theta}) = E[(\hat{\theta} - E(\theta)]^2 + [E(\hat{\theta}) - (\theta)]^2 ,$$

where the first term is the sampling variance and the second term is the bias squared.

It was usually possible, by truncating some of the largest weights and smoothing (distributing) the truncated portions over all the weights, to reduce the mean square error by substantially reducing the variance and slightly increasing the bias in the weights. However, the subsequent nonresponse and poststratification adjustments reduced the bias.

To evaluate the weight variation, the unequal weighting effects on the variance were computed for the ultimate strata defined by the cross of institution type and student type, as follows:

$$UWE = n\Sigma w^2 / (\Sigma w)^2.$$

When the large sampling weights and the cumulative effect of the weight adjustment factors caused the unequal weighting effects to be unreasonably large, an upper limit was established for truncation of the largest weights. To distribute the truncated portions, a smoothing adjustment ratio was calculated as the sum of the original weights over the sum of the truncated weights for each class, as follows.

Table 6-3.—Weight adjustment factors for unknown student eligibility status

Weighting class (institution level, by student type, by matching status to financial aid files)	Number adjusted for unknown eligibility	Weight adjustment factor (WT10)
Total	12,543	†
Public less than 2-year		
Matched Pell or Stafford file	81	0.85
Matched CPS file only	32	0.80
No matches	177	0.57
Public 2-year		
Matched Pell or Stafford file	492	0.93
Matched CPS file only	222	0.85
No matches	1,319	0.79
Public 4-year non-doctorate-granting		
Undergraduates: Matched Pell or Stafford file	566	0.97
Matched CPS file only	112	0.90
No matches	662	0.85
Graduates: Matched Pell or Stafford file	24	0.99
Matched CPS file only	4	0.87
No matches	132	0.88
Public 4-year doctorate-granting		
Undergraduates: Matched Pell or Stafford file	1,092	0.98
Matched CPS file only	219	0.93
No matches	1,399	0.91
Graduates: Matched Pell or Stafford file	220	0.99
Matched CPS file only	19	0.87
No matches	681	0.91
Private not-for-profit less-than-4-year		
Matched Pell or Stafford file	264	0.95
Matched CPS file only	36	0.85
No matches	132	0.70
Private not-for-profit 4-year, non-doctorate-granting		
Undergraduates: Matched Pell or Stafford file	577	0.97
Matched CPS file only	91	0.87
No matches	447	0.85
Graduates: Matched Pell or Stafford file	40	0.95
Matched CPS file only	9	0.93
No matches	97	0.92
Private not-for-profit 4-year doctorate-granting		
Undergraduates: Matched Pell or Stafford file	405	0.98
Matched CPS file only	71	0.82
No matches	430	0.85
Graduates: Matched Pell or Stafford file	199	0.99
Matched CPS file only	25	0.84
No matches	459	0.85

Table 6-3.—Weight adjustment factors for unknown student eligibility status —Continued

Weighting class (institution level, by student type, by matching status to financial aid files)	Number adjusted for unknown eligibility	Weight adjustment factor (WT10)
Private for-profit less-than-2-year	874	0.94
Matched Pell or Stafford file	139	0.68
Matched CPS file only	200	0.76
No matches		
Private for-profit 2-year		
Matched Pell or Stafford file	225	0.94
Matched CPS file only	29	0.64
No matches	64	0.60
Private for-profit 4-year		
Undergraduates: Matched Pell or Stafford file	102	0.97
Matched CPS file only	11	0.88
No matches	110	0.79
Graduates: Matched Pell or Stafford file	18	0.99
Matched CPS file only/ No matches combined	36	0.96

†Not applicable.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Postsecondary Student Aid Study, 1999–2000 (NPSAS:2000).

$$S_c = \frac{\sum_{i \varepsilon c} W_O(i)}{\sum_{i \varepsilon c} W_T(i)}$$

where

$W_O(I)$ = the original weight (WT1•WT2•...WT10), and

$W_T(I)$ = the truncated weight (the minimum of the original weight and the upper limit).

The truncation and smoothing steps were then combined into one adjustment factor by defining the weight component as

$$WT13 = \frac{W_T(i)}{W_O(i)} \bullet S_c \ .$$

(12) Adjustment for Study Nonresponse (WT12)

The first type of adjustment for student nonresponse was adjustment for study nonresponse, i.e., insufficient CADE or CATI data. These weight adjustments were made to compensate for the potential study nonresponse bias. Adjustment factors were inverses of predicted response propensities derived from a logistic regression model. The logistic

procedure, developed by Folsom,[3] adjusts the weights of respondents so that the adjusted weight sums of respondents reproduce the unadjusted weight sums of respondents and nonrespondents for the categorical predictor variables included in the model. To avoid excessive weight variation, the procedure also constrains the adjustment factors to be within specified lower and upper bounds.

Candidate predictor variables were chosen that were thought to be predictive of response status and were nonmissing for both study respondents and nonrespondents. The candidate predictor variables included

- institution type,
- Region,
- institution enrollment from IPEDS IC file (categorical),
- student type,
- Social Security number indicator,
- CPS record indicator,
- Pell grant status,
- Pell grant amount (categorical),
- Stafford Loan status,
- Stafford Loan amount (categorical), and
- federal aid receipt status.

To detect important interactions for the logistic models, a Chi-squared automatic interaction detector analysis was performed on the predictor variables. The CHAID analysis divided the data into segments that differed with respect to the response variable, study response. The segmentation process first found the variable that was the most significant predictor of response within each category or collapsed set of categories of this variable, it looked for the next most significant predictor of response. This process continued until no more statistically significant predictors were found (or until some other stopping rule was met). The interactions from the final CHAID segments were then defined from the final nesting of the variables.

The interaction segments and all the main effect variables were then subjected to variable screening in the logistic procedure. Variables significant at the 15 percent level were retained, with the exception of institution type and student type, which were retained regardless of their significance.

From the logistic models, the predicted probability that student *j* was a study respondent was given by

$$\hat{p}_{rj} = \left[1 + \exp(-\boldsymbol{x}_j \beta)\right]^{-1},$$

where

$\mathbf{x}_j$ = the row vector of predictor variables, and

[3] Folsom, R.E. (1991). "Exponential and Logistic Weight Adjustments for Sampling and Nonresponse Error Reduction." *Proceedings of the Social Statistics Section of the American Statistical Association*, pp. 197–202.

B = the column vector of regression coefficients.

The logistic adjustment factor is then simply the reciprocal of this predicted probability of being a student respondent, or

$$WT12 = 1/\hat{p}_{rj} .$$

Table 6-4 presents the final predictor variables used in the logistic model 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: 1.03
- maximum: 1.71.

(13) Poststratification Adjustment for Study Respondents (WT13)

To ensure population coverage, the study weights were further adjusted to control totals with a generalized raking procedure that derived adjustment factors from an exponential regression model.[4] The algorithm for this procedure was similar to the algorithm used in the logistic procedure for the nonresponse adjustments.

Control totals were established for annual student enrollment, by institution type; total number of Pell Grants awarded; amount of Pell Grants awarded, by institution type; and amount of Stafford Loans awarded, by institution type.

The annual enrollment control totals were estimated by multiplying the "known" fall enrollment totals from the 1997–98 Fall Enrollment Survey[5] by the estimated ratio (based on NPSAS:2000 data) of annual enrollment over fall enrollment. Specifically, the annual enrollment control totals were computed as

$$A_{control} = \frac{A_{npsas}}{F_{npsas}} \bullet F_{known} ,$$

[4] R.E. Folsom. "Exponential and Logistic Weight Adjustments for Sampling and Nonresponse Error Reduction." *Proceedings of the Social Statistics Section of the American Statistical Association,* 1991, 197–202.

[5] The 1997–98 Fall Enrollment Survey was used to estimate fall enrollment since that is what was available on the sampling frame. The IPEDS fall 1999 enrollments were not imputed, so they would not provide reliable estimates. It was determined that using fall 1997 estimates was sufficient since fall enrollments did not change significantly over this period.

Table 6-4.—Average weight adjustment factors from logistic model used to adjust study weights for student nonresponse

Logistic model predictor variables	Number of respondents	Weighted response rate	Average weight adjustment factor (WT12)
Total	61,770	97.1	1.03
Institutional sector			
Public less-than-2-year	1,060	95.4	1.04
Public 2-year	8,930	97.2	1.03
Public 4-year non-doctorate-granting	8,950	97.0	1.03
Public 4-year doctorate-granting	19,730	97.1	1.03
Private not-for-profit 2-year or less	1,510	98.4	1.02
Private not-for-profit 4-year, non-doctorate-granting	7,190	97.2	1.03
Private not-for-profit 4-year doctorate-granting	8,410	97.4	1.03
Private for-profit less-than-2-year	3,630	93.2	1.07
Private for-profit 2-year	1,170	97.7	1.02
Private for-profit 4-year	1,170	99.6	1.00
Region			
New England	3,580	98.7	1.01
Great Lakes	10,000	98.7	1.01
Plains	4,660	98.7	1.01
Rocky Mountains	2,460	99.8	1.00
AK, HI, PR	1,660	96.7	1.02
Other	39,410	96.3	1.04
Student type			
Baccalaureate, business major	1,330	96.0	1.04
Baccalaureate, other major	13,710	97.8	1.02
Other undergraduate	35,510	97.2	1.03
Master's	5,370	97.4	1.03
Doctor's	3,450	94.2	1.06
Other graduate	1,190	96.6	1.03
First-professional	1,200	95.5	1.05
SSN preloaded			
Yes	59,750	97.2	1.03
No	2,020	94.8	1.05
CHAID segments			
1 = No CPS match, SSN not preloaded, New England	110	96.8	1.04
2 = No CPS match, SSN not preloaded, Mid East	380	94.2	1.07
3 = No CPS match, SSN not preloaded, Great Lakes, Plains	280	99.5	1.01
4 = No CPS match, SSN not preloaded, Southeast	210	86.7	1.16
5 = No CPS match, SSN not preloaded, Southwest, Rocky Mountains, Far West	280	98.6	1.02
6 = No CPS match, SSN not preloaded, AK, HI, PR	50	61.3	1.63
7 = No CPS match, SSN preloaded, ENTOTCAT=3,4	17,170	96.7	1.04
8 = CPS match, AK, HI, PR, enrollment <= 3,267	520	100.0	1.00
9 = CPS match, New England, 3267 < enrollment <24,120	1,000	100.0	1.00
10 = CPS match, Rocky Mountains, 3267 < enrollment <24,120	590	100.0	1.00
11 = CPS match, AK, HI, PR, 3267 < enrollment <24,120	620	100.0	1.00
12 = CPS match, New England, enrollment > 24,120	200	100.0	1.00
13 = CPS match, Plains, enrollment > 24,120	400	99.9	1.00
14 = CPS match, Southeast, enrollment > 24,120	1,270	90.1	1.11
15 = CPS match, Southwest, Rocky Mountains, Far West, AK, HI, PR, enrollment > 24,120	2,480	99.7	1.00
16 = Other	36,210	97.4	1.03

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Postsecondary Student Aid Study, 1999–2000 (NPSAS:2000).

where

$A_{control}$ = annual enrollment control total,

A_{npsas} = annual enrollment estimated from NPSAS:2000,

F_{npsas} = fall enrollment estimated from NPSAS:2000, and

F_{known} = fall enrollment from the 1997–98 Fall Enrollment Survey.

The exponential adjustment satisfies the following constraints:

$$\sum_j W_j \lambda_j \boldsymbol{x}_j^T = \eta_O^T \ ,$$

where

W_j = the cumulative weight (WT1•WT2•....•WT12),

λ_j = $\exp(\alpha + \mathbf{x}_j B)$,

α = model intercept

β = vector of parameters that specify the nature of the relationship between λ_j and $\mathbf{x}_j$

$\mathbf{x}_j$ = the vector of regressors associated with the domains to be controlled, and

η_o = the set of control totals.

The exponential adjustment factor for student *j* is then simply

WT13 = λ_j .

Tables 6-5 and 6-6 present the average weight adjustment factor for each variable in the model. Table 6-5 presents the variables associated with the student enrollment control totals and the average weight adjustment factors by these variables. Similarly, table 6-6 presents the variables associated with the Pell Grant and Stafford Loan control totals and the average weight adjustment factors. The weight adjustment factors from the exponential adjustment are summarized below, and met the following constraints:

- minimum: 0.53
- median: 0.99
- maximum: 2.36.

Table 6-5.—Average weight adjustment factors from exponential models for Poststratifying to student enrollment totals

Exponential model variable	Fall enrollment from 1997–1998 fall enrollment survey	Ratio of NPSAS:2000 annual over fall enroll-ment	Control total for annual enrollment[1]	Average weight adjustment factor (WT13)	Average weight adjustment factor (WT17)
Student type					
Undergraduate	†	†	16,538,472	†	1.00
Graduate	†	†	2,332,233	†	1.00
First-professional	†	†	325,301	†	1.00
Institutional sector					
Public less-than-2-year	84,498	1.33	112,533	2.08	0.99
Public 2-year	5,378,376	1.41	7,568,455	1.09	1.00
Public 4-year non-doctorate-granting	1,935,294	1.19	2,307,422	1.00	1.00
Public 4-year doctorate-granting	4,011,997	1.16	4,657,446	1.01	1.00
Private not-for-profit 2-year or less	104,077	1.30	135,742	1.25	1.00
Private not-for-profit 4-year, non-doctorate-granting	1,478,483	1.18	1,738,463	0.92	1.00
Private not-for-profit 4-year doctorate-granting	1,546,883	1.15	1,780,664	0.94	1.00
Private for-profit less-than-2-year	164,123	2.01	329,751	0.92	1.01
Private for-profit 2-year	227,659	1.40	318,488	0.89	1.01
Private for-profit 4-year	190,371	1.30	247,043	0.75	1.02

† Not applicable.

[1] Control total is not the exact product of the fall enrollment from 1995–1996 fall enrollment survey and the ratio of NPSAS:2000 annual over fall enrollment, due to rounding of the ratio.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Postsecondary Student Aid Study, 1999–2000 (NPSAS:2000).

After this weight adjustment was performed, the final study weights (STUDYWT) were computed as the product of the 13 weight components and then rounded to the nearest integer.

(14) Adjustment for Not Locating Students (WT14)

The final (unrounded) study weights were further adjusted to produce the CATI analysis weights. The adjustment for CATI nonresponse was performed in three stages because the predictors of response propensity were potentially different at each stage:

- inability to locate the student,
- refusal to be interviewed, and
- other non-interview.

Using these three 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.

Table 6-6.—Average weight adjustment factors from exponential model for poststratifying to Pell grant and Stafford loan control totals

Exponential model variable	Control total	Average weight adjustment factor (WT13)	Average weight adjustment factor (WT17)
Pell grants			
Total *number* awarded	3,759,000	1.00	1.01
Total *dollars* awarded			
Public 4-year	2,771,723,587	1.01	1.01
Public 2-year	2,156,165,970	1.15	0.98
Private not-for-profit 4-year	1,223,434,200	0.87	1.01
Private not-for-profit 2-year	103,619,419	1.08	1.02
Private for-profit	927,331,131	0.98	1.03
Stafford Loans			
Total *dollars* awarded – study weights			
Undergraduate			
Public 4-year	9,812,004,437	1.06	†
Public 2-year	1,594,864,801	1.03	†
Private not-for-profit 4-year	6,084,095,282	0.98	†
Private not-for-profit 2-year	201,342,429	1.04	†
Private for-profit	3,269,427,995	1.08	†
Graduate/first-professional			
Public 4-year	4,238,972,034	1.04	†
Public 2-year	5,071,137	0.61	†
Private not-for-profit 4-year	6,285,676,620	1.03	†
Private not-for-profit 2-year	†	†	†
Private for-profit	377,462,273	0.93	†
Total *dollars* awarded — CATI weights			
Public 4-year	14,050,976,471	†	1.00
Public 2-year	1,599,935,938	†	0.96
Private not-for-profit 4-year	12,369,771,902	†	1.01
Private not-for-profit 2-year	201,342,429	†	0.98
Private for-profit	3,646,890,268	†	0.99

† Not applicable.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Postsecondary Student Aid Study, 1999–2000 (NPSAS:2000).

The same logistic regression procedure used to adjust for study nonresponse (WT12) was again used to adjust for inability to locate (contact) the student. Candidate predictor variables were chosen that were thought to be predictive of CATI nonresponse and were missing for 5 percent or fewer of all study respondents. The candidate predictor variables included

- age (categorical),
- any aid receipt indicator,
- fall attendance status,
- citizenship,
- CPS record indicator,
- institution enrollment from IPEDS IC file (categorical),
- fall enrollment status,
- federal aid receipt indicator,
- sex,

- Hispanic indicator,
- institutional aid receipt indicator,
- OBE region,
- student date of birth preloaded into CATI,
- parent data preloaded into CATI,
- total number of phone numbers obtained for student,
- Social Security number indicator,
- Pell Grant status,
- Pell Grant amount (categorical),
- Stafford Loan status,
- Stafford Loan amount (categorical),
- institution type,
- state aid receipt indicator,
- number of institutions attended in 1999–2000, and
- student type.

Other variables that were considered but not included because they were missing for more than 5 percent of all study respondents included

- dependents indicator,
- dependency status,
- number of dependents,
- full-year attendance status,
- high school degree indicator and type,
- high school graduation year,
- local residence,
- parents' income,
- parents' family size,
- parent's marital status,
- student's marital status,
- student's income, and
- race.

As in the study 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 subjected to variable screening in the logistic procedure. Variables significant at the 15 percent significance level were retained, with the exception of institution type, student type, Pell Grant status, and Stafford Loan status, which were retained regardless of the significance level.

Table 6-7 presents the final predictor variables used in the logistic model to adjust the CATI weights and the average weight adjustment factors resulting from these variables. As in the study nonresponse adjustment, the weighting adjustment factor for student *j* was the reciprocal of the predicted response probability, or

$$WT14 = 1/\hat{p}_{rj} \ .$$

Table 6-7.—Average weight adjustment factors from logistic model used to adjust CATI weights for student location nonresponse

Logistic model predictor variables	Number of located respondents	Weighted response rate	Average weight adjustment factor (WT14)
Total	50,764	82.7	1.19
Institutional sector			
Public less-than-2-year	850	83.8	1.19
Public 2-year	7,062	81.5	1.22
Public 4-year non-doctorate-granting	7,578	84.9	1.16
Public 4-year doctorate-granting	16,554	83.6	1.18
Private not-for-profit 2-year or less	1,120	77.6	1.29
Private not-for-profit 4-year, non-doctorate-granting	6,064	83.7	1.18
Private not-for-profit 4-year doctorate-granting	7,077	84.4	1.17
Private for-profit less-than-2-year	2,676	75.7	1.31
Private for-profit 2-year	882	77.9	1.28
Private for-profit 4-year	901	78.9	1.25
Region			
Southwest	5,348	79.2	1.24
AK, HI, PR	1,147	71.4	1.42
Other	44,269	83.4	1.18
Student type			
Confirmed baccalaureate	11,803	86.8	1.15
Other undergraduate	28,854	81.7	1.22
Graduate	9,075	86.1	1.16
First-professional	1,032	86.7	1.15
Age group			
Less than 30	36,430	81.3	1.21
30 or older	14,334	85.9	1.15
Sex			
Male	21,007	81.1	1.21
Female	29,757	83.9	1.18
Received institutional aid			
Yes	11,647	85.2	1.16
No	39,117	82.2	1.20
Pell Grant recipient			
Yes	10,780	80.6	1.23
No	39,984	83.2	1.18
Stafford Loan recipient			
Yes	17,940	83.5	1.18
No	32,824	82.3	1.20
Citizenship			
U.S. citizen or resident	48,892	83.1	1.19
Visa	1,872	70.6	1.38
Fall enrollment			
Not enrolled	8,253	80.7	1.23
Enrolled at NPSAS institution	41,380	83.1	1.19
Enrolled at other institution	1,131	87.0	1.14
Number of phone numbers			
0–4	49,863	82.8	1.19
5	666	77.1	1.28
More than 5	235	71.3	1.37

Table 6-7.—Average weight adjustment factors from logistic model used to adjust CATI weights for student location nonresponse —Continued

Logistic model predictor variables	Number of located respondents	Weighted response rate	Average weight adjustment factor (WT14)
Number of schools attended			
1	45,918	82.0	1.21
2	4,535	92.7	1.07
3 or 4	311	98.1	1.02
Date of birth preloaded in CATI			
Yes	46,963	82.4	1.20
No	3,801	86.8	1.15
Parent information preloaded in CATI			
Yes	46,865	82.6	1.19
No	3,899	84.3	1.18
CHAID segments			
1 = Non-Hispanic, no institutional aid, attended 2 schools	3,376	93.2	1.06
2 = Other	47,388	82.2	1.20

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Postsecondary Student Aid Study, 1999–2000 (NPSAS:2000).

The resulting weight adjustment factors are

- minimum: 1.00
- median: 1.18
- maximum: 1.84.

(14) Adjustment for CATI Refusals (WT15)

The second stage of student CATI nonresponse adjustment was an adjustment for refusal during CATI, given that the student was located. This additional type of nonresponse adjustment was made to further compensate for the potential CATI nonresponse bias. The same logistic regression procedure was used as in the adjustment for study nonresponse and not locating students (WT12 and WT14). Candidate predictor variables were the same as those used in the location nonresponse adjustment, with the addition of student marital status and dependency status (2 levels). These additional variables were missing for 5 percent or fewer of all located study respondents.

As in the other two nonresponse adjustments, 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 subjected to variable screening in the logistic procedure. Variables significant at the 15 percent significance level were retained, with the exception of institution type, student type, Pell Grant status, and Stafford Loan status, which were retained regardless of the significance level.

Table 6-8 presents the final predictor variables used in the logistic model to adjust the CATI weights and the average weight adjustment factor resulting from these variables. As in the previous nonresponse adjustments, the weighting adjustment factor for student *j* was the reciprocal of the predicted response probability, or

$$WT15 = 1/\hat{p}_{rj} \ .$$

The resulting weight adjustment factors are

- minimum: 1.00
- median: 1.08
- maximum: 1.37.

(16) Adjustment for Other CATI Nonresponse (WT16)

The third, and final, stage of adjustment for student CATI nonresponse was adjustment for a student not responding to CATI, given that the student was located and did not refuse. This additional type of CATI nonresponse adjustment was made to further compensate for the potential CATI nonresponse bias. The same logistic regression procedure was used as in the adjustment for study nonresponse, not locating students, and CATI refusals (WT12, WT14, and WT15). Candidate predictor variables were the same as those used in the CATI refusal nonresponse adjustment, using three-level dependency status rather than two-level dependency status. This new variable was missing for fewer than 5 percent of all located and nonrefusal study respondents.

As in the other three nonresponse adjustments, 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 subjected to variable screening in the logistic procedure. Variables significant at the 15 percent significance level were retained, with the exception of institution type, student type, Pell Grant status, and Stafford Loan status, which were retained regardless of the significance level.

Table 6-9 presents the final predictor variables used in the logistic model to adjust the CATI weights and the average weight adjustment factor resulting from these variables. As in the previous nonresponse adjustments, the weighting adjustment factor for student *j* was the reciprocal of the predicted response probability, or

$$WT16 = 1/\hat{p}_{rj} \ .$$

Table 6-8.—Average weight adjustment factors from logistic model used to adjust CATI weights for student refusal nonresponse

Logistic model predictor variables	Number of nonrefusal respondents	Weighted response rate	Average weight adjustment factor (WT15)
Total	46,340	89.6	1.10
Institutional sector			
Public less-than-2-year	780	89.7	1.11
Public 2-year	6,240	87.5	1.13
Public 4-year non-doctorate-granting	6,920	91.1	1.09
Public 4-year doctorate-granting	15,180	90.9	1.09
Private not-for-profit 2-year or less	1,040	92.0	1.08
Private not-for-profit 4-year, non-doctorate-granting	5,590	91.4	1.09
Private not-for-profit 4-year doctorate-granting	6,460	90.6	1.10
Private for-profit less-than-2-year	2,500	93.0	1.08
Private for-profit 2-year	800	91.8	1.09
Private for-profit 4-year	810	90.3	1.11
Region			
Southeast	10,320	91.6	1.08
Rocky Mountains	1,910	90.6	1.09
AK, HI, PR	1,120	96.9	1.03
Other	32,990	88.8	1.10
Student type			
Confirmed baccalaureate	10,830	92.3	1.08
Other undergraduate	26,230	89.1	1.10
Graduate	8,320	91.2	1.09
First-professional	950	91.4	1.09
Age group			
Less than 30	33,370	90.2	1.09
30 or older	12,960	88.3	1.11
Sex			
Male	19,090	89.0	1.10
Female	27,250	90.1	1.09
Federal aid recipient			
Yes	21,110	93.2	1.07
No	25,230	87.4	1.12
Pell Grant recipient			
Yes	10,170	94.5	1.05
No	36,170	88.4	1.11
Stafford Loan recipient			
Yes	16,710	92.9	1.07
No	29,630	88.4	1.11
Citizenship			
U.S. citizen	42,600	89.3	1.10
Resident	1,980	94.3	1.05
Visa	1,760	93.5	1.06
Hispanic			
Yes	4,840	92.5	1.06
No	41,490	89.3	1.10

Table 6-8.—Average weight adjustment factors from logistic model used to adjust CATI weights for student refusal nonresponse—Continued

Logistic model predictor variables	Number of nonrefusal respondents	Weighted response rate	Average weight adjustment factor (WT15)
Enrollment[1]			
Less than or equal to 3,267	11,140	92.3	1.08
Greater than 3,267	35,200	89.1	1.10
Number of schools attended			
1	41,600	89.2	1.10
2	4,430	97.0	1.03
3 or 4	310	100.0	1.00
CPS match			
Yes	24,370	92.7	1.07
No	21,970	87.0	1.12
Date of birth preloaded in CATI			
Yes	42,720	89.2	1.10
No	3,620	95.1	1.05
Marital status			
Single	33,940	89.5	1.10
Married	11,740	90.0	1.09
Separated	660	90.0	1.09
CHAID segments[2]			
1 = No aid, attended 1 school, attended full time in fall	7,230	88.7	1.12
2 = No aid, attended 1 school, attended half time in fall	2,970	86.8	1.14
3 = No aid, attended 1 school, attended less than half time or not at all in fall	6,940	83.2	1.19
4 = No aid, attended more than 1 school	1,950	100.0	1.00
5 = Received aid, New England, enrollment <=11,096	990	90.4	1.10
6 = Received aid, New England, 11,096 < enrollment < 24,120	280	87.4	1.14
7 = Received aid, Plains, Southeast, Southwest, Rocky Mountains, Far West, attended less than full time in fall	2,050	91.3	1.09
8 = Received aid, Plains, Southeast, Southwest, Rocky Mountains, Far West, did not attend in fall	1,970	92.6	1.07
9 = Received aid, AK, HI, PR, 15-23 years old	510	99.7	1.00
10 = Other	21,450	93.2	1.07

[1]Enrollment categories were defined by quartiles and then collapsed in the model.

[2]Enrollment categories were defined by quartiles and then collapsed in the Chi-squared automatic interaction detection (CHAID) analysis.

NOTE: To protect confidentiality, some numbers have been rounded.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Postsecondary Student Aid Study, 1999–2000 (NPSAS:2000).

Table 6-9.—Average weight adjustment factors from logistic model used to adjust CATI weights for student other nonresponse

Logistic model predictor variables	Number of respondents	Weighted response rate	Average weight adjustment factor (WT16)
Total	44,490	95.5	1.04
Institutional sector			
Public less-than-2-year	740	93.4	1.06
Public 2-year	5,950	94.7	1.05
Public 4-year non-doctorate-granting	6,730	96.9	1.03
Public 4-year doctorate-granting	14,640	96.2	1.04
Private not-for-profit 2-year or less	980	94.2	1.06
Private not-for-profit 4-year, non-doctorate-granting	5,410	96.4	1.03
Private not-for-profit 4-year doctorate-granting	6,150	95.1	1.05
Private for-profit less-than-2-year	2,350	94.7	1.05
Private for-profit 2-year	780	97.9	1.02
Private for-profit 4-year	760	94.4	1.06
Region			
New England	2,540	95.2	1.05
Southwest	4,650	94.4	1.05
Other	37,310	95.7	1.04
Student type			
Confirmed baccalaureate	10,400	96.2	1.04
Other undergraduate	25,130	95.3	1.04
Graduate	8,040	96.6	1.03
First-professional	920	96.9	1.03
Gender			
Male	18,240	94.9	1.05
Female	26,250	96.1	1.04
Institutional aid recipient			
Yes	10,450	96.4	1.04
No	34,040	95.4	1.04
Pell Grant recipient			
Yes	9,730	95.8	1.04
No	34,760	95.5	1.04
Stafford Loan recipient			
Yes	16,180	97.0	1.03
No	28,310	95.0	1.05
Fall attendance			
Full time	27,730	96.4	1.03
Half time	5,710	95.5	1.04
Less than half time	4,040	94.0	1.05
None	7,020	94.2	1.05
Enrollment			
Less than or equal to 11,096	22,260	96.6	1.03
Between 11,096 and 24,120 (not inclusive)	11,060	95.0	1.04
Greater than or equal to 24,120	11,170	94.4	1.05
Number of schools attended			
1	39,790	95.3	1.04
2	4,390	99.2	1.01
3 or 4	310	100.0	1.00

Table 6-9.—Average weight adjustment factors from logistic model used to adjust CATI weights for student other nonresponse—Continued

Logistic model predictor variables	Number of respondents	Weighted response rate	Average weight adjustment factor (WT16)
Number of phone numbers			
0	150	71.4	1.39
1 or 2	34,890	95.8	1.04
3	6,700	95.1	1.04
4	2,010	95.3	1.04
5	560	94.5	1.05
More than 5	190	90.4	1.09
Marital status			
Single	32,460	95.3	1.04
Married or separated	12,030	96.3	1.03
Dependency			
Dependent	24,970	95.9	1.04
Independent	19,520	95.1	1.04
Date of birth preloaded in CATI			
Yes	40,990	95.4	1.04
No	3,500	97.6	1.02
Parent information preloaded in CATI			
Yes	3,440	96.9	1.03
No	41,060	95.5	1.04
CHAID segments			
1 = U.S. citizen, attended 1 school, Hispanic	3,500	93.1	1.07
2 = U.S. citizen, attended more than 1 school, no federal aid	2,240	100.0	1.00
3 = Resident or visa, public 2-year or less, attended 1 school	380	84.0	1.19
4 = Resident or visa, public 4-year attended 1 school	1,450	92.1	1.08
5 = Resident or visa, Private not-for-profit 2-year or less, full-time in fall	50	71.0	1.38
6 = Resident or visa, Private not-for-profit 4-year, single	550	85.6	1.16
7 = Resident or visa, Private not-for-profit 4-year, married or separated	260	92.1	1.08
8 = Resident or visa, Private for-profit less-than-2-year, enrolled at NPSAS institution or not at all in fall	110	89.7	1.11
9 = Private for-profit 2-year or more, resident	80	94.8	1.05
10 = Private for-profit 2-year or more, visa	60	82.4	1.22
11 = Other	35,810	96.4	1.03

NOTE: To protect confidentiality, some numbers have been rounded.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Postsecondary Student Aid Study, 1999–2000 (NPSAS:2000).

The resulting weight adjustment factors are

- minimum: 1.00
- median: 1.03
- maximum: 1.49.

(17) Poststratification Adjustment for CATI Respondents (WT17)

To ensure population coverage, the CATI weights were adjusted to control totals with the same generalized raking procedure used to adjust the study weights. The control totals established for the study weights also were used for the CATI weights. To help reduce nonresponse bias further, we additionally formed control totals for annual enrollment by student type as well as control totals by

- sex,
- age group (<24, 24–29, and 30+),
- federal aid applicant,
- federal aid receipt,
- state aid receipt,
- institution aid receipt, and
- fall attendance status.

The annual enrollment control totals by student type were formed using the study weights so that estimates of the annual enrollment using the study or CATI weights would be the same. The other (new) control totals were also computed using the study weights because these variables were known for most CATI respondents and nonrespondents. As in the previous poststratification adjustment (WT13).

The exponential adjustment satisfies the following constraints:

$$\sum_j W_j \lambda_j \boldsymbol{x}_j^T = \eta_O^T \ ,$$

where

W_j = the cumulative weight (WT1•WT2•....•WT12),

λ_j = $\exp(\alpha + \mathbf{x}_j B)$,

α = model intercept

β = vector of parameters that specify the nature of the relationship between λ_j and $\mathbf{x}_j$

$\mathbf{x}_j$ = the vector of regressors associated with the domains to be controlled, and

η_o = the set of control totals.

WT17 = λ_j .

Table 6-5 presented the student enrollment control totals by student type and institution type and the average weight adjustment factors by these variables. Similarly, Table 6-6 presented the variables associated with the Pell Grant and Stafford Loan control totals and the average weight adjustment factors. Table 6-10 displays seven variables by institution type associated with the student enrollment control totals and the average weight adjustment factors for these variables. The weight adjustment factors from the exponential adjustment are summarized below, and met the constraints

- minimum: 0.55
- median: 0.99
- maximum: 1.36.

After this last weight adjustment was performed, the final CATI weights (CATIWT) were computed as the product of the unrounded study weights and the remaining four weight components and then rounded to the nearest integer.

The two statistical analysis weights on the analysis files are the study weight (STUDYWT) and the CATI weight (CATIWT). The study weight is the product of weight components WT1-WT13 and should be used when no data items in the analysis are based entirely on CATI data or require CATI data to be reliable. The CATI weight is the product of all weight components (WT1-WT17) and should be used when at least one data item in the analysis is based entirely on CATI data or requires CATI data to be reliable.

The distributions of the study weights and the CATI weights are summarized in Tables 6-11 and 6-12, respectively. These tables also summarize the variance inflation due to unequal weighting, i.e., the unequal weighting effect. It can be seen that the unequal weighting effects are slightly higher for the CATI weights than for the study weights (2.00 versus 1.83). The lowest design effects are for students from public 2-year institutions, and the highest design effects are for students from private for-profit less-than-2-year institutions.

Table 6-10.—Average weight adjustment factors from exponential model for poststratifying to study weight control totals

Exponential model variables	Control total	Average weight adjustment factor (WT17)
Fall attendance by institutional sector		
Full-time		
Public less-than-2-year	50,618	0.96
Public 2-year	2,376,264	0.95
Public 4-year non-doctorate-granting	1,345,611	0.98
Public 4-year doctorate-granting	3,069,092	0.98
Private not-for-profit 2-year or less	87,384	1.00
Private not-for-profit 4-year, non-doctorate-granting	1,110,598	0.98
Private not-for-profit 4-year doctorate-granting	1,162,583	0.98
Private for-profit less-than-2-year	143,473	1.02
Private for-profit 2-year	191,160	1.03
Private for-profit 4-year	146,104	1.08
Half-time		
Public less-than-2-year	17,738	1.09
Public 2-year	1,648,417	1.03
Public 4-year non-doctorate-granting	370,970	1.05
Public 4-year doctorate-granting	585,981	1.13
Private not-for-profit 2-year or less	13,695	0.97
Private not-for-profit 4-year, non-doctorate-granting	230,795	1.04
Private not-for-profit 4-year doctorate-granting	232,861	1.09
Private for-profit less-than-2-year	22,251	1.04
Private for-profit 2-year	33,212	1.00
Private for-profit 4-year	36,175	1.06
Less than half time		
Public less-than-2-year	16,182	0.98
Public 2-year	1,540,201	1.06
Public 4-year non-doctorate-granting	242,822	1.03
Public 4-year doctorate-granting	402,605	1.00
Private not-for-profit 2-year or less or 4-year, non-doctorate-granting	155,002	1.05
Private not-for-profit 4-year doctorate-granting	165,969	1.05
Private for-profit less-than-2-year	5,251	0.76
Private for-profit 2-year and 4-year	21,883	0.98
None		
Public less-than-2-year	27,992	1.02
Public 2-year	2,003,574	1.01
Public 4-year non-doctorate-granting	348,018	1.03
Public 4-year doctorate-granting	599,767	1.00
Private not-for-profit 2-year or less	29,965	1.02
Private not-for-profit 4-year, non-doctorate-granting	246,762	1.03
Private not-for-profit 4-year doctorate-granting	219,251	0.98
Private for-profit less-than-2-year	158,775	1.01
Private for-profit 2-year	86,992	0.98
Private for-profit 4-year	50,002	0.87

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Postsecondary Student Aid Study, 1999–2000 (NPSAS:2000).

Table 6-10.—Average weight adjustment factors from exponential model for poststratifying to study weight control totals —Continued

Exponential model variables	Control total	Average weight adjustment factor (WT17)
Age group by institutional sector		
Less than 24 years old		
Public less-than-2-year	35,286	1.01
Public 2-year	3,481,994	0.98
Public 4-year non-doctorate-granting	1,284,235	1.00
Public 4-year doctorate-granting	2,688,476	0.99
Private not-for-profit 2-year or less	90,507	1.00
Private not-for-profit 4-year, non-doctorate-granting	941,304	0.98
Private not-for-profit 4-year doctorate-granting	848,262	1.01
Private for-profit less-than-2-year	140,826	0.99
Private for-profit 2-year	153,360	0.98
Private for-profit 4-year	76,616	1.11
24-29 years old		
Public less-than-2-year	22,563	1.01
Public 2-year	1,391,321	1.03
Public 4-year non-doctorate-granting	446,216	1.01
Public 4-year doctorate-granting	1,007,081	1.03
Private not-for-profit 2-year or less	19,311	0.90
Private not-for-profit 4-year, non-doctorate-granting	272,413	1.04
Private not-for-profit 4-year doctorate-granting	441,175	1.00
Private for-profit less-than-2-year	91,421	1.02
Private for-profit 2-year	91,794	1.11
Private for-profit 4-year	68,627	1.03
30 years old or older		
Public less-than-2-year	54,683	0.97
Public 2-year	2,695,140	1.00
Public 4-year non-doctorate-granting	576,970	0.98
Public 4-year doctorate-granting	961,888	0.99
Private not-for-profit 2-year or less	25,922	1.11
Private not-for-profit 4-year, non-doctorate-granting	524,744	1.00
Private not-for-profit 4-year doctorate-granting	491,226	0.99
Private for-profit less-than-2-year	97,502	1.03
Private for-profit 2-year	73,333	0.97
Private for-profit 4-year	101,798	0.97

Table 6-10.—Average weight adjustment factors from exponential model for poststratifying to study weight control totals —Continued

Exponential model variables	Control total	Average weight adjustment factor (WT17)
Gender by institutional sector		
Males		
Public less-than-2-year	55,370	1.01
Public 2-year	3,274,820	1.01
Public 4-year non-doctorate-granting	942,920	0.98
Public 4-year doctorate-granting	2,140,714	1.00
Private not-for-profit 2-year or less	58,247	1.00
Private not-for-profit 4-year, non-doctorate-granting	708,495	0.99
Private not-for-profit 4-year doctorate-granting	821,063	0.98
Private for-profit less-than-2-year	121,612	0.98
Private for-profit 2-year	112,219	1.00
Private for-profit 4-year	127,325	1.00
Females		
Public less-than-2-year	57,162	0.98
Public 2-year	4,293,635	0.99
Public 4-year non-doctorate-granting	1,364,501	1.01
Public 4-year doctorate-granting	2,516,732	1.00
Private not-for-profit 2-year or less	77,494	1.01
Private not-for-profit 4-year, non-doctorate-granting	1,029,968	1.00
Private not-for-profit 4-year doctorate-granting	959,600	1.01
Private for-profit less-than-2-year	208,138	1.02
Private for-profit 2-year	206,268	1.01
Private for-profit 4-year	119,717	1.04
CPS match by institutional sector		
Matched CPS		
Public less-than-2-year	41,733	0.95
Public 2-year	2,537,146	1.00
Public 4-year non-doctorate-granting	1,220,921	0.99
Public 4-year doctorate-granting	2,252,757	0.99
Private not-for-profit 2-year or less	93,083	1.00
Private not-for-profit 4-year, non-doctorate-granting	1,042,320	0.99
Private not-for-profit 4-year doctorate-granting	938,019	1.01
Private for-profit less-than-2-year	276,380	1.01
Private for-profit 2-year	283,412	1.01
Private for-profit 4-year	163,223	0.98
Did not match CPS		
Public less-than-2-year	70,800	1.03
Public 2-year	5,031,309	1.00
Public 4-year non-doctorate-granting	1,086,501	1.01
Public 4-year doctorate-granting	2,404,689	1.01
Private not-for-profit 2-year or less	42,659	1.01
Private not-for-profit 4-year, non-doctorate-granting	696,143	1.01
Private not-for-profit 4-year doctorate-granting	842,645	0.98
Private for-profit less-than-2-year	53,371	0.98
Private for-profit 2-year	35,076	1.01
Private for-profit 4-year	83,820	1.08

Table 6-10.—Average weight adjustment factors from exponential model for poststratifying to study weight control totals —Continued

Exponential model variables	Control total	Average weight adjustment factor (WT17)
Federal aid recipient by institutional sector		
Received federal financial aid		
Public less-than-2-year	29,806	0.95
Public 2-year	1,725,729	0.99
Public 4-year non-doctorate-granting	1,013,460	1.00
Public 4-year doctorate-granting	1,926,288	1.00
Private not-for-profit 2-year or less	78,783	0.99
Private not-for-profit 4-year, non-doctorate-granting	928,595	0.99
Private not-for-profit 4-year doctorate-granting	843,977	1.02
Private for-profit less-than-2-year	265,349	1.03
Private for-profit 2-year	276,166	1.00
Private for-profit 4-year	162,384	0.98
Did not receive federal financial aid		
Public less-than-2-year	82,727	1.01
Public 2-year	5,842,726	1.00
Public 4-year non-doctorate-granting	1,293,962	1.00
Public 4-year doctorate-granting	2,731,158	1.00
Private not-for-profit 2-year or less	56,959	1.03
Private not-for-profit 4-year, non-doctorate-granting	809,868	1.00
Private not-for-profit 4-year doctorate-granting	936,687	0.98
Private for-profit less-than-2-year	64,402	0.95
Private for-profit 2-year	42,322	1.04
Private for-profit 4-year	84,659	1.08
State aid recipient by institutional sector		
Received state financial aid		
Public less-than-2-year	7,222	0.97
Public 2-year	993,524	0.98
Public 4-year non-doctorate-granting	410,207	0.99
Public 4-year doctorate-granting	626,012	1.02
Private not-for-profit 2-year or less	27,114	0.95
Private not-for-profit 4-year, non-doctorate-granting	363,646	0.96
Private not-for-profit 4-year doctorate-granting	199,701	0.98
Private for-profit less-than-2-year	12,942	0.98
Private for-profit 2-year	53,653	0.91
Private for-profit 4-year	11,875	0.76
Did not receive state financial aid		
Public less-than-2-year	105,311	0.99
Public 2-year	6,574,931	1.00
Public 4-year non-doctorate-granting	1,897,215	1.00
Public 4-year doctorate-granting	4,031,434	1.00
Private not-for-profit 2-year or less	108,628	1.02
Private not-for-profit 4-year, non-doctorate-granting	1,374,817	1.01
Private not-for-profit 4-year doctorate-granting	1,580,963	1.00
Private for-profit less-than-2-year	316,809	1.01
Private for-profit 2-year	264,835	1.03
Private for-profit 4-year	235,168	1.04

Table 6-10.—Average weight adjustment factors from exponential model for poststratifying to study weight control totals —Continued

Exponential model variables	Control total	Average weight adjustment factor (WT17)
Institutional aid recipient by institutional sector		
Received institutional financial aid		
Public 2-year-or-less	306,645	1.01
Public 4-year non-doctorate-granting	283,801	1.03
Public 4-year doctorate-granting	983,407	1.00
Private not-for-profit 2-year or less	44,809	1.01
Private not-for-profit 4-year, non-doctorate-granting	678,407	0.97
Private not-for-profit 4-year doctorate-granting	715,038	1.01
Private for-profit less-than-2-year	19,664	0.98
Private for-profit 2-year	19,846	1.07
Private for-profit 4-year	23,903	1.10
Did not receive institutional financial aid		
Public 2-year-or-less	9,290,254	1.00
Public 4-year non-doctorate-granting	2,023,621	0.99
Public 4-year doctorate-granting	3,674,039	1.00
Private not-for-profit 2-year or less	90,933	1.00
Private not-for-profit 4-year, non-doctorate-granting	1,060,056	1.02
Private not-for-profit 4-year doctorate-granting	1,065,626	0.99
Private for-profit less-than-2-year	310,087	1.01
Private for-profit 2-year	298,642	1.00
Private for-profit 4-year	223,140	1.01

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Postsecondary Student Aid Study, 1999–2000 (NPSAS:2000).

Table 6-11.—Study weight distribution and unequal weighting effects for study respondents

Analysis Domain	Minimum	First Quartile	Median	Third Quartile	Maximum	Mean	Unequal weighting effect[1]
Total	2.53	93.18	255.23	395.83	2862.53	310.78	1.83
Student type							
Undergraduate	2.53	89.49	292.41	413.49	2862.53	331.21	1.83
Graduate	10.34	97.67	225.94	289.92	2592.78	219.30	1.54
First-professional	25.91	204.17	278.96	339.23	1071.49	271.54	1.18
Institutional sector							
Public less-than-2-year	2.53	24.92	91.80	181.87	260.08	105.86	1.59
Public 2-year	50.39	754.92	884.41	998.65	2100.35	847.34	1.07
Public 4-year non-doctorate-granting	10.34	87.45	268.19	366.98	2862.53	257.81	1.58
Public 4-year doctorate-granting	10.22	100.11	213.72	379.26	1829.84	236.06	1.50
Private not-for-profit 2-year or less	6.29	57.31	86.03	127.03	170.17	89.84	1.24
Private not-for-profit 4-year, non-doctorate-granting	6.51	96.59	255.36	371.57	988.83	241.79	1.39
Private not-for-profit 4-year doctorate-granting	13.89	71.69	213.49	315.56	1549.54	211.68	1.53
Private for-profit less-than-2-year	3.27	53.35	67.05	96.30	876.59	90.79	2.26
Private for-profit 2-year	34.60	205.64	254.12	325.31	815.41	271.28	1.19
Private for-profit 4-year	13.87	118.03	195.84	265.25	1520.44	210.61	1.54

[1]Unequal weighting effect calculated as $n \Sigma(Wt)^2 / (\Sigma Wt)^2$.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Postsecondary Student Aid Study, 1999–2000 (NPSAS:2000).

Table 6-12.—CATI weight distribution and unequal weighting effects for CATI respondents

Analysis Domain	Minimum	First Quartile	Median	Third Quartile	Maximum	Mean	Unequal weighting effect[1]
Total	2.53	93.18	255.23	395.83	2862.53	310.78	2.00
Student type							
Undergraduate	2.95	116.99	378.39	579.72	3696.58	465.41	2.00
Graduate	10.23	123.61	285.08	389.45	2908.80	290.19	1.60
First-professional	25.99	248.99	356.54	440.64	1754.40	353.96	1.22
Institutional sector							
Public less-than-2-year	2.95	31.30	106.35	265.10	615.24	151.66	1.71
Public 2-year	52.92	1012.93	1358.66	1578.59	3387.62	1271.15	1.13
Public 4-year non-doctorate-granting	10.23	112.61	338.65	504.48	3696.58	343.11	1.65
Public 4-year doctorate-granting	9.25	125.97	225.24	527.66	2173.21	318.07	1.58
Private not-for-profit 2-year or less	9.70	80.91	137.01	192.45	393.98	138.65	1.29
Private not-for-profit 4-year, non-doctorate-granting	8.07	119.87	317.84	501.01	1620.23	321.52	1.49
Private not-for-profit 4-year doctorate-granting	13.21	94.20	265.57	440.63	2740.76	289.59	1.58
Private for-profit less-than-2-year	3.18	77.60	106.61	146.08	1618.00	140.62	2.38
Private for-profit 2-year	81.55	307.69	386.56	482.34	1166.44	406.75	1.15
Private for-profit 4-year	12.26	176.68	262.79	431.88	2229.27	323.35	1.55

[1]Unequal weighting effect calculated as $n \Sigma(Wt)^2 / (\Sigma Wt)^2$.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Postsecondary Student Aid Study, 1999–2000 (NPSAS:2000).

6.2 Baccalaureate (B&B) Weights

Because baccalaureate status was known only for CATI respondents, the CATI weights (WT17) are the appropriate analysis weights for students known to be baccalaureate recipients.

In addition, base weights were needed for all students who belonged to the base-year cohort of the Baccalaureate and Beyond (B&B)-longitudinal follow-up study. The sampling frame for the B&B follow-up included all NPSAS CATI respondents confirmed to be baccalaureate recipients, as well as all study respondents who were sampled as potential baccalaureate recipients but who were CATI nonrespondents. Hence, the NPSAS study weight should be used as the base weight to develop statistical analysis weights for the Baccalaureate and Beyond Longitudinal Study.

6.3 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 common procedures for estimating variances of survey statistics are the Taylor series linearization procedure and the balanced repeated replication (BRR) procedure, which are both available on the NPSAS data files. Section 6.3.1 discusses the analysis strata and replicates created for the Taylor series procedure, and Section 6.3.2 discusses the replicate weights created for the BRR procedure.

Also, to measure the effects that complex sample design features had on the variances of survey estimates, Section 6.3.3 presents design effect estimates for several key statistics within each of several analysis domains.

6.3.1 Taylor Series

The Taylor series variance estimation procedure is a well-known technique 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 appropriate variance formula based on the sample design. Woodruff[6] presented the mathematical formulation of this procedure.

For stratified multistage surveys, the Taylor series procedure requires analysis strata and analysis primary sampling units (PSUs) defined from the sampling strata and PSUs used in the first stage of sampling. For NPSAS:2000, analysis strata and analysis PSUs were defined separately for each domain for which separate analyses were anticipated: all students combined, all undergraduate students, all graduate/first-professional students, and all baccalaureate students.

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

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 796 noncertainty institutions. For the 287 certainty institutions, however, the students represent the first stage of sampling. In order to obtain appropriate degrees of freedom for variance estimation, the students selected from each certainty institution were partitioned into two, three, or four pseudo-PSUs by random assignment of sample students into approximately equal-sized groups. The number of pseudo-PSUs formed was based on the institution's measure of size for first-stage sampling.

The next step was to sort the PSUs and pseudo-PSUs by the 22 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 and pseudo-PSUs as required so each analysis PSU contained at least four CATI respondents. This sample size requirement satisfied the requirements of the NCES DAS and ensured stable variance estimates. Analysis PSUs were then paired to form analysis strata. Certainty institutions that included three or four pseudo-PSUs were made a single analysis stratum. This process resulted in 624 analysis strata for all students, 623 analysis strata for undergraduate students, 361 analysis strata for graduate/first-professional students, and 396 analysis strata for baccalaureates.

The names of the analysis strata and analysis PSU variables are:

- **ANALSTR, ANALPSU:** Analysis strata and analysis PSUs for all students
- **UANALSTR, UANALPSU:** Analysis strata and analysis PSUs for undergraduate students
- **GANALSTR, GANALPSU:** Analysis strata and analysis PSUs for graduate/first-professional students
- **BANALSTR, BANALPSU:** Analysis strata and analysis PSUs for baccalaureate recipients.

6.3.2 Balanced Repeated Replication

The BRR procedure is an alternative variance estimation procedure that computes the variance based on a balanced set of pseudo-replicates. BRR weights were computed because of concern that the variances for medians and other quantiles might not be appropriate when computed using Taylor series or other methods such as the Jackknife procedure. The BRR variance estimation process involved modeling the design as if it were a two-PSU-per-stratum design. Variances were then calculated using a random group type of variance estimation procedure, with a balanced set of replicates as the groups. Balancing was done by creating replicates using an orthogonal matrix and allowed the use of less than the full set of 2^L possible replicates, where L is the number of analysis strata.

To form pseudo replicates for BRR variance estimation, the Taylor Series analysis strata were collapsed. The number of Taylor Series analysis strata.and PSUs were different for all

students combined, graduates/first-professionals, and baccalaureate recipients, so the collapsing was done independently and, hence, with different results. The goal of the collapsing was to get 50 to 120 replicates and not necessarily the same number of replicates for each domain. A common rule is to have at least 50 replicates; the gain in efficiency with more than 120 replicates does not justify the extra effort.[7] The analysis strata defined for the Taylor series were collapsed to form the BRR analysis strata, which included

- 52 BRR strata for all students combined,
- 60 BRR strata for graduate/first-professional students, and
- 64 BRR strata for baccalaureate students.

Then, two BRR pseudo-PSUs were created within each stratum by collapsing the Taylor series analysis PSUs.

Based on the BRR strata and PSU definitions, we created replicate weights associated with the two analysis weights: study weights and CATI weights. For the study weights, this included separate replicate weights for all students and for graduate/first-professional students only; for the CATI weights, this included separate replicate weights for all students, graduate/first-professional students only, and baccalaureates only. Thus, a total of five replicate weight sets were created:

- **BRSWT01–BRSWT52:** Study BRR weights for all students
- **BRSGWT01–BRSGWT60:** Study BRR weights for graduate/first-professional students
- **BRCWT01–BRCWT52:** CATI BRR weights for all students
- **BRCGWT01–BRCGWT60:** CATI BRR weights for graduate/first-professional students
- **BRCBWT01–BRCBWT64:** CATI BRR weights for baccalaureate students.

To create the replicate weights, student-level replicate weights were defined. For each replicate set, student weights of one PSU within each analysis stratum were set to zero and the student weights of the other PSUs were doubled to approximately preserve the population weight total. The number of replicates was set equal to the number of analysis strata to achieve the correct degrees of freedom for variance estimation. Then each set of replicate weights was poststratified to the control totals, similar to the description in Section 6.1, with a couple of exceptions to allow the models to converge. First, there were model convergence problems for some replicates when we attempted to control to total Pell grant recipients and also to Pell grant amounts. Therefore, we could not control the mean value and could only control to Pell amounts. Second, for several of the replicates, we had to collapse some control totals, such as

[7] Babu V. Shah. Personal correspondence, 2001

enrollment by sector, for two sectors because some replicates had small sample sizes for certain poststratification groups.

6.3.3 Design Effects

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. Also, weight adjustments for nonresponse, which are performed to reduce nonresponse bias, increase the variance by increasing the weight variation. Because of these effects, most complex multistage sampling designs, like NPSAS:2000, result in design effects greater than one. 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

$$Deff(\hat{\theta}) = \frac{Var_{design}(\hat{\theta})}{Var_{srs}(\hat{\theta})}.$$

Also, the square root of the design effect is another useful measure, which can also be expressed as the ratio of the standard errors, or

$$Deft(\hat{\theta}) = \frac{SE_{design}(\hat{\theta})}{SE_{srs}(\hat{\theta})}.$$

In Appendix I, design effect estimates are presented to summarize the effects of stratification, multistage sampling, unequal probabilities of selection, and the nonresponse weight adjustments. These design effects were estimated using SUDAAN, which uses the Taylor series variance estimation procedure.[8] If one must perform a quick analysis of NPSAS:2000 data without using one of the software packages for analysis of complex survey data, the design effect tables in this appendix can be used to make approximate adjustments to the standard errors of survey statistics computed using the standard software packages that assume simple random sampling designs. However, one cannot be confident regarding the actual design-based standard errors without performing the analysis using one of the software packages specifically designed for analysis of data from complex sample surveys.

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 under 2.0 is low, 2.0 to 3.0 is moderate, and above 3.0 is high. Moderate and high design effects often occur in complex surveys such as NPSAS, and the design effects in appendix I are consistent with those in past NPSAS studies. Unequal weighting causes large design effects and

[8] B.V Shah, B.G Barnwell, and G.S Bieler. *SUDAAN User's Manual*. Research Triangle Park, NC: Research Triangle Institute, 1995.

is often due to nonresponse adjustments. However, in NPSAS, the unequal weighting is due to the sample design and different sampling rates between institution strata and also different sampling rates between student strata. The median design effects in appendix I are generally lower when based on CATI weights rather than study weights. However, estimates based on CATI weights have smaller sample sizes, so the precision is not necessarily better than for estimates based on study weights with larger sample sizes.

Appendix I presents tables of design effect estimates for important survey estimates among undergraduate students, graduate students, and first-professional students, along with a discussion of statistical analysis considerations and specifications for the generic program code. The tables include design effects based on the study weights and on the CATI weights. Specifically, these tables are:

- Tables I.1–I.19: Design effects for undergraduates based on study weights
- Tables I.20–I.38: Design effects for undergraduates based on CATI weights
- Tables I.39–I.41: Design effects for graduates (excluding first-professionals) based on study weights
- Tables I.42–I.44: Design effects for graduates (excluding first-professionals) based on CATI weights
- Tables I.45–I.47: Design effects for first-professionals based on study weights
- Tables I.48–I.50: Design effects for first-professionals based on CATI weights.