

## Appendix II, Part I

# Sampling and Estimation Procedures

Leela M. Aertker and Robert P. Agans

*The Survey Research Unit*

*The University of North Carolina, Chapel Hill, North Carolina*

### Overview

A stratified, simple random sample was employed in the CBMS 2005 survey, and strata were based on three variables: curriculum, highest degree level offered, and total institutional enrollment. A paper-and-pencil data collection method was implemented between the months of September 2005 and May 2006, and all resulting estimates were generated in an SAS-Callable version of SUDAAN using a stratified-sampling-without-replacement design. This report is divided into the following two sections: Sampling Approach and Survey Design.

### Sampling Approach

A stratified, simple random sample of 600 two-year and four-year colleges and universities was employed in CBMS 2005. A compromise mix of statistically optimum Neyman allocations based on two key outcome variables was used to determine targeted sample sizes for the 24 sampling strata.

### Target Population and Sampling Frames

The target population of the CBMS 2005 survey consisted of undergraduate mathematics and statistics programs at two-year and four-year colleges and universities in the United States. In most cases, these programs were established academic departments whereas others were fledgling departments or other types of curriculum concentrations. A total of 2,459 programs were identified as eligible for participation in the survey. Sample selection was made from a merged program frame of 1,417 mathematics programs at four-year colleges and universities, 67 statistics programs at four-year colleges and universities, and 975 mathematics programs at two-year colleges.

### Selection of Stratification Variables

Prior to selecting the sample for the CBMS 2005 and CBMS 2000 surveys, the stratification variables used in the CBMS 1995 survey were examined to

determine their significance in predicting specific key outcome variables in each of the programs surveyed and thus, their utility for stratification in future CBMS surveys. This was done because the utility of a variable for stratification in generating estimates from a stratified sample depends on its statistical correlation with important measurements made on the sample.

Stratification in the CBMS 1995 survey was accomplished as follows: universities and colleges were separately divided into 20 strata based on curriculum (four-year mathematics programs, four-year statistics programs, or two-year mathematics programs), control (publicly or privately funded), level (the highest degree offered—BA, MA, or PhD), and enrollment (total institutional enrollment for Fall 1995). Our analysis of the CBMS 1995 data showed that curriculum, level, and enrollment would be the best stratification variables for producing estimates for future CBMS target populations. It was, therefore, decided not to stratify by each program's public or private classification as only minimal strength in predicting key outcome variables was gained by using this stratification variable.

The final stratum designations for the CBMS 2005 survey follow the exact stratum designations for the CBMS 2000 survey and very closely follow the stratum designations for the CBMS 1995 survey with the exception of control as a stratification variable. The four-year mathematics programs were divided into 12 strata, the four-year statistics programs were divided into five strata, and the two-year programs were divided into seven strata. Table A2.1 displays the overall stratum breakdown (24 strata total).

### Allocation Process

For purposes of consistency in design development strategy, the same approach as used in CBMS 2000 was followed to determine the allocation of the CBMS 2005 sample. For CBMS 2005, stratum designations were assigned, key outcome variables were selected, and a multi-variable Neyman allocation was implemented in two iterations so that comparable precision

was produced for each frame with the same number of schools expected to respond as in CBMS 2000.

Three program frames were sent to us by the study directors. Each frame included colleges and universities who were thought to offer undergraduate programs in four-year mathematics, four-year statistics, and two-year mathematics programs. The goal of sample selection was to select a representative sample of programs from each of the three frames. The sample was stratified by curriculum (four-year mathematics programs, four-year statistics programs, or two-year mathematics programs), level (the highest degree offered—BA, MA, or PhD), and enrollment (total institutional enrollment for Fall 2005).

The same key outcome variables from CBMS 2000 were once again proposed by the study directors in CBMS 2005; namely, total fall enrollment and number of full-time faculty. An additional outcome variable, number of baccalaureate degrees awarded, was also proposed, but this information was only collected for strata involving four-year institutions (i.e., strata 1–17). The variances of the two key outcome variables that were considered for purposes of allocation decisions, total fall enrollment and total full-time faculty, were estimated for each stratum using CBMS 2000 respondent data.

A multi-variable Neyman allocation was implemented to determine the optimum sample sizes for the strata within each frame, which would produce the most cost-effective allocation of the sample. This type of allocation samples more intensely from strata with more diversity or variability. The sample allocation intended to produce estimates of comparable precision for each of the three frames (four-year mathematics programs, four-year statistics programs, or two-year mathematics programs). This was done so that estimates aimed at the three frames would have approximately equal precision.

For CBMS 2005, it was determined that the same number of schools would be selected as in CBMS 2000 (i.e.,  $n = 600$ ). Due to refusals and unforeseen ineligibles, not all institutions selected would consequently respond. Thus, we intended to select a sample for CBMS 2005 that was *expected* to produce the same number of participating institutions as in CBMS 2000 (i.e.,  $m = 392$ ). The simple variance of each key outcome variable in each frame was calculated by using CBMS 2000 respondent data. The expected number of participating programs in each frame ( $m_g$ ) was determined by the constraint that the variances of each frame were equivalent ( $V_1 = V_2 = V_3$ ). A weighted average of the subgroup allocations was computed; however, this compromise mix of subgroup allocations called for sampling more four-year statistics programs than were on the frame. Therefore, the expected number to respond in the four-year statistics programs was set to the maximum expected number

to respond ( $m_2 = 47$ ) based on a realistic response rate for the particular subgroup.

The number expected to respond in the four-year and two-year mathematics program frames was then determined by the constraint that the variances of the four-year and two-year mathematics programs were equivalent ( $V_1 = V_3$ ). A compromise mix of the expected number of programs to respond in the subgroup allocations was determined by giving the subgroup allocation based on total fall enrollment a relative weight of 0.75 and the subgroup allocation based on the number of full-time faculty a relative weight of 0.25. A larger relative weight was given to the subgroup allocation based on total fall enrollment since this variable, according to the study directors, was more salient to the study. The resulting subgroup allocation was as follows: expected number to respond for the four-year mathematics programs ( $m_1$ ) = 202, expected number to respond for the two-year mathematics programs ( $m_3$ ) = 143, and expected number to respond for the four-year statistics programs ( $m_2$ ) = 47.

Separate Neyman allocations were then conducted for the four-year and two-year mathematics programs. The first Neyman allocation iteration produced two different sets of allocations among the strata—one based on total fall enrollment and the other based on full-time faculty. A minimum expected number of seven responding programs in each stratum was set unless seven exceeded the total stratum size times the CBMS 2000 response rate. In the latter case, the minimum expected number was the maximum number of expected respondents. By applying this rule, we set the minimum expected number of responding programs and computed a second iteration of the Neyman allocation for the 15 strata whose first iteration allocations exceeded the minimum standard.

The final sample allocation was anchored to the allocation produced by the key outcome variable, total fall enrollment, since this outcome variable was more salient to the study, according to the study directors. Modifications to the allocation based on total fall enrollment were made in consideration of sample size needs vis-à-vis the allocation based on total full-time faculty. Accordingly, a weighted average of the two second iteration allocations was computed based on total fall enrollment (given a relative weight of 0.75) and total full-time faculty (given a relative weight of 0.25) to produce the compromise mix of allocations in the four-year and two-year mathematics categories. Once the optimum allocation was determined, the number of selected programs in each stratum was calculated based on CBMS 2000 response rates. To obtain comparable precision for estimates aimed at the three frames, more participating four-year statistics programs were called for than were on the frame. Thus, for the four-year statistics frame, we simply

took the maximum number of programs expected to respond and selected all programs in the frame. Table A2.1 lists the final agreed allocation and the sampling

rate of the 600 selected programs for the CBMS 2005 survey.

Table A2.1 Stratum Designations and Final Agreed Allocation for the CBMS 2005 Study

Stratum	Curriculum	Level	Enrollment	Final Agreed Allocation	Sampling Rate
1	Four-Year Math	PhD	0 – 14,999	37	0.3627
2			15,000 – 24,999	54	0.8438
3			25,000 – 34,999	15	0.7500
4			35,000 +	6	1.0000
5		MA	0 – 6,999	17	0.2208
6			7,000 – 14,999	21	0.2414
7			15,000 +	12	0.4800
8		BA	0 – 999	16	0.0874
9			1,000 – 1,499	17	0.0846
10			1,500 – 2,499	30	0.1024
11			2,500 – 4,999	26	0.1130
12			5,000 +	41	0.3178
13	Four-Year Statistics	PhD	0 – 14,999	20	1.0000
14			15,000 – 24,999	23	1.0000
15			25,000 – 34,999	9	1.0000
16			35,000 +	3	1.0000
17		MA/BA	All	12	1.0000
18	Two-Year Schools	N/A	0 – 999	12	0.1519
19			1,000 – 1,999	16	0.1096
20			2,000 – 3,999	35	0.1378
21			4,000 – 7,999	64	0.2540
22			8,000 – 14,999	51	0.3312
23			15,000 – 19,999	26	0.6500
24			20,000+	37	0.7400
				600 programs	

### Sample Selection

The SurveySelect procedure in SAS Version 8.2 was used to select the allocation from the merged program frame. We employed a stratified simple random sample design with three stratification variables (i.e., curriculum, level, and enrollment). The N= option specified the sample sizes for each of the 24 strata.

### Survey Design

This section describes data collection, analysis procedures, and final weight construction.

### Survey Implementation

Data collection occurred over a nine-month period. An advance letter was sent out to all respondents informing them that they were selected to participate and that they would receive the CBMS 2005 questionnaire within the next couple of weeks. All questionnaires were mailed out August 29, 2005 and a postcard was sent out at the end of October to either remind participants to respond or to thank them for their participation. A second batch of questionnaires was mailed out to all nonrespondents in the beginning of November. Questionnaires were accepted until an extended deadline of May 15, 2006.

## Data Analysis

SUDAAN is a statistical package of choice when analyzing data from complex sample surveys. This software is advantageous since it allows the user to compute not only estimates such as totals and ratios, but also the standard errors of those estimates in accordance with the sample design. Many statistical packages are capable of computing population estimates, but the standard errors are based on simple random sampling; thus, they produce standard errors that are inappropriate for more complex designs. SUDAAN uses first-order Taylor series approximation procedures in generating the standard errors, which tend to be more accurate than estimates from other statistical packages. The sample design used in this study and incorporated in SUDAAN was stratified sampling without replacement (STRWOR).

For quality control purposes, all questionnaires were doubly entered by data entry personnel at the Survey Research Unit (SRU) at the University of North Carolina at Chapel Hill, and most discrepancies between the two files were settled by review of the original document. In a few cases, however, the respondents had to be contacted to clarify discrepancies. The bulk of data cleaning occurred between the

months of May and July 2006. Data analysis took place between the months of May and August 2006.

## Sample weights

For any respondent in the  $h^{th}$  stratum, the nonresponse adjusted sample weight was computed as follows:

- Raw Weight =  $N_h / n_h$
- Response Rate (RR) =  $m_h / (n_h - i_h)$
- Adjusted weight = Raw Weight \* (1/RR)

where,

$N_h$  = the total number of programs in the  $h_{th}$  stratum

$n_h$  = the number of selected programs in the  $h_{th}$  stratum

$m_h$  = the number of (eligible) respondents in the  $h_{th}$  stratum

$i_h$  = the number of study ineligibles in the sample for the  $h_{th}$  stratum

See Tables A2.2, A2.3, and A2.4 for the weights used in the four-year mathematics, four-year statistics, and two-year mathematics categories, respectively.

Table A2.2 Nonresponse Adjusted Sample Weights Used in the Four-Year Mathematics Questionnaire

Stratum	Total ( $N_h$ )	Number Selected ( $n_h$ )	Number of completes ( $m_h$ )	Number of ineligibles ( $i_h$ )	Response rate (RR)	Program level raw weight	Program level adjusted weight
1	102	37	30	0	0.811	2.757	3.400
2	64	54	34	1	0.642	1.185	1.847
3	20	15	10	0	0.667	1.333	2.000
4	6	6	3	1	0.600	1.000	1.667
5	77	17	14	0	0.824	4.529	5.500
6	87	21	14	0	0.667	4.143	6.214
7	25	12	6	0	0.500	2.083	4.167
8	183	16	8	0	0.500	11.438	22.875
9	201	17	8	0	0.471	11.824	25.125
10	293	30	14	0	0.467	9.767	20.929
11	230	26	13	1	0.520	8.846	17.012
12	129	41	22	0	0.537	3.146	5.864
<b>Total</b>	<b>1417</b>	<b>292</b>	<b>176</b>	<b>3</b>	<b>0.609</b>	-	-

Table A2.3 Nonresponse Adjusted Sample Weights Used in the Statistics Questionnaire

Stratum	Total ( $N_h$ )	Number Selected ( $n_h$ )	Number of completes ( $m_h$ )	Number of ineligible ( $i_h$ )	Response rate (RR)	Program level raw weight	Program level adjusted weight
13	20	20	12	0	0.600	1.000	1.667
14	23	23	12	2	0.571	1.000	1.750
15	9	9	7	0	0.778	1.000	1.286
16	3	3	2	0	0.667	1.000	1.500
17	12	12	6	0	0.500	1.000	2.000
<b>Total</b>	<b>67</b>	<b>67</b>	<b>39</b>	<b>2</b>	<b>0.600</b>	-	-

Table A2.4 Nonresponse Adjusted Sample Weights Used in the Two-Year Mathematics Questionnaire

Stratum	Total ( $N_h$ )	Number Selected ( $n_h$ )	Number of completes ( $m_h$ )	Number of ineligible ( $i_h$ )	Response rate (RR)	Program level raw weight	Program level adjusted weight
18	79	12	6	0	0.500	6.583	13.167
19	146	16	9	0	0.563	9.125	16.222
20	254	35	18	0	0.514	7.257	14.111
21	252	64	30	0	0.469	3.938	8.400
22	154	51	29	0	0.569	3.020	5.310
23	40	26	15	1	0.600	1.538	2.564
24	50	37	23	0	0.622	1.351	2.174
<b>Total</b>	<b>975</b>	<b>241</b>	<b>130</b>	<b>1</b>	<b>0.542</b>	-	-

### Analysis Plan

To expedite analysis, protocols were developed in advance. Each protocol identified the variables involved, any mathematical transformations, the type of parameter being estimated, the procedure used to estimate the parameter, the units in which the estimate was to be reported, and any domain variables

used to compartmentalize the variables. All protocols were subject to review by the CBMS director and approved before any estimates were generated. Table A2.5 is an example of the protocol used to construct a portion of the table FY.1 on page 114. All variables and resulting calculations were defined in an attempt to eliminate ambiguity.

TABLE A2.5 Example of Analysis Protocol: Portion of Table FY.1 (page 114).

Key <sup>1</sup>	Variable	Description	Numerator	Denominator	Parameter Type	SUDAAN Procedure <sup>3</sup>	Unit	Domain Variable <sup>4</sup>
4M	C64	No. of Mathematics for Liberal Arts sections	C64					
4M	C65	No. of sections taught by tenured/tenure eligible faculty	C65		Percentage	Ratio	Sections	HDO_Math
4M	PC65	Percentage of sections taught by tenured/tenure eligible faculty	C65	C64	Percentage	Ratio	Sections	HDO_Math
4M	C66_67	No. of sections taught by other full-time (total) faculty	SUM(C66,C67);	C64	Percentage	Ratio	Sections	HDO_Math
4M	PC66_67	Percentage of sections taught by other full-time (total) faculty	C66_67	C64	Percentage	Ratio	Sections	HDO_Math
4M	C66	No. of sections taught by other full-time (doctoral) faculty	C66		Percentage	Ratio	Sections	HDO_Math
4M	PC66	Percentage of sections taught by other full-time (doctoral) faculty	C66	C64	Percentage	Ratio	Sections	HDO_Math
4M	C68	No. of sections taught by part-time faculty	C68		Percentage	Ratio	Sections	HDO_Math
4M	PC68	Percentage of sections taught by part-time faculty	C68	C64	Percentage	Ratio	Sections	HDO_Math
4M	C69	No. of sections taught by graduate teaching assistants	C69		Percentage	Ratio	Sections	HDO_Math
4M	PC69	Percentage of sections taught by graduate teaching assistants	C69	C64	Percentage	Ratio	Sections	HDO_Math
4M	C6UNK	No. of sections taught by unknown faculty	C64 - (SUM (C65, C66,C67,C68,C69));	C64	Percentage	Ratio	Sections	HDO_Math
4M	PC6UNK	Percentage of sections taught by Unknown faculty	C6UNK	C64	Percentage	Ratio	Sections	HDO_Math
4M	C63	Total On-campus enrollment in Mathematics for Liberal Arts	C63		Average	Ratio	Sections	HDO_Math
4M	AC64	Average size of Mathematics for Liberal Arts sections	C63	C64	Average	Ratio	Sections	HDO_Math

<sup>1</sup> Key: 4M=4-Yr Math, 4S=4-Yr Stat, 2M=2-Yr Math

<sup>2</sup> Blank boxes in the questionnaire tables were interpreted zeros

<sup>3</sup> Estimates weighted by final adjusted weight to produce national estimates

<sup>4</sup> HDO\_Math (Highest Degree Offered for Four Year Math Departments)

### Manipulation Checks

Because of the complex nature of the questionnaire, several manipulation checks were performed on the data before analyses proceeded. If a discrepancy could not be settled by reviewing the questionnaire, the respondent was called or emailed to settle it. No imputations were made for missing data. In fact, blank boxes in questionnaire tables were interpreted as zeros since many respondents refused to fill in all of the boxes. Hence, it was impossible to tell the difference between missing values and zeros in the questionnaire tables.

### Generation of Information Products

All analyses were generated using a SAS-Callable version of SUDAAN (Version 9.01). To ease interpretation, the SUDAAN output was exported to Excel spreadsheets and sent to the CBMS director, which were transferred into production table shells. See Table A.2.6 for an example of the SUDAAN output that refers to the percentage of sections of one particular course taught by faculty with various appointments and the average section size in four-year mathematics departments by school type (or highest degree offered—HDO). All estimates were produced in a similar manner.

TABLE A2.6 Example of SUDAAN Output: Portion of Table FY.1 (page 114).

Estimate Description	Highest Degree Offered			OVERALL
	PhD	MA	BA	
Four-Year Mathematics	Percentage (SE)	Percentage (SE)	Percentage (SE)	Percentage (SE)
Mathematics for Liberal Arts				
Percentage of sections taught by tenured/tenure eligible faculty	18.04% (3.09%)	35.55% (5.66%)	42.93% (5.70%)	36.71% (3.53%)
Percentage of sections taught by other full-time (total) faculty	18.93% (3.99%)	13.12% (4.09%)	15.65% (3.71%)	15.50% (2.45%)
Percentage of sections taught by other full-time (doctoral) faculty	5.29% (1.55%)	3.81% (1.74%)	4.11% (1.82%)	4.22% (1.15%)
Percentage of sections taught by part-time faculty	27.62% (3.55%)	37.90% (6.15%)	31.93% (7.03%)	32.86% (4.29%)
Percentage of sections taught by graduate teaching assistants	24.86% (4.72%)	3.47% (2.40%)	0% (0%)	5.13% (1.36%)
Percentage of sections taught by Unknown faculty	10.54% (4.55%)	9.96% (4.69%)	9.49% (4.03%)	9.80% (2.70%)
	Average (SE)	Average (SE)	Average (SE)	Average (SE)
Average size of Mathematics for Liberal Arts sections	45.95 (3.10)	33.87 (2.29)	24.75 (1.18)	30.83 (1.07)

## Appendix II, Part II

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

James W. Maxwell

*American Mathematical Society*

### Overview

In all previous CBMS surveys, data on the faculty were collected on the CBMS form. For CBMS 2005, the information on the faculty at four-year colleges and universities provided in this report is derived from a separate survey conducted by the American Mathematical Society under the auspices of the AMS-ASA-IMS-MAA-SIAM Data Committee. The “Departmental Profile – Fall 2005” is one of a series of surveys of mathematical sciences departments at four-year institutions conducted annually as part of the *Annual Survey of the Mathematical Sciences*. In 2005 this survey was expanded to gather data on the age and the race/ethnicity of the faculty, in addition to the usual 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 in the 2005 Annual Survey reduced the size of the 2005 CBMS survey form. Furthermore, it eliminated the collection of the same faculty data on both surveys. 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 2005 Departmental Profile survey are very similar to

those used in CBMS 2005, described in detail in the preceding pages of this appendix. The primary characteristic used to group the departments for survey and reporting purposes is the highest mathematical sciences degree offered by the department: doctoral, masters, or bachelors, the same groupings used by CBMS 2005. There are some notable differences. The Departmental Profile survey uses a census of the doctoral mathematics and statistics departments, and it surveys only the doctoral statistics departments. There were twelve departments in the CBMS 2005 sample frame of statistics departments that offered at most a bachelors or masters degree. These departments are not represented in the description of the faculty at the doctoral statistics departments.

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

Table AS.1 demonstrates that the sample frames of four-year mathematics departments used in the two surveys are in extremely close alignment. 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. Furthermore, the estimates produced from each of the surveys may be applied interchangeably to these groupings of departments.

**Table AS.1 Comparability of 2005 Annual Survey Sample Frame and the 2005 CBMS Sample Frame for Four-Year Mathematics Departments**

<b>Dept. Grouping</b>	<b>Annual Survey Count</b>	<b>CBMS Count</b>	<b>Overlap Count</b>
Bachelors Depts.	1036	1036	1030
Masters Depts.	190	189	188
Doctoral Depts.	196	192	188
Total	1422	1417	1406

### **Sampling Masters and Bachelors Departments at Four-Year Institutions**

While the Annual Survey employs a census of the doctoral mathematics and statistics departments, it uses a stratified, random sample of the masters and bachelors departments. The masters and bachelors departments are stratified by control (public or private)

and by total institutional undergraduate enrollment. Table AS.2 summarizes the stratifications used for the Departmental Profile and the allocation of the sample to the strata for the masters and bachelors departments.

**Table AS.2 Stratum Designations and Allocations for the 2005 Departmental Profile Survey**

<b>Stratum</b>	<b>Curriculum</b>	<b>Level</b>	<b>Institutional Enrollment</b>	<b>Sample Allocation</b>	<b>Sampling Rate</b>
1	Four-Year Math	PhD	All	196	1.0000
2	(Public)	MA	0 – 5,999	12	0.4444
3			6,000 – 8,999	21	0.5526
4			9,000 – 11,999	21	0.5833
5			12,000 – 17,999	22	0.5641
6			18,000 +	10	0.5559
7	(Private)	MA	0 – 3,999	6	0.5000
8			4,000 – 7,999	5	0.4545
9			8,000 +	3	0.3333
10	(Public)	BA	0 – 1,999	22	0.3548
11			2,000 – 3,999	31	0.3605
12			4,000 – 6,999	40	0.5063
13			7,000 – 11,999	21	0.7778
14			12,000 +	10	0.6667
15			Military academies	2	0.6667
16	(Private)		0 – 999	48	0.2667
17			1,000 – 1,499	49	0.3161
18			1,500 – 1,999	65	0.4815
19			2,000 – 3,999	70	0.3483
20			4,000 – 6,999	24	0.3582
21			7,000 – 8,999	10	0.6250
22			9,000 +	4	0.4000
23	Four-Year Statistics	PhD	All	56	1.0000
				748 departments	

## Survey Implementation

Departmental Profile forms were mailed in late September 2005 with a due date of October 30th to all doctoral-granting mathematics and statistics departments and to a sampling of the masters- and bachelors-granting departments of mathematical sciences at four-year colleges and universities in the U.S. A second mailing of forms was sent to non-responders in early November with a due date of December 6th. A third mailing was sent via email at the end of January 2006 providing a link to an interactive PDF version of the form with a due date in early February. The final effort to obtain responses took place during February through March in the form of phone calls to non-responding departments. The final efforts were concentrated on the stata with the lowest response rates.

## Data Analysis

The data analysis used with the 2005 Departmental Profile survey parallels that used by CBMS 2005. 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 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 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 2005 is supported by

a 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 highly likely to be smaller than the department's variation from the mean data for that department's stratum. Since the technique used to estimate a total for a stratum is equivalent to replacing the missing data with the average for the responding departments in that stratum, using prior responses to the same question is likely to produce a more accurate estimate of the total.

Table AS.3 lists the program-level adjusted sample weights used to produce the estimates within each stratum of counts of faculty by rank, type-of-appointment and gender. The column "Number of Completes" displays the total of the forms returned plus the responses from prior years when available. (Compare with Table A2.2 in Appendix II.) The adjusted weights used to produce estimates of age distribution and race/ethnicity distributions are slightly higher 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]. For the doctoral mathematics departments, use of prior-year responses produced a 100% response rate for certain items, hence the contribution of the doctoral mathematics departments to the standard errors for those items was zero.

**Table AS.3 Nonresponse Adjusted Sample Weights Used with the 2005 Departmental Profile Questionnaire.**

Stratum	Total	Number Selected	Number of completes	Number of Prior-year Resp. used	(Final) Response rate	Program level raw weight	Program level adjusted weight
1	196	196	163	33	1.000	1.000	1.000
2	27	12	5	3	0.667	2.250	3.375
3	38	21	13	3	0.762	1.810	2.375
4	36	21	13	2	0.714	1.714	2.400
5	39	22	12	5	0.773	1.773	2.294
6	18	10	7	3	1.000	1.800	1.800
7	12	6	2	1	0.500	2.00	4.000
8	11	5	4	0	0.800	2.200	2.750
9	9	3	3	0	1.000	3.000	3.000
10	62	22	2	2	0.182	2.818	15.500
11	86	31	13	1	0.452	2.774	6.143
12	79	40	23	6	0.725	1.975	2.724
13	27	21	14	2	0.762	1.286	1.688
14	15	10	5	1	0.600	1.500	2.500
15	3	2	2	0	1.000	1.500	1.500
16	180	48	15	1	0.333	3.750	11.250
17	155	49	16	3	0.388	3.163	8.158
18	135	65	26	6	0.492	2.077	4.219
19	201	70	34	5	0.557	2.871	5.154
20	67	24	13	5	0.750	2.792	3.722
21	16	10	4	1	0.500	1.600	3.200
22	10	4	2	0	0.500	2.500	5.000
23	56	56	39	16	0.982	1.000	1.018