DEVELOPING A MODEL FOR NONRESPONSE BIAS ANALYSIS

Maxime Bokossa, Sameena Salvucci, Dhirendra Ghosh, Synectics for Management Decisions Inc.;
Marilyn Seastrom, Steven Kaufman, National Center of Education Statistics; Daniel Kasprzyk,
Mathematica Policy Research, Inc.
Sameena Salvucci, Synectics, 1901 N. Moore Street, Suite 900, Arlington, Virginia 22209

KEY WORDS: Unit nonresponse, Item nonresponse, Relative bias, Bias ratio, Induced bias

1. Introduction

In this paper we present techniques for conducting unit and item level nonresponse bias analyses to be used by National Center for Education Statistics (NCES) program managers when response rates for their surveys fall below NCES minimum standards. Data from NCES' 1999–2000 Schools and Staffing Survey (SASS) were used to illustrate these techniques.

The 1999-2000 SASS provides data on public and private schools, principals, school districts, and teachers. SASS gathers information about many topics, including various characteristics of elementary and secondary students, some of the professional and paraprofessional staff who serve them, the programs offered by schools, principals' and teachers' perceptions of school climate and problems in their schools, teacher compensation, and district hiring practices. SASS is a unified set of surveys that facilitates comparison between public and private schools and allows linkages of teacher, school, school district, and principal data. SASS has been administrated four times since 1987–88.

2. SASS Sample Design

SASS uses a stratified probability sample design. Details of stratification variables, sample selection, and frame sources are provided below.

Schools are selected first. For the *public school* sample, the first level of stratification is by these five types of schools: a) BIA schools; b) Native American schools (i.e., schools with 19.5 percent or more Native American students); c) schools in Delaware, Nevada, and West Virginia (where it is necessary to implement a different sampling methodology to select at least one school from each LEA in the state); d) charter schools; and e) all other schools. Schools falling into more than one group are assigned in hierarchical order in the second level of stratification, first by state and then by LEA. There are three grade level strata (elementary, secondary, and combined schools).

Within each stratum, all non-BIA and non-Charter schools are systematically selected using a probability proportionate to size algorithm. The measure of size used in SASS for schools in the SASS frame, i.e. the Common Core of Data (CCD), was the square root of the number of teachers in the school. Any school with a measure of size larger than the sampling interval was excluded from the probability sampling operation and included in the sample with certainty.

3. Unit Response Rates

The NCES standard for calculating unit response rates is as follows:

- *Unit Response Rate (RR)*

$$\text{Weighted RR} = \frac{\text{Weighted \# of completed interviews}}{\text{Wtd \# of units sampled - Wtd \# out of scope}}$$

$$\text{Unweighted RR} = \frac{\text{\# of completed interviews}}{\text{\# of units sampled - \# out of scope}}$$

- *Overall Response Rate for 2-stage sampling*

$$\text{Overall RR} = (\text{unit RR at stage 1}) \times (\text{unit RR at stage 2})$$

NCES standards for minimum unit response rates is 75 percent. Thus, the first step in examining bias included identifying which SASS questionnaire response rates fell below the NCES minimum standards.

Overall weighted (by base weight) unit response rates were calculated for each of 16 SASS questionnaires as well as four teacher listing forms (table 1) as defined by the six different types of SASS respondents and four types of SASS sectors:

Table 1.—SASS Questionnaire Weighted Unit Response Rates

Respondent	Sector			
	Public	*Private*	*Indian*	*Public Charter*
District	School District Questionnaire: 88.56	N/A	N/A	N/A
Principal	Public School Principal Questionnaire: 90.01	Private School Principal Questionnaire: 84.81	Indian School Principal Questionnaire: 93.28	Public Charter School Principal Questionnaire: 90.18
School	Public School Questionnaire: 88.45	Private School Questionnaire: 79.78	Indian School Questionnaire: 96.67	Public Charter School Questionnaire: 86.14
Teacher Listing Form	Public School Teacher Listing Form: 92.42	Private School Teacher Listing Form: 87.02	Indian School Teacher Listing Form: 97.50	Public Charter School Teacher Listing Form: 91.46
Teacher	Public School Teacher Questionnaire: 83.08	Private School Teacher Questionnaire: 77.29	Indian School Teacher Questionnaire: 87.40	Public Charter School Teacher Questionnaire: 78.57
Overall	76.78	67.26	85.22	71.86
Library	Public School Library Media Center Questionnaire: 94.73	Private School Library Media Center Questionnaire: 87.72	Indian School Library Media Center Questionnaire: 95.41	N/A

N/A = Not applicable

Within each type of questionnaire, weighted response rates were calculated by key characteristics such as region, community type, school level, and student enrollment categories.

In addition, weighted unit response rates were calculated for:
- Each *state* for the **all public and sector charter school** questionnaires
- Each *typology* for **all private school sector** questionnaires

The following summarizes response rates for SASS questionnaires where the unit response rates examined fell below 75 percent.

3.1 Public school sector questionnaires.
Regarding the region, community type, school level, and student enrollment characteristics, all of the school district, public school principal, and public school weighted unit response rates were above 75 percent. The overall public school teacher weighted response rate was below 75

percent in three instances: under region, the response rate was 69.82 percent for the Northeast; for community type, the response rate for central city was 71.58 percent; and for student enrollment, the response rate was 73.85 percent for the category 1,000 or more.

All four questionnaires had response rates below 75 percent for at least one state. For the School District Questionnaire, response rates for four states fell below 75 percent: Maryland (74.63 percent), New Hampshire (73.83 percent), Rhode Island (73.26 percent), and Vermont (68.89 percent).

The Public School Principal Questionnaire and the Public School Questionnaire both had a response rate below 75 percent for one state: Maryland (72.54 and 67.99 percent, respectively).

While only one unit response rate for the Public School Teacher Listing Form or the Public

School teacher Questionnaire fell below 75 percent, the overall weighted unit response rates for the Public School Teacher Questionnaire fell below 75 percent for 13 states.

3.2 Private school sector questionnaires and Teacher Listing Form.

Regarding the region, community type, school level, and student enrollment characteristics, the private school principal and the private school library media center questionnaires as well as the Private School Teacher Listing Form had no response rates below 75 percent. The Private School Questionnaire had two rates below 75 percent: for the student enrollment category "less than 50," the response rate was 74.99 percent, and for the school level category "combined," the response rate was 72.63 percent.

The weighted unit response rates for the Private School Teacher Listing Form fell below 75 percent for three affiliations: Other Jewish (71.00 percent), Hebrew Day (70.25 percent), and American Association of Christian Schools (65.38 percent).

The weighted unit response rates for the Private School Teacher Questionnaire fell below 75 percent for the same three affiliations, plus one other: American Association of Christian Schools (69.79 percent), Montessori (68.41), Hebrew Day (59.84 percent), and Other Jewish (55.89 percent). The weighted unit response rates for the Private School Teacher Questionnaire also fell below 75 percent for the NCES typology "Nonsectarian—Special emphasis" (66.26 percent) and for the student enrollment category "Less than 100" (69.54 percent).

The overall weighted unit response rates for the Private School Teacher Questionnaire fell below 75 percent for at least one category in all four characteristics. For region, the response rates were below 75 percent for the West (66.50 percent), South (63.75 percent), and the Northeast (63.18 percent). The response rates were below 75 percent for all three categories of community type: rural/small town (71.24 percent), urban fringe/large town (67.56 percent), and central city (65.43 percent). The response rate was below 75 percent for two categories of school level: elementary (72.01 percent) and secondary (69.99 percent). The response rate was below 75 percent for all categories of student enrollment: 1,000 or more (73.85 percent), 100 to 199 (71.44 percent), 200

to 499 (68.20 percent); 500 to 749 (65.20 percent), 750 to 999 (61.48 percent), and fewer than 100 (60.35 percent).

Response rates were also calculated for affiliation and NCES typology for the private school questionnaires. The response rates that fell below 75 percent are shown in the table 2 below.

3.3 Public charter school sector questionnaire.

The weighted response rates for the Public Charter Teacher Questionnaire and for the Public Charter School Teacher Listing Form were all above 75 percent, but the overall weighted unit response rate fell below 75 percent for at least one category of all the characteristics examined.

Region:	Northeast (72.55%, West (70.69%) South (69.72%)
Community type:	Urban fringe/large town (74.97%), Central city (68.40%)
School level:	Elementary (73.91%) Secondary (70.71%) Combined (70.54 %)
Student enrollment:	500 or more (72.52%) 100 to 145 (71.09%) 150 to 499 (70.03%)
States:	"Others"[1] (71.03%) California (70.35%) Arizona (69.68%) Texas (65.30%)

3.4. Comparison of Sample Estimates with Frame Estimates.

Further bias analyses included comparing sample and frame estimates for characteristics identified above where the weighted unit response rate fell below 75 percent on a specific questionnaire. In each case where a significant difference between the sample and frame estimate was found, further analysis was conducted in order to understand the source of the difference. Analysis included examining the components of the sample (i.e. respondents, nonrespondents, out-of-scopes) and the corresponding weighted percentage distribution. In almost all cases, the difference was found to be due to very small cell sizes or a disproportionate level of out-of-scopes.

---

[1] There were sufficient numbers of charter schools to calculate response rates for the states of Arizona, California, Colorado, Florida, Michigan, and Texas. All other states were collapsed into the "others" category.

Table2.—Private School sector questionnaires and Private School Teacher Listing Form: Weighted unit response rates below 75 percent by affiliation and NCES typology

Characteristic		Private School Principal	Private School	Private School Teacher Listing Form	Private School Teacher	Private School Teacher (overall)	Private School Library Media Center
Affiliation	Catholic	†	†	†	†	73.99	†
	Episcopal	†	†	†	†	74.80	†
	Friends	†	74.67	†	†	66.35	†
	Seventh-day Adventist	†	†	†	†	†	†
	Hebrew Day	†	70.69	70.25	59.84	42.04	73.40
	Solomon Schechter	†	†	†	†	64.29	†
	Other Jewish	73.74	73.42	71.00	55.89	39.68	74.65
	Lutheran, Missouri Synod	†	†	†	†	†	†
	Lutheran, Wisconsin Synod	†	†	†	†	†	†
	Evangelical Lutheran	†	†	†	†	74.25	†
	Other Lutheran	†	†	†	†	67.79	†
	Christian School International	†	†	†	†	†	†
	Amer. Assoc. of Christian Sch.	71.30	66.57	65.38	69.79	45.63	†
	Assoc. of Christian Sch. Intl.	†	†	†	†	63.03	†
	Montessori	†	†	†	68.41	59.65	†
	Natl. Assoc. Pr. Sch. For Exp.	†	†	†	†	74.48	†
	Independent Schools	†	†	†	†	71.13	†
	National Ind. Priv. Sch. Assoc.	†	†	†	†	60.47	†
	Others	74.95	70.59	†	†	58.34	†
NCES typology	Catholic	†	†	†	†	74.61	†
	Parochial	†	†	†	†	†	†
	Diocesan	†	†	†	†	74.97	†
	Private	†	†	†	†	71.75	†
	Other religious	†	†	†	†	66.84	†
	Conservative Christian	†	†	†	†	61.04	†
	Affiliated	†	†	†	†	72.09	†
	Unaffiliated	†	†	†	†	67.77	†
	Nonsectarian	†	†	†	†	67.99	†
	Regular	74.32	71.55	†	†	66.53	†
	Special emphasis	†	†	†	66.26	58.51	†
	Special education	†	†	†	†	†	†

† Not applicable

SOURCE: U.S. Department of Education, National Center for Education Statistics, Schools and Staffing Survey (SASS), 1999–2000, "Private School Principal Questionnaire," "Private School Questionnaire," "Teacher Listing Form for Private Schools." "Private School Teacher Questionnaire" and "Private School Library Media Center."

4. Item Nonresponse Rates

Item response rates (RRI) were calculated as the ratio of the number of respondents for whom an in-scope response was obtained (I^x for item x) to the number of respondents who are asked to answer that item. The number asked to answer an item is the number of unit level respondents (I) minus the number of respondents with a valid skip for item x (V^x):

$$RRI^x = \frac{I^x}{I - V^x}$$

Items with item response rates less than 75 percent were identified and evaluated for the potential for nonresponse bias.

4.1 Comparison of Sample and Frame Estimates

For each item with "high" nonresponse, the percentage of respondents by their key characteristics were calculated and compared to the percentage in the population with these same key characteristics. Significant differences in these percentages were identified.

4.2 Estimation of Item Nonresponse Bias

As a last step in this nonresponse bias analysis, an estimate of the magnitude of the relative bias due to item nonresponse was developed for some of the variables with item response rates below 75 percent. The following describes the methodology used in developing these estimates.

Let A and B be two similar items from the same questionnaire. Item A has a low response rate (below 75%) while item B has a high response rate (greater than 75%) and is highly correlated with item A. We make the assumption that if the nonrespondents of item A had also been nonrespondents of item B then the nature and the direction of the bias in item B due to such hypothetical nonresponse would be the same as the actual nonresponse bias in item A. Under this assumption we calculate an estimate of item B based only on the respondents of item A. We will call the difference between the actual estimate of item B derived from all respondents to this item and the estimate based only on the cases who have responded to item A as the "bias in item B *induced* by the nonrespondents of item A".

We divide this difference by the actual estimate of item B and call this the "relative bias of item B induced by the nonrespondents of item A". This bias is expressed in percentage form.

We further assume that that the relative bias in item A due to nonresponse is equal or close to equal the relative bias in item B induced by the nonrespondents of item A.

To test the validity of our assumptions, we examine the following two estimates of item A.
1. Estimate based on the respondents without any imputation.
2. Estimate based on the respondents augmented by the usual imputation of missing values.

We claim that if the imputation is adequately performed then the difference between the above two estimates reflects the magnitude of the bias due to nonresponse in item A. When we divide the above difference by the estimate based on the imputed item; the ratio (expressed in percent) would be an estimate of the magnitude of the relative bias of item A. Thus, we have developed two estimates of the relative bias of A. When our assumptions are sound, the two estimates of the relative bias of item A due to nonresponse should be of the same order of magnitude.

In order to test our methodology and assumptions we selected a set of A items from each of the SASS questionnaires with response rates below 75 percent. We also selected a highly correlated B item for each selected A item. For example, Table 3 below list a set of A items that

we selected from the District Questionnaires with less than 75 percent response. In addition, the table lists the highly correlated item B selected for each of these A items.

Table 3: District Questionnaire Items

Q6. Around the first of October, how many students in grades K-12 and comparable ungraded levels were:			
A Variables		**B Variables**	
6a.	Hispanic, regardless of race (Mexican, Puerto Rican, Cuban, Central or South American, or other Hispanic culture or origin)?	6f.	Total K-12 students
6b.	White, not of Hispanic origin?	6f.	Total K-12 students
6c.	Black, not of Hispanic origin?	6f.	Total K-12 students

Table 4 presents the calculations that lead to developing the two estimates of the relative bias for all A items selected in this study (shaded columns).

We found that the order of magnitude of the two estimates of the relative bias of A in all cases were the same. Thus, we have shown that our assumptions were valid and we conclude that the final estimates of item A (i.e., calculated after imputations) are largely free from any nonresponse bias. Note, however, the magnitude of these estimates of relative bias correspond to the magnitude of nonresponse bias in the unimputed items.

Table 4: Estimates of Relative Bias for Selected Items

Item A	Imputed Estimate of Item A	Unimputed Estimate of Item A	Bias in Item A	First Estimate of Relative bias of Item A	Item B	Imputed Estimate of Item B	Estimate of Item B induced by unimputed Item A	Bias in Item B induced by Item A	Relative Bias of Item B (induced by Item A) =Second Estimate of Relative Bias of Item A	Correlation of Unimputed Item B and Imputed Item A
				District Questionnaire						
6a	499.3	385.4	113.9	22.81%	6f	3151.9	2600.0	551.9	17.51%	0.84
6b	1966.7	1730.8	236.0	12.00%	6f	3151.9	2646.2	505.8	16.05%	0.71
6c	526.6	371.4	155.3	29.48%	6f	3151.9	2601.0	550.9	17.48%	0.71
				Private School Questionnaire						
8b	148.9	114.3	34.6	23.25%	24b	14.5	11.9	2.7	18.33%	0.84
8f	193.3	154.5	38.9	20.11%	24f	16.5	13.7	2.8	16.97%	0.85
52c	22472.6	22671.4	-198.8	-0.88%	52a	20302.1	20537.8	-235.6	-1.16%	0.90
52d	23176.8	23592.4	-415.6	-1.79%	52a	20302.1	20626.1	-323.9	-1.60%	0.86
52e	31303.2	31868.4	-565.2	-1.81%	52b	25358.7	25727.5	-368.8	-1.45%	0.89
52f	34347.6	34729.8	-382.2	-1.11%	52b	25358.7	25542.2	-183.5	-0.72%	0.81
				Public Charter School Questionnaire						
10b	122.5	109.4	13.2	10.75%	10f	264.1	241.7	22.4	8.47%	0.70
46b	12.7	11.0	1.7	13.27%	46f	17.3	15.8	1.5	8.73%	0.82
69d	31190.8	31441.6	-250.7	-0.80%	69c	30083.0	30318.8	-235.8	-0.78%	0.91
69e	41881.3	42116.8	-235.6	-0.56%	69b	34263.5	34543.7	-280.2	-0.82%	0.83