Appendix II, Part I

Sampling and Estimation Procedures

Diana Stukel, Annie Lo, and **Bradford Chaney, Westat**

Overview

A stratified, simple random sample was employed in the CBMS 2010 survey, and strata were based on three variables: curriculum, highest degree level offered, and total institutional enrollment. Data were collected using an online survey with email and telephone followup.

Sampling Approach

For CBMS 2010, the basic design was a stratified simple random sample of institutions. A compromise mix of statistically optimum Neyman allocations based on two key outcome variables was used to determine targeted sample sizes for the 28 sampling strata.

Target Population and Sampling Frames

The Integrated Postsecondary Education Data System (IPEDS), a database maintained by the National Center for Education Statistics within the U.S. Department of Education, was used as a basis for building a frame for this survey. For the academic year 2008-2009, there were approximately 6,800 colleges and universities across the country, according to IPEDS. Of these, 2,593 had mathematics or statistics departments (or both). AMS conducts annual surveys of four-year institutions, and thus has reasonably current information for four-year institutions; this information was used as a basis for updating the IPEDS frame. AMS and Westat also contacted two-year institutions to obtain updated information for them. Two primary considerations with regard to two-year institutions were determining how the institutions organized mathematics within departments or divisions (e.g., there may be a combined division of science and mathematics), and whether the systems were centralized (so that one institution had all required information) or decentralized (so that each campus must be surveyed separately, and the sampling unit would be the campus rather than the institution). In the case of decentralization, IPEDS generally has information for the overall institution rather than for each individual campus, so the IPEDS-based frame was modified to include the individual campuses.

In 2010, the Mathematical Association of America (MAA) also conducted a survey of faculty and students of two-year and four-year colleges and universities where calculus is taught. Although the two surveys (CBMS and MAA) were administered separately and at different times, and although the surveys targeted somewhat different respondents (department heads for the CBMS survey, and faculty and students for the MAA survey), a joint sampling plan to serve both surveys simultaneously was developed. Thus, the overarching aim was to optimize the allocation for both surveys while minimizing overlap between them wherever possible.

The target population of the CBMS 2010 survey consisted of undergraduate mathematics and statistics programs at two-year and four-year colleges and universities in the United States. Thus the frame for the CBMS 2010 survey was divided into three parts: (A) 1,393 institutions having four-year math programs, (B) 79 institutions having four-year statistics programs, and (C) 1,121 institutions having two-year math programs, for a total of 2,593 institutions having programs eligible for participation in the survey. In most cases, these programs were within established academic departments, but at times they were part of more comprehensive departments (i.e., covering more topics than mathematics and/or statistics) or divisions. Note that parts A and B did not necessarily consist of mutually exclusive institutions since some institutions had both four-year math programs and four-year statistics programs. However, this was not problematic since the math and statistics programs within these institutions were the targets of interest, and the departments were sampled independently.

Sampling Strata

The three parts of the frame were each stratified using the same two variables that were used in the previous two rounds of the CBMS survey, that is, "Highest Degree Granted by the Institution" (PhD, MA, or BA) and "Institutional Full Time Equivalent (FTE) Undergraduate Fall Enrollment." This initially resulted in the same 24 strata that were used for CBMS 2005. For this round of the survey, however, the values for the stratification variables were taken from IPEDS 2008. A further refinement to the stratification was made by calculating the standard errors for each of the strata in parts A and C using data for two key outcome variables from CBMS 2005 – "Total Student Enrollment in Math Department's Undergraduate Courses, Previous Fall (2004)" and "Number of Full-Time Tenured or Mathematics Faculty in Fall 2005." The standard errors were used as a gauge to assess how homogeneous the strata were. Based on this calculation, four additional strata were created (for a total of 28 strata) by splitting four of the original strata. The four original strata had the highest standard errors for both variables considered amongst the 24 strata, and thus it was felt that splitting them would create more homogeneous strata. The final stratification can be seen in the first four columns of Table 1. The four-year mathematics programs were divided into fifteen strata, the four-year statistics programs were divided into five strata, and the two-year programs were divided into eight strata. Note that the four pairs of strata labeled 1 and 2; 3 and 4; 8 and 9; and 25 and 26 were originally combined in CBMS 2005. The stratification for part B of the frame remained unchanged from CBMS 2005.

Allocation Process

For the CBMS 2010 survey, a stratified simple random sample of 600 institutions was drawn from parts A, B, and C. For CBMS 2010, since there were only 79 institutions within part B of the frame (four-year Statistics), and since each of the five strata within part B had fewer than 25 institutions, a decision was made to sample all 79 institutions, forcing strata 16-20 to be certainty strata. Thus, the remaining 521 sampled institutions for CBMS 2010 were sampled from parts A and C of the frame.

In order to allocate the sample optimally to each of the 23 strata, Neyman allocation was used. This form of allocation distributes sample to the strata proportionately to the overall number of institutions on the frame belonging to each stratum, while adjusting the allocation to give more sample to those strata with greater variability (larger standard errors) with respect to key variables. The standard errors for the same two key variables that were used in the stratification process ("Total Student Enrollment in Math Department's Undergraduate Courses, Previous Fall (2004)" and "Number of Full-Time Tenured (four-year) or Mathematics (or two-year) Faculty in Fall 2005") were used for this purpose. The same basic methodology that was used in CBMS 2005 was followed here. That is, separate Neyman allocations were calculated based on standard errors of the two key variables and then a composite combination of the two allocations was calculated by giving the Neyman allocation based on the first variable (enrollment) a relative weight of 0.75 and the Neyman allocation based on the second variable (faculty) a relative weight of 0.25. Giving a higher relative weight to enrollment seemed to be a reasonable approach given that this variable was deemed to be more salient to the study and had greater variability in the stratum-level standard errors than the faculty variable.

Given that one of the interests of this study was to obtain estimates at the level of the three program types (A, B, and C), it was necessary to ensure estimates of roughly equal precision (i.e., having the same variances) at these three levels. However, given that a "census" of institutions from the four-year statistics part of the frame was taken, there was no sampling error associated with estimates from part B. Therefore, the sampling strategy was limited to ensuring equivalent precisions for estimates coming from the two other levels (parts A and C), and the Neyman allocation was constrained to ensure this. For the purposes of this exercise, the precision under the composite Neyman allocation was approximated by using variances of the same two key variables as were used above from CBMS 2005, at the aggregate part A and C levels. Variances for each of the two key variables under the composite allocation were considered separately. Given the identical sample sizes for the two surveys, it seemed reasonable that the allocations for the two surveys should be identical as well. Thus, the constrained allocation was achieved by initially allocating roughly half of the 521 institutions to each of parts A and C, performing Neyman allocations to the fifteen strata in part A and eight strata in part C, and computing the two variances for parts A and C. The above process was iteratively reworked until approximate equivalence between the variances for parts A and C was achieved. That is, the fifty percent allocation of the 521 sampled institutions to each of parts A and B was re-adjusted to be disproportionate, the Neyman allocation to the fifteen plus eight strata was recomputed, and the variances of parts A and B were also re-computed until the variances roughly matched.

Because another important aspect of the design was the need to minimize the overlap between the CBMS 2010 survey and the MAA 2010 survey while minimizing the overall aggregate level variances, four scenarios were considered under the constrained Neyman allocation procedure described in the above paragraph. The first scenario forced strata 6 and 28 to be certainty strata, since their universe sizes were so small (six institutions each) and since a Neyman allocation would force such a capping regardless. This scenario was compared with three other scenarios where strata 5, 9, and 27 were successively also added as certainty strata to the two initial certainty strata, since their individual stratum level variances were greatest amongst all strata and since their universe sizes were smallest. Of the four scenarios considered above, the one having the minimum overall variances (for both enrollment and number of faculty) while maintaining the smallest amount of overlap was the one for which strata 5, 6, and 28 were designated as certainty strata. Note that this design assumed that both the CBMS 2010 and the MAA 2010 surveys would be administered to each institution within each certainty stratum (and therefore each institution within such strata would be visited twice). Note that this design also generated additional overlap from strata where greater than half the institutions comprising the universe were sampled – stated otherwise, where the sampling rate, given by the number of sampled institutions divided by the number of institutions in the universe, was greater than 0.5.

The resultant "optimal" sample for both CBMS 2010 and MAA 2010 consisted of 314 institutions sampled from part A (including the two certainty strata, strata 5 and 6, of size nineteen and six, respectively), and 207 institutions (including the one certainty stratum, stratum 29, of size six), for a total of 521 institutions. See Table 1 below for details of the final allocation given in the columns labeled "Universe" (or number of institutions on the frame), "Final Sample Allocation", and "Sampling Rate". Note that, apart from the three certainty strata, where there was 100% overlap between the two samples, there were also five strata where the sampling rate was greater than 0.5, indicating partial overlap between the two samples. The overall number of overlapped institutions between the two samples was 75; that is, 31 from the three certainty strata and 44 from the five strata where the sampling rate was greater than 0.5. The overlap of 75 institutions represented roughly 15% of the 521 sampled institutions; it was not possible to reduce this any further given the modest universe sizes within each stratum.

For each of CBMS 2010 and MAA 2010, 314 institutions were drawn from part A (drawing separately for each of the fifteen strata in accordance with the specific allocation in Table 1), and 207 institutions from part C (drawing separately for each of the eight strata in accordance with the specific allocation in Table 1). Additionally, for CBMS 2010, the 79 certainty institutions from part B (with sampling rates of 1.0) were added to the 521 institutions drawn from parts A and C, giving a total sample size of 600 institutions.

The final column of Table 1 also gives the "Raw Sampling Weights" which were adjusted for non-response after the surveys were conducted. In so doing, final sampling weights were produced, which can be used for estimation purposes.

TABLE 1: Stratum Designations and Final Allocation for the CBMS 2010 Study (Program Types A, B, and C) and the MAA 2010 Study (Program Types A and C)

Stratum	Program Type	Highest Degree Granted	FTE Undergraduate Fall Enrollment	Universe (N)	Final Sample Allocation (n)	Sampling Rate (n/N)	Raw Sampling Weights (N/n)
1	Four-Year Math (A)	PhD	0-7,499	49	18	0.37	2.72
2			7,500-14999	55	35	0.64	1.57
3			15,000-19,999	43	25	0.58	1.72
4			20,000-24,999	25	17	0.68	1.47
5			25,000-34,999	19	19	1.00	1.00
6			35,000+	6	6	1.00	1.00
7		MA	0-6,999	85	28	0.33	3.04
8			7,000-10,999	52	13	0.25	4.00
9			11,000-14,999	23	16	0.70	1.44
10			15,000+	21	3	0.14	7.00
11		BA	0-999	193	16	0.08	12.06
12			1,000-1,499	201	14	0.07	14.36
13			1,500-2,499	271	25	0.09	10.84
14			2,500-4,999	244	39	0.16	6.26
15			5,000+	106	40	0.38	2.65
16	Four-Year Statistics (B)	PhD	0-14,999	17	17	1.00	1.00
17			15,000-24,999	23	23	1.00	1.00
18			25,000-34,999	11	11	1.00	1.00
19			35,000+	4	4	1.00	1.00
20		MA/BA	All	24	24	1.00	1.00
21	Two-Year Schools (C)	N/A	0-999	162	7	0.04	23.14
22			1,000-1,999	246	17	0.07	14.47
23			2,000-3,999	310	54	0.17	5.74
24			4,000-7,999	265	69	0.26	3.84
25			8,000-11,499	81	31	0.38	2.61
26			11,500-14,999	33	12	0.36	2.75
27			15,000-19,999	18	11	0.61	1.64
28			20,000+	6	6	1.00	1.00

Weighting Approach

Sampling weights adjusted for non-responding institutions were created for weighted data analysis. To facilitate the calculation of standard errors of estimates derived from the CBMS using the stratified jackknife method, replicate weights were created. Nonresponse adjustments were also applied to each set of replicate weights.

Sampling Weights

The raw sampling weight in the h^{th} stratum was computed as N_h/n_h, where N_h is the total number of institutions in the h^{th} stratum and n_h is the number of selected institutions in the h^{th} stratum. After the sample had been selected, a number of programs were identified as ineligible in their sampling strata, for the following reasons:

- Institutions have graduate programs only but were classified as a four-year program based on the sampling frame;
- Institutions no longer had mathematics (statistics) programs but were classified as a mathematics (statistics) program;
- Institutions were reclassified from a two-year mathematics program to a four-year mathematics program;
- Duplicate institutions (with different IDs on the sampling frame) were found.

In the weighting process, N_h and n_h were adjusted accordingly to account for these ineligible units. For example, N_h was reduced by the number of ineligible institutions in the h^{th} stratum. In the event that the ineligible institutions were sampled, n_h was also reduced by the number of ineligible institutions.

To remove bias from the estimates and reduce variability of the estimates, the raw sampling weights were adjusted for nonresponse. Within stratum h, a nonresponse adjustment factor, f_h, was calculated as

$$f_h = \frac{\sum_{eligible} W_h}{\sum_{responding} W_h}$$

where W_h is the raw sampling weight. Small cells in a stratum with less than 10 institutions or large nonresponse adjustment exceeding 2.5 were collapsed with an adjacent cell within program type and highest degree granted. The analysis weight, W_h^*, for any respondent in the h^{th} stratum was computed as

$$W_h{}^* = W_h f_h.$$

See Tables 2, 3, and 4 for the weights used in the four-year mathematics, four-year statistics, and two-year mathematics categories, respectively. Note that N_h' and n_h' in the tables reflect the number of eligible institutions in the h^{th} stratum.

Table 2. Final sampling weights used in the four-year mathematics questionnaire

Stratum (h)	Universe (N_h')	Number selected (n_h')	Number of completes (m_h)	Number of ineligibles	Response rate	Raw sampling weight (W_h)	Nonresponse adjusted factor (f_h)	Final sampling weight (W_h*)
1	49	18	14	1	0.778	2.722	1.286	3.500
2	55	35	26	1	0.743	1.571	1.346	2.115
3	43	25	21	0	0.840	1.720	1.190	2.048
4	25	17	11	0	0.647	1.471	1.545	2.273
5	18	18	14	1	0.778	1.000	1.200	1.200
6	6	6	6	0	1.000	1.000	1.200	1.200
7	85	28	18	0	0.643	3.036	1.658	5.032
8	52	13	7	0	0.538	4.000	1.658	6.631
9	23	16	12	0	0.750	1.438	1.408	2.024
10	21	3	2	0	0.667	7.000	1.408	9.856
11	192	15	8	1	0.533	12.800	1.510	19.323
12	201	14	11	0	0.786	14.357	1.510	21.674
13	270	25	19	0	0.760	10.800	1.316	14.211
14	244	39	27	0	0.692	6.256	1.444	9.037
15	106	40	34	0	0.850	2.650	1.176	3.118
Total	1,390	312	230	4	0.737			

Table 3. Final sampling weights used in the four-year statistics questionnaire

Stratum (h)	Universe (N_h')	Number selected (n_h')	Number of completes (m_h)	Number of ineligibles	Response rate	Raw sampling weight (W_h)	Nonresponse adjusted factor (f_h)	Final sampling weight (W_h*)
16	17	17	12	0	0.706	1.000	1.417	1.417
17	23	23	17	0	0.739	1.000	1.375	1.375
18	10	10	7	1	0.700	1.000	1.375	1.375
19	4	4	4	0	1.000	1.000	1.000	1.000
20	22	22	13	2	0.591	1.000	1.692	1.692
Total	76	76	53	3	0.697			

Table 4. Final sampling weights used in the two-year mathematics questionnaire

Stratum (h)	Universe (N_h')	Number selected (n_h')	Number of completes (m_h)	Number of ineligibles	Response rate	Raw sampling weight (W_h)	Nonresponse adjusted factor (f_h)	Final sampling weight (W_h*)
21	157	6	4	0	0.667	26.167	1.315	34.404
22	243	18	15	0	0.833	13.500	1.315	17.750
23	309	54	32	0	0.593	5.722	1.688	9.656
24	265	68	27	1	0.397	3.897	2.466	9.611
25	80	30	13	1	0.433	2.667	2.466	6.577
26	33	12	5	0	0.417	2.750	2.145	5.900
27	18	11	6	0	0.545	1.636	2.145	3.511
28	6	6	3	0	0.500	1.000	2.145	2.145
Total	1,111	205	105	2	0.512			

Replicate Weights

Weighted estimates and standard errors were calculated using a replication method, JKn (Jackknife method n, or the stratified jackknife method). The idea behind replication is to select subsamples (replicates) repeatedly from the whole sample, calculate the statistic of interest for each subsample, and then use these subsamples or replicate statistics to estimate the variance of the full-sample statistics. The JKn method divides the sample into subsamples by excluding one unit at a time.

For the CBMS, 68 replicates were created for the four-year mathematics program, and 60 replicates were created for the two-year mathematics programs. The replicates were designed in such a way so that on average, each replicate contained four to five sampled institutions. For the four-year statistics program, each sampled institution constituted a replicate, resulting in 75 replicates. The same nonresponse adjustment used for the full sample was applied to each replicate.

In stratum 6 and stratum 19, all the institutions were selected and all of them responded. These self-representing institutions were excluded from the computations involved in creating the replicate weights for non-self-representing institutions. Replicate weights associated with self-representing institutions were set equal to their full-sample weights. By handling the self-representing institutions in this manner, they were included in the population estimates but did not contribute to the resulting variance.

See Tables 5, 6, and 7 for the replicates for the four-year mathematics, four-year statistics, and two-year mathematics categories, respectively.

For variance estimation purposes, the "Stratum" in Tables 5, 6, and 7 is referred to as the variance stratum (VarStrat). The sampled institutions in a VarStrat are the variance units (VarUnits). For the first replicate weight, the full sample of institutions in the first VarStrat and VarUnit were multiplied by 0 and the weights associated with the other VarUnits in the same VarStrat and adjusted by $n_h'/(n_h'-1)$ to account for reducing the sample. The weights of the institutions in other VarStrat were not changed. The remaining replicates were formed in the same manner by systematically dropping each of the remaining VarUnits and computing the replicate weights as described for the first replicate.

Variance Estimation

Suppose that $\hat{\theta}$ is the full-sample estimate of some population parameter θ. The variance estimator using the JKn method, $v(\hat{\theta})$, is

$$v(\hat{\theta}) = \sum_{g=1}^{G} f_g h_g (\hat{\theta}_{(g)} - \theta)^2,$$

where

$\hat{\theta}_{(g)}$ is the estimate of θ based on the observations included in the g-th replicate,

G is the number of replicates formed,

f_g is the finite population correction (FPC) factors for replicate g, and

h_g is the JKn factors for replicate g.

The FPC is an adjustment to the estimated variance that accounts for how large a fraction of the population is selection for the sample. For replicate g, the FPC factor is $f_g = 1 - m_h/N_h'$, where m_h is the number of completes shown in Tables 2, 3, and 4. The JKn factor is computed as $h_g = (n_h' - 1)/n_h'$.

See Tables 5, 6, and 7 for the JKn factors and FPC factors for the four-year mathematics, four-year statistics, and two-year mathematics categories, respectively.

Table 5. Replicates, JKn factors, and FPC factors for the four-year mathematics program

Stratum (*h*)	Replicate (*g*)	Number of replicates	JKn factors	FPC factors
1	1-4	4	0.750	0.71
2	5-12	8	0.875	0.53
3	13-17	5	0.800	0.51
4	18-21	4	0.750	0.56
5	22-25	4	0.750	0.26
7	26-31	6	0.833	0.79
8	32-34	3	0.667	0.87
9	35-37	3	0.667	0.48
10	38-40	3	0.667	0.90
11	41-43	3	0.667	0.96
12	44-46	3	0.667	0.95
13	47-51	5	0.800	0.93
14	52-59	8	0.875	0.89
15	60-68	9	0.889	0.68

Table 6. Replicates, JKn factors, and FPC factors for the four-year statistics program

Stratum (*h*)	Replicate (*g*)	Number of replicates	JKn factors	FPC factors
16	1-17	17	0.941	0.29
17	18-40	23	0.957	0.26
18	41-51	11	0.909	0.36
20	52-75	24	0.958	0.46

Table 7. Replicates, JKn factors, and FPC factors for the two-year statistics program

Stratum (*h*)	Replicate (*g*)	Number of replicates	JKn factors	FPC factors
21	1-6	6	0.833	0.97
22	7-10	4	0.750	0.94
23	11-23	13	0.923	0.90
24	24-40	17	0.941	0.90
25	41-48	8	0.875	0.84
26	49-51	3	0.667	0.85
27	52-54	3	0.667	0.67
28	55-60	6	0.833	0.50

WesVar, a variance estimation software designed for complex surveys, was used to calculate estimates and standard errors of the estimates for the CBMS using the JKn replication method. WesVar can be used with a wide range of complex sample designs, including multistage, stratified, and unequal probability samples. The replicate variance estimates can reflect many types of estimation schemes, including nonresponse adjustment, poststratification, raking, and ratio estimation. It computes variance estimates for medians, percentiles, ratios, difference of ratios, and log-odds ratios.

Appendix II, Part II

Sampling and Estimation Procedures: Four-Year Mathematics and Statistics Faculty Profile

James W. Maxwell
American Mathematical Society

Overview

In CBMS surveys prior to 2005, information on the faculty was based on data collected on the CBMS form. Starting with the 2010 CBMS survey, the information on the faculty at four-year colleges and universities was based on a separate survey conducted by the American Mathematical Society under the auspices of the AMS-ASA-MAA-SIAM Data Committee. The Departmental Profile Survey is one of several surveys of mathematical sciences departments at four-year institutions conducted annually as part of the *Annual Survey of the Mathematical Sciences.* For 2010 the Departmental Profile Survey form was expanded to gather data on the age and the race/ethnicity of the faculty in addition to the data collected annually on rank, tenure status, and gender. The information on the four-year mathematics and statistics faculty derived from this data is presented in Chapters 1 and 4 of this report.

Using the faculty data collected by the 2010 Departmental Profile Survey reduced the size of the 2010 CBMS survey form. Furthermore, it eliminated the collection of the same faculty data on both surveys. In addition, coordination between the administrators of the Annual Survey and the CBMS survey allowed for minimizing the number of departments that were asked to complete both surveys.

Target Populations and Survey Approach

The procedures used to conduct the 2010 Departmental Profile survey are parallel to those used in CBMS 2010 as described in detail in Part I of this appendix. As with the CBMS 2010 survey, the primary characteristics used to stratify the departments for survey and reporting purposes are program type (four-year mathematics or four-year statistics) and the highest mathematical sciences degree offered by the department: doctoral, masters, or bachelors. The Departmental Profile survey employs a census of the doctoral mathematics departments whereas the CBMS survey samples these departments. In addition, the CBMS 2010 sample frame of statistics departments included twenty-four departments that offered at most a masters degree in statistics. These departments are not part of the regular Annual Survey sample frame but were included in the 2010 Departmental Profile survey. The Annual Survey reports separately on doctorate-granting departments of applied mathematics, but these departments are grouped with the doctoral departments of mathematics for the CBMS 2010 analysis. Finally, the Departmental Profile survey was sent to all masters-level mathematics departments and to double the number of bachelor-level departments: 267 compared to 134 for the CBMS survey.

Comparison of the Annual Survey Sample Frame with the CBMS Sample Frame

Table AS.1 demonstrates that the sample frames of four-year mathematics and statistics departments used in the two surveys closely align. As a consequence of this alignment, the distinction between the terms "Bachelors", "Masters", and "Doctoral" Mathematics Departments as defined in the two surveys is immaterial.

Table AS.1 Comparability of 2010 Annual Survey Sample Frame and the 2010 CBMS Sample Frame for Four-Year Mathematics Departments & Statistics Departments

Dept. Grouping	Annual Survey Count	CBMS Count	Overlap Count
Doctoral Math. Depts.	193	197	193
Masters Math. Depts.	180	181	177
Bachelors Math. Depts.	1012	1015	1011
Doctoral Stat. Depts.	54	55	54
Masters Stat. Depts.	22	24	22
Total	1461	1472	1457

Table AS.2 summarizes the stratifications used with the Departmental Profile and the allocation of the sample to the strata for the bachelors departments. This is the same stratification scheme used for CBMS 2010 and described in Part I of this appendix.

Survey Implementation

Departmental Profile forms were distributed in early January of 2011 asking departments to report on their fall-term 2010 faculty. Follow-up requests were sent to non-responding departments over the winter of 2011. The final effort to obtain responses took place during April in the form of phone calls to non-responding departments. The final efforts were concentrated on the strata with the lowest response rates.

Data Analysis

The data analysis used with the 2010 Departmental Profile survey parallels that used for CBMS 2010. The only notable variation is that if a non-responding department had completed a Departmental Profile survey within the previous three years, data from that survey was used to replace as much of the missing data for fall 2010 as feasible. This previously reported data consisted of the department's counts of faculty by rank, tenure-status, and gender. This technique was not possible for the fall 2010 data on faculty age and race/ethnicity since this information is not a part of previous Departmental Profile surveys.

The use of a department's prior-year faculty data to replace missing data for fall 2010 is supported by an ongoing review of annual faculty data from departments responding to the Departmental Profile in multiple years. Analysis of these data series demonstrates that the year-to-year variations in a given department's faculty data are, in general, much smaller than the department's deviation from the means for that department's stratum. Since the technique used to estimate the total for a stratum is equivalent to replacing the missing data with the average for the responding departments in that stratum, using prior responses is likely to produce a more accurate estimate of the total.

Table AS.2 lists the final sample weights used to produce the estimates within each stratum of the counts of faculty by rank, type of appointment, and gender. The column "Response rate" reflects the sum of the forms returned and the responses from prior years, when available. The sample weights used to produce estimates of age distribution and race/ethnicity distributions are higher in some strata since responses to those items were not available for prior years.

The standard errors reported for the faculty data are computed using the formulas described on pages 83-84 and 97-98 of [SMO].

Table AS.2 Stratum designations and allocations and nonresponse adjusted sample weights used with Annual Survey Data analyzed for CBMS 2010 report.

Stratum	Program Type	Highest Degree	Universe (N)	Number selected (n)	Sampling Rate (n/N)	Raw Sampling Weights (N/n)	Number of Responses	Number prior-year responses	Response rate	Nonresponse adjusted factor	Final sampling weights
1	4-year Math	PhD	48	48	1.000	1.000	42	3	0.938	1.067	1.067
2			54	54	1.000	1.000	46	6	0.963	1.038	1.038
3			43	43	1.000	1.000	40	3	1.000	1.000	1.000
4			24	24	1.000	1.000	22	1	0.958	1.043	1.043
5			18	18	1.000	1.000	15	2	0.944	1.059	1.059
6			6	6	1.000	1.000	6	0	1.000	1.000	1.000
7		MA	84	84	1.000	1.000	37	18	0.655	1.527	1.527
8			52	52	1.000	1.000	29	12	0.788	1.268	1.268
9			23	23	1.000	1.000	16	4	0.870	1.150	1.150
10			21	21	1.000	1.000	13	4	0.810	1.235	1.235
11		BA	191	35	0.183	5.457	8	2	0.286	3.500	19.100
12			201	35	0.174	5.743	14	3	0.486	2.059	11.824
13			270	50	0.185	5.400	19	5	0.480	2.083	11.250
14			244	85	0.348	2.871	37	11	0.565	1.771	5.083
15			106	62	0.585	1.710	37	15	0.839	1.192	2.038
16	4-year Stat	PhD	17	17	1.000	1.000	12	3	0.882	1.133	1.133
17			23	23	1.000	1.000	18	3	0.913	1.095	1.095
18			10	10	1.000	1.000	6	3	0.900	1.111	1.111
19			4	4	1.000	1.000	4	0	1.000	1.000	1.000
20		MA	22	22	1.000	1.000	10	0	0.455	2.200	2.200