DOCUMENT RESUME

ED 360 380 TM 020 384

AUTHOR Kaufman, Steven; Huang, Hertz
TITLE Schools and Staffing Survey, 1990-91: Sample Design and Estimation. Technical Report.
INSTITUTION National Center for Education Statistics (ED), Washington, DC.
REPORT NO NCES-93-449
PUB DATE Jul 93
NOTE 137p.
PUB TYPE Statistical Data (110) -- Reports - Descriptive (141)

EDRS PRICE MF01/PC06 Plus Postage.
DESCRIPTORS Administrator Characteristics; Data Collection; Elementary School Teachers; Elementary Secondary Education; *Estimation (Mathematics); *National Surveys; Private Schools; Public Schools; Research Design; *Research Methodology; Sample Size; *Sampling; School Statistics; *School Surveys; Secondary School Teachers; Teacher Characteristics; Teacher Supply and Demand
IDENTIFIERS Common Core of Data Program; *Schools and Staffing Survey (NCES)

ABSTRACT

The Schools and Staffing Survey (SASS) represents the union of three surveys by the National Center for Education Statistics (NCES), the Teacher Demand and Shortage Survey, the School and School Administrator Surveys, and the Teacher Survey. The SASS measures critical aspects of teaching supply and demand, the composition of the teacher and administrator workforce, and the status of teaching and schooling in general. The merger of the studies produces one database. SASS is designed to provide estimates to meet its analytical goals: (1) national estimates for public and private schools; (2) state estimates for public schools; (3) state/elementary, state/secondary, and national combined public school estimates; (4) detailed association estimates and grade school estimates for private schools; (5) estimates of change from 1988 to 1991 in school level characteristics; and (6) national estimates for schools with greater than 25% American Indian enrollment. This report describes the procedures used in the following areas: (1) school and teacher sample stratum allocation; (2) overlapping 1988 and 1991 SASS samples; (3) public school sample design; (4) local education agency sample design; (5) private school sample design; (6) teacher sample design (including within school teacher allocation); (7) weighting; (8) imputation; (9) variance estimation techniques; and (10) frame evaluation. Changes in the study design since the inception of the study in 1987-88 are reviewed. Twenty figures and 24 tables illustrate the study design and changes. Four appendixes describe the Common Core of Data; and explain sample reallocation, school overlap, and the effect of a population correction. (SLD)

NATIONAL CENTER FOR EDUCATION STATISTICS

Technical Report

July 1993

1990–91 Schools and Staffing Survey: Sample Design and Estimation

U.S. Department of Education
Office of Educational Research and Improvement

NCES 93-449

NATIONAL CENTER FOR EDUCATION STATISTICS

Technical Report July 1993

1990–91 Schools and Staffing Survey: Sample Design and Estimation

Steven Kaufman
National Center for Education Statistics

Hertz Huang
Bureau of the Census

U.S. Department of Education
Office of Educational Research and Improvement NCES 93-449

U.S. Department of Education
Richard W. Riley
Secretary

Office of Educational Research and Improvement
Emerson J. Elliott
Acting Assistant Secretary

National Center for Education Statistics
Emerson J. Elliott
Commissioner

National Center for Education Statistics

"The purpose of the Center shall be to collect, analyze, and disseminate statistics and other data related to education in the United States and in other nations."—Section 406(b) of the General Education Provisions Act, as amended (20 U.S.C. 1221e–1).

July 1993

Contact:
Steven Kaufman
(202) 219–1337

TABLE OF CONTENTS

APPENDICES

LIST OF TABLES

LIST OF FIGURES

1. Introduction

1.1 General Goals

The Schools and Staffing Survey (SASS) is the integration by the National Center for Education Statistics (NCES) of three of its Elementary and Secondary Education Surveys. These SASS components are: "Teacher Demand and Shortage Survey", the "School and School Administrator Surveys", and the "Teacher Survey". The survey was designed to measure the critical aspects of teaching supply and demand, the composition of the administrator and teaching work force, and the status of teaching and schooling generally. The merger produces one database that can provide linkage of data between the LEAs (local education agencies), schools, and teachers. In addition, these three surveys have the same reference period in SASS. To accomplish this:

1) Schools were selected first. Each selected school received a school questionnaire and an administrator questionnaire.

2) A sample of teachers was selected within each selected school. Each selected teacher received a teacher questionnaire.

3) For public schools, the LEAs associated with the selected schools received a Teacher Demand and Shortage (TDS) questionnaire. An additional sample of districts not associated with schools was selected and received the TDS questionnaire. The school questionnaire for the selected private schools included TDS questions for the school.

The SASS was first conducted by the Bureau of the Census during the 1987-88 school year. This report documents the second SASS collection. It was conducted during the 1990-1991 school year, and is referred to as 1991 SASS in this document. Some 12,958 schools and administrators; and 65,217 teachers were selected. 5,424 local education agencies associated with the selected schools and 135 districts not associated with schools were selected in 1991 SASS.

The SASS is designed to provide the following estimates to meet its analytical goals:

1) National estimates for public and private schools;

2) state estimates for public schools;

3) state/elementary, state/secondary, and national combined public school estimates (see section 5.1.3 for the definition of elementary, secondary and combined schools);

4) detailed association estimates and grade level estimates for private schools;

5) estimates of change from 1988 to 1991 in school level characteristics;

6) national estimates for schools with greater than 25% Indian enrollment.

This report describes the procedures used for the: 1) school and teacher sample stratum allocation, 2) overlapping 1988 and 1991 SASS samples, 3) public school sample design, 4) LEA sample design, 5) private school sample design, 6) teacher sample design (including within school teacher allocation), 7) weighting, 8) imputation, 9) variance estimation techniques, and 10) frame evaluation.

1.2 Response Rates

Below are the unweighted and weighted questionnaire response rates for the SASS components. The unweighted response rates are defined as the number of in-scope responding questionnaires divided by the number of in-scope sample cases. The weighted response rates are defined the same way, using the weighted instead of unweighted numbers. The overall response rate for a particular item (questionnaire response rate times item response rate) may be lower than the respective response rates given below because the item nonresponse rates are not included in the figures below.

Table 1.--Weighted and Unweighted Questionnaire response rates

Survey Type	Unweighted Response Rate	Weighted Response Rate[1]
Teacher Demand and Shortage (LEA)	93.0	93.5
Public School Administrator	96.9	96.7
Private School Administrator	91.0	90.1
Public School	95.0	95.3
Private School	84.8	83.9
Public Teacher[2]	91.5	91.0
Private Teacher[3]	83.0	83.6

[1]Weighted using the inverse of the probability of selection.

[2]These rates do not include the 5 percent of the public schools that did not provide teacher lists.

[3]These rates do not include the 11 percent of the private schools that did not provide teacher lists.

A future report will examine survey response rates and possible bias in more detail.

Table 2 provides the district and school 1991 unweighted response rates for units being asked to respond to SASS in: 1) both 1988 and 1991 (overlap units), and 2) 1991 only (nonoverlap). See section 4 for more information concerning the selection of overlap schools.

Table 2.-- Unweighted Overlap/Nonoverlap Questionnaire response rates

Survey Type	Overlap Response Rate	Nonoverlap Response Rate
Teacher Demand and Shortage (LEA)	93.3	94.3
Public School	95.0	95.1
Private School	87.1	84.5

1.3 Changes in SASS Design from 1988 to 1991

After the experiences with the first SASS collection, a statistical team was set up to evaluate the 1988 sample design and make appropriate changes. Below is a summary of the changes made in the 1991 sample design.

1) Instead of using the QED (Quality Education Data)[1] file as a public school frame, the Center's CCD (Common Core Data) school file was used. See Appendix 1 for a description of the CCD file. This was done because the definition of a school was different between QED and CCD. See Section 5.1.2 for further discussion of definitional differences. To make SASS school estimates consistent with CCD school estimates, the SASS frame was changed. Additionally, it was felt that the CCD teacher counts were more current.

   To measure the impact of the school definition difference on SASS school estimates, the 1991 survey was designed to produce estimates using either the QED or CCD definition. The default definition is the CCD's.

2) To improve the precision of the 1991 private sector estimates, the 1989 Private School Survey (PSS) was used as a sampling frame. See appendix 1 for more information about PSS. The PSS has 123 area frame PSUs, instead of the 1988 SASS's 75, and updated measures of size for schools on the frame.

[1]The QED (Quality Education Data) file was produced by Quality Education Data, Inc., a company that produces mailing lists of educational institutions.

The 1991 private school stratum definitions were based on the PSS school reports of association membership and affiliation, instead of association membership and affiliation lists, as was done for the 1988 private school stratum definitions.

3) To increase the level of detail available for the public sector, the school sample was reallocated to produce state/elementary and state/secondary estimates. In 1988, public sector estimates were only designed to be reliable at the state level (elementary, secondary and combined schools as one group).

   For the private sector, the sample was reallocated to publish estimates for five additional associations - for a total of 18 associations.

4) To improve the precision of SASS change estimates from 1988 to 1991, 30% of the 1991 public school sample was also in the sample in 1988. For the private school sample, associations with high response rates had a 30% overlap like the public schools. Associations with lower response rates had smaller percent school overlaps. Associations with poor response rates had the school overlap minimized. See Section 4 for further discussion of this issue.

5) In 1988, bilingual and new teachers (teachers with less than three years teaching experience) were oversampled. In 1991, Native Americans, and Asian and Pacific Islanders were oversampled, in addition to bilingual and new teachers.

   In 1991, schools with 25% or more Native American enrollment were oversampled. Also, a large sample of BIA (Bureau of Indian Affairs) schools was included.

6) In 1988, the SASS administrator and teacher files were not imputed for missing data. In 1991, they were imputed. In both 1988 and 1991, the TDS and school files were imputed for item nonresponse.

7) To make the SASS estimated teacher counts from the school and teacher files more consistent, the teacher file weights were adjusted so they equaled the teacher estimate (headcount) from the school file.

The 1991 survey differs from the 1988 survey primarily in that in 1988 the sample of private schools was given separate Teacher Demand and Shortage questionnaires. In 1991, the private school Teacher Demand and Shortage questionnaire was incorporated into the Private School questionnaire and the Teacher Demand and Shortage questionnaire was administered only to public school LEAs.

1.4 Concerns about SASS change estimates from 1988 to 1991

Care must be taken when estimating change from 1988 to 1991 in a SASS data element, because some of the change may be due to changes in the sample design, as opposed to change in the education system (like a 3% drop in enrollment). Below are sample design changes that might affect the measurement of change over time.

1.4.1 Changing the sampling frame from QED to CCD

This is a concern because the definition of a school is different between the two frames. In the 1988 SASS (QED) a school was defined as a physical location, while in the 1991 SASS (CCD) a school was defined as an administrative unit with a principal. In states which have multiple administrative units in a single physical location, the estimated change in the number of schools could increase. This increase is at least partially caused from the definition difference.

It is possible to collapse the 1991 SASS school sample to reflect the QED definition of school as was defined in the 1988 SASS, thereby eliminating this concern. However, resulting estimates may no longer be consistent with CCD estimates.

To the extent that the coverage between CCD and QED are different, then part of the change in school related estimates can be attributable to this coverage difference.

1.4.2 Adjusting the estimated number of teachers from the teacher file to the estimated number of teachers from the school file

This was done to make estimates in the files more consistent. Since this was not done in the 1988 survey, some observed distributional differences between the 1988 and 1991 teacher files may be partially attributable to this adjustment. In the public 1988 SASS files, the teacher counts on the teacher file are smaller than the counts on the school file. In the 1991 SASS files, the teacher file counts are increased to equalize the estimates. This increase does not reflect a change in the educational system, but a bias correction differentially applied between the files.

1.4.3 Imputing for missing data on the administrator and teacher files

All data files are adjusted for complete refusals. However, for the 1988 administrator and teacher files, missing data elements within responding units are not imputed. Hence, estimates of totals implicitly use a value of zero for all missing data elements (i.e., 1988 totals are underestimates whenever there are missing data). The 1991 estimates of totals use imputed values for missing data

elements. Therefore, some of the measured change between the 1988 and 1991 totals is due to imputing one year, but not the other. This part of the change is not due to a change in the educational system.

Change estimates for ratios and averages are also affected by imputations in one year, but not the other. However, the magnitude and direction of the bias is unknown and dependent on the variable of interest. This part of the change is not due to a change in the educational system.

1.4.4 Questionnaire and conceptual differences

Care must be observed in the interpretation of change estimates between 1988 and 1991 since specific questions are not always worded the same from the first SASS survey to the second[2]. Both major and minor changes in wording of specific items occur; the ordering of items may be different and concepts can be different.

As an example, in both the 1987-88 and 1990-91 SASS, the question, "Which best describes the community in which the school is located?" was asked of the principal (for the administrator/principal survey) and the respondent to the school survey. The SASS reinterview program in both 1988 and 1991 determined the responses to this item were highly subjective and exhibited moderate response variance. As a result of this finding, the 1991 SASS micro-data files contain an "urbanicity" code (Locale) developed by Johnson (1989)[3]. This code is believed to be a more accurate description of the community than the self-reports on SASS. This methodology assigns "type of locale codes" based on the school mailing address matched to Bureau of the Census data files containing population density data, Standard Metropolitan Statistical Area (SMSA) codes, and a Census code defining urban and rural areas.

This rigorously defined locale code on the 1991 SASS files may be different from the self-report of community type.

---

[2]See SASS: 1990-91 Data File User's Manual, Volume I Appendix B, f a crosswalk of the changes between the 1988 and 1991 questionnaires.

[3]Johnson, F. (1989), Assigning Type of Locale Codes to the 1987-88 CCD Public School Universe, National Center for Education Statistics Technical Report, Data Series: SP-CCD-87188-7.4, CS 89-194.

2. Defining the Universe for the 1991 SASS: Scope of 1991 SASS

In the 1991 SASS, CCD was used as a sampling frame for public schools. CCD defined a school differently than QED which was used as a frame in 1988 SASS. The following terms define the scope of the components of the 1991 SASS.

2.1 Teacher Demand and Shortage Survey

Local Education Agency (LEA). An LEA, or public school district, is defined by CCD as a government agency administratively responsible for providing public elementary and/or secondary instruction and educational support services. The agency or administrative unit must operate under a public board of education. Districts which do not operate schools but do hire teachers are included.

Out-of-Scope. An LEA was considered out-of-scope for the Teacher Demand and Shortage Survey if it did not employ elementary or secondary teachers of any kind, including special education teachers and itinerant teachers.

2.2 School Survey

Public School. The CCD defines a public school as an institution which provides educational services, has one or more teachers to give instruction, is located in one or more buildings, receives public funds as primary support, has an assigned administrator, and is operated by an education agency. Prison schools, schools operated by Department of the Defense and the Bureau of Indian Affairs are included.

Out-of-Scope. A public CCD school was considered out-of-scope for SASS if it did not have any students in any grades 1-12. Schools offering only kindergarten and pre-kindergarten were deleted from the sampling frame before the sample was selected. If a school was determined to be out-of-scope after editing its questionnaire, it was deleted from the data file.

Private School. A private school is defined by PSS as a school not in the public system that provides instruction for any of grades 1-12 where the instruction was not given exclusively in a private home.

Out-of-Scope. A private school was considered out-of-scope for SASS if it did not have any students in any of grades 1-12, if it operated in a private home that was used as a family residence, or if it was undetermined whether it operated in a private home and its size was very small (enrollment less than 10 or only one teacher). Out-of-scope schools were deleted from the sampling frame before the sample was selected. If a school was determined to be out-of-scope after editing its questionnaire, it was deleted from the data file.

2.3 School Administrator Survey

Administrator. A school administrator questionnaire was sent to the person who is primarily responsible for overseeing the administrative operations and actions of the school.

Out-of-Scope. A school administrator sample case was considered out-of-scope if the school did not have an administrator. Also, if a sample administrator's school is considered out-of-scope, the administrator is automatically classified as out-of-scope.

2.4 Teacher Survey

Teacher. A teacher is defined as any full-time or part-time teacher whose primary assignment is teaching in any of grades K-12. Itinerant teachers are included, as well as long-term substitutes who were filling the role of a regular teacher on an indefinite basis. An itinerant teacher is defined as a teacher who teaches at more than one school.

Out-of-Scope. A sample teacher was considered out-of-scope if he/she is a short-term substitute, a student teacher, a nonteaching specialist (e.g., guidance counselor, librarian, nurse, psychologist), an administrator (e.g., principal, assistant principal), a teacher's aide, or in some other professional or support staff position (cooks, custodian, bus driver, dietician, secretary). If a sample school is out-of-scope, all teachers from that school are also considered out-of-scope.

If an LEA was classified as out-of-scope, its teachers, administrators and schools are also classified as out-of-scope. Likewise if a school is classified as out-of-scope, its teachers and administrators are also considered out-of-scope.

3. School and Teacher Allocation

This section discusses the allocation of the public and private school sample, as well teacher samples. The CCD file was used as the public school frame. The private school sample was based on the list and area frame design from the PSS. See the sections noted below for more information concerning the SASS frames and selection procedures.

3.1 Public School Allocation (See section 5.1)

3.1.1 SASS Public School Sample Goals

The goals for the public school sample of the 1991 SASS were:

1. use the Common Core of Data (CCD)[4] file as a frame instead of the Quality Education Data (QED)[5] file used for 1988 SASS,

2. produce state estimates,

3. produce state/elementary and state/secondary estimates (See Appendix 2 for a graphical explanation of how this allocation rule is expected to change the precision of the state-level estimates.),

4. produce national combined school estimates,

5. produce overall national estimates by sector,

6. overlap a certain percentage of the 1991 SASS school sample with the 1988 school sample to improve 1988 to 1991 change estimates, over what they would be without overlap,

7. measure the effect of the differences in the QED and CCD school definitions on the 1991 SASS data, and

8. oversample schools with greater than 25% Native American enrollment, so that national estimates can be produced.

---

[4]The Common Core of Data is a file of all public schools and school districts compiled by the National Center for Education Statistics from data supplied by all state agencies.

[5]The Quality Education Data is a file of public and private schools and public school districts that was purchased from Quality Education Data, Inc.

3.1.2 Allocation Methodology

The 1988 SASS, produced the following observations:

1. state estimates from the states with smaller populations had higher than expected standard errors,

2. state estimates from the states with larger populations had lower than expected standard errors,

3. state elementary and state secondary estimates could not be made except for the largest states, and

4. the overall national estimates had much lower than expected standard errors.

To improve the analytical capabilities for state estimates in the 1991 SASS, the sample was reallocated so that elementary and secondary state estimates for schools could be made. Since the states with the largest population had considerable sample in 1988 SASS, one way of achieving this goal was to reduce sample in the largest states and reallocate it to the smallest states. This process would reduce the reliability of the overall national estimates, but since the absolute reliability of these estimates was high, the ability to provide reliable national estimates would not be reduced appreciably.

The basic approach for the allocation was:

1. Assign a minimum number of schools to each stratum (state/level). For the combined school strata, the minimum was 10. For elementary/secondary strata the school minimum was 80. (With eighty schools in a stratum most elementary/secondary strata CVs should be 15% or less.) Since Alaska has many combined schools, the minimum sample size was set at 70 rather than 10.

2. Reduce the state collection burden. No stratum should have a sample size larger than 40% of the total number of schools in the stratum.

3. Use a total public school sample size in the 1991 SASS of 9330, as it was in the 1988 SASS.

4. Allocate 1500 schools proportional to the 1988 SASS unit standard errors for the state/combined school strata to achieve maximum precision for national combined school estimates.

5. Allocate the remainder of the school sample proportional to the 1988 SASS unit standard errors for the state/elementary and state/secondary school strata.

The allocation process described above could be done using any SASS variable. Total teachers, total enrollment and total number of schools were used to do separate allocations. Because the primary objective in SASS is to estimate teacher characteristics and because the allocations based on enrollment and school estimates produced similar allocations to the one based on teacher estimates, the teacher allocation was used as the final allocation.

3.1.3 Allocation Results

Table 5 provides the final stratum allocation of the 1991 SASS public school sample.

To control the school overlap, the 1988 and 1991 frames were matched. While matching, the QED school definition was used to define the school. The reasons for using the QED definition are explained in Section 5.1.2. The sample allocation in Table 5 was based on the CCD school definition. As described in section 5.1.2, each selected school may require collection of multiple CCD schools for each selected school. To reduce this overcollection, the designated sample sizes were reduced in some states. The results of this sample reduction did not work well - 209 more schools required collection than designated in the original allocation.

3.1.4 Oversampling of Schools with More Than 25% Native American Enrollment

To improve Native American school estimates, schools with American Indian/Aleut/Eskimo student populations greater than or equal to 25% (Native American strata) were placed into their own strata. Arizona, North Dakota and Oklahoma had individual Native American strata. The rest of the states were placed into an "all other states" Native American stratum. The Native American strata were also stratified by school level. These strata were allocated 251 schools proportional to the number of schools in the stratum. The sample sizes are provided below. Since most Alaskan schools have at least 25% Native American students, they were not included in this stratification.

Table 3.--American Indian/Aleut/Eskimo Stratum Sample Sizes by State and School Level

State	Total	Elementary	Secondary	Combined
Total	251	162	81	8
Arizona	33	22	10	1
North Dakota	15	8	6	1
Oklahoma	117	77	40	-
All Other	86	55	25	6

- means there were no schools in the frame

3.1.5 Selection of Bureau of Indian Affairs (BIA) Schools

The BIA schools were generally not listed on the CCD so they had to be selected from the universe of BIA schools. BIA schools were consequently placed in their own strata. Arizona, New Mexico, and South Dakota had individual BIA strata. The rest of the states were placed into "all other" BIA strata. School level was also used to define the BIA strata. These strata were allocated 101 schools proportional to the number of schools in the stratum. The sample sizes are provided below.

Table 4.--Bureau of Indian Affairs School Stratum Sample Sizes by State and School Level

State	Total	Elementary	Secondary	Combined
Total	101	67	14	20
Arizona	32	26	3	3
New Mexico	26	21	3	2
South Dakota	16	8	2	6
All Other	27	12	6	9

Table 5a.--Public School Stratum Sample Sizes by State and School Level

State	Combined	Elementary	Secondary	Total
Total United States	1508	4203	3625	9336
Alabama	70	80	80	230
Alaska	70	66	25	161
Arizona	10	80	80	170
Arkansas	8	76	76	160
California	132	127	111	370
Colorado	19	80	80	179
Connecticut	10	80	80	170
Delaware	8	44	20	72
District of Columbia	10	48	16	74
Florida	76	121	80	277
Georgia	32	80	80	192
Hawaii	5	66	21	92
Idaho	10	80	68	158
Illinois	80	81	80	241
Indiana	40	80	80	200
Iowa	14	80	80	174
Kansas	1	80	80	161
Kentucky	20	80	80	180
Louisiana	69	80	80	229
Maine	10	80	55	145
Maryland	18	80	80	178
Massachusetts	15	80	80	175
Michigan	51	80	80	211
Minnesota	29	80	80	189
Mississippi	75	80	80	235
Missouri	42	80	80	202
Montana	3	80	80	163

Table 5a--Public School Stratum Sample Sizes by State and School Level (Continued)

State	Combined	Elementary	Secondary	Total
Nebraska	4	80	80	164
Nevada	4	80	30	114
New Hampshire	4	80	34	118
New Jersey	23	95	80	198
New Mexico	2	80	68	150
New York	98	85	80	263
North Carolina	10	80	80	170
North Dakota	20	80	80	180
Ohio	43	80	80	203
Oklahoma	5	80	80	165
Oregon	8	80	80	168
Pennsylvania	77	94	80	251
Rhode Island	4	80	24	108
South Carolina	26	80	80	186
South Dakota	7	80	80	167
Tennessee	45	80	80	205
Texas	105	180	129	414
Utah	14	80	80	174
Vermont	6	80	19	105
Virginia	42	80	80	202
Washington	16	80	80	176
West Virginia	10	80	80	170
Wisconsin	6	80	80	166
Wyoming	2	80	49	131

Table 5b.--Proportion of Public School Frame Selected in SASS Sample by State

State	Sample Size	Percent of Frame in Sample
Total United States	9336	11.3%
Alabama	230	17.8%
Alaska	161	35.6%
Arizona	170	16.7%
Arkansas	160	15.1%
California	370	5.1%
Colorado	179	13.4%
Connecticut	170	17.7%
Delaware	72	43.6%
District of Columbia	74	40.0%
Florida	277	11.5%
Georgia	192	11.1%
Hawaii	92	40.0%
Idaho	158	28.3%
Illinois	241	5.9%
Indiana	200	10.4%
Iowa	174	10.8%
Kansas	161	11.0%
Kentucky	180	12.9%
Louisiana	229	14.6%
Maine	145	19.5%
Maryland	178	15.0%
Massachusetts	175	9.8%
Michigan	211	6.5%
Minnesota	189	12.2%
Mississippi	235	24.7%
Missouri	202	9.5%
Montana	163	21.2%

Table 5b.--Proportion of Public School Frame Selected in SASS Sample by State (Continued)

State	Sample Size	Percent of Frame in Sample
Nebraska	164	10.9%
Nevada	114	36.2%
New Hampshire	118	27.2%
New Jersey	198	8.8%
New Mexico	150	23.1%
New York	263	6.7%
North Carolina	170	8.7%
North Dakota	180	26.5%
Ohio	203	5.4%
Oklahoma	165	9.1%
Oregon	168	13.9%
Pennsylvania	251	7.6%
Rhode Island	108	36.0%
South Carolina	186	16.9%
South Dakota	167	21.1%
Tennessee	205	13.2%
Texas	414	7.1%
Utah	174	23.9%
Vermont	105	31.8%
Virginia	202	11.4%
Washington	176	9.5%
West Virginia	170	16.0%
Wisconsin	166	8.4%
Wyoming	131	31.8%

Table 5c.--Proportion of Public School Frame Selected in SASS Sample by School Level

School Level	Sample Size	Percent of Frame in Sample
Total	9336	11.3%
Combined	1508	32.1%
Elementary	4203	7.4%
Secondary	3625	17.6%

3.2 Private School Allocation for the List Sample (See section 5.3)

The goals for the 1991 SASS private school allocation for the most part remained the same as the 1988 goals.

1. produce detailed Private School Association group estimates
2. produce national private sector estimates
3. produce national private sector school level estimates
4. produce estimates for national public vs private sector comparisons

The following changes were made to the 1988 goals for the private school sample:

1. Instead of using State as a stratification variable, census region was used. This was done to improve the efficiency of the allocation and to allow for more private school association strata. The four census regions are defined below:

   a) Northeast consists of the states of Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont.

   b) Midwest consists of the states of Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin.

   c) South consists of the states of Alabama, Arkansas, Delaware, District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, and West Virginia.

d) West consists of the states of Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming.

2. Increase the number of association strata from 13 to 18. The 1988 Lutheran stratum was split into 4 strata-Lutheran, Missouri Synod; Lutheran, Wisconsin Synod; Evangelical Lutheran Church in America; and Other Lutheran; and the Jewish stratum was split into three strata - National Society of Hebrew Day Schools; Solomon Schechter; and Other Jewish.

3. Identify school association membership from the 1989 Private School Survey (PSS)[6] data, instead of the 1988 association lists.

4. Identify school level using the PSS data, instead of the QED data.

The allocation procedure used the 1988 estimated number of teachers from the SASS teacher file to proportionally allocate the sample. The file was stratified by association/level/region. First, the 1988 SASS variance estimates for numbers of teachers were used to determine that approximately 100 schools per association should provide association strata CVs of less than 15%. Next, the proportional allocation was determined without regard to minimum stratum sample size. Any association with less than 100 schools from this proportional allocation was targeted for 100 schools. The allocation was then redone within each of the targeted associations as well as the rest of the associations.

The private school sample size selected from the list frame was 2670 schools. In addition, 600 schools were selected from the area frame to make up for coverage deficiencies in these lists. See section 5.3.3 for more detailed discussion of the Private School Area Frame.

Table 6a provides the allocation for the list frame. The table includes allocations for the association/level/region strata, as well as for marginal aggregate groupings. Table 6b shows the allocation by association/level, as well as the marginal aggregate groupings.

[6]The Private Schools Survey (PSS) is a file of all private schools obtained from QED, list frame updates, and area search frame updates. See Appendix 1 for a description of the PSS file.

Table 6a.--Private School Stratum Sample Sizes by Association, Region and School Level

	North East				North Central			
Association	Elementary	Secondary	Combined	Total	Elementary	Secondary	Combined	Total
Total	348	152	259	759	458	125	148	731
Catholic	145	69	13	227	182	72	10	264
Friends	18	3	19	40	2	2	1	5
Episcopal	3	6	9	18	2	0	2	4
National Hebrew Day	36	18	12	66	2	2	9	13
Solomon Schechter	28	0	2	30	7	0	1	8
Other Jewish	25	6	32	63	4	4	3	11
Lutheran - Missouri Synod	5	2	2	9	38	5	2	45
Lutheran - Wisconsin Synod	1	0	0	1	75	5	2	82
Evangelical Lutheran Church in America	17	0	2	19	24	1	2	27
Other Lutheran	6	0	2	8	41	3	8	52
Seventh Day Adventist	10	3	10	23	10	5	10	25
Christian Schools International	9	3	4	16	22	6	13	41
American Association of Christian Schools	7	0	10	17	9	1	11	21
National Association of Private Schools for Exceptional Children	1	1	46	48	0	0	5	5
Military	0	2	0	2	0	3	2	5
Montessori	8	0	3	11	17	0	20	37
National Association of Independent Schools	10	13	49	72	10	6	21	37
Other	19	26	44	89	13	10	26	49

Table 6a.--Private School Stratum Sample Sizes by Association, Region and School Level (Continued)

	South				West			
Association	Elementary	Secondary	Combined	Total	Elementary	Secondary	Combined	Total
Total	302	71	337	710	247	75	148	470
Catholic	91	33	12	136	53	26	10	89
Friends	4	2	8	14	5	2	1	8
Episcopal	26	4	27	57	6	2	11	19
National Hebrew Day	10	1	4	15	2	3	2	7
Solomon Schechter	10	0	1	11	4	1	1	6
Other Jewish	4	2	11	17	3	2	4	9
Lutheran - Missouri Synod	14	2	2	18	24	2	2	28
Lutheran - Wisconsin Synod	4	0	1	5	8	2	2	12
Evangelical Lutheran Church in America	11	0	0	11	36	0	0	36
Other Lutheran	20	1	2	23	12	1	2	15
Seventh Day Adventist	10	5	10	25	10	10	10	30
Christian Schools International	13	0	9	22	10	2	10	22
American Association of Christian Schools	10	2	34	46	4	1	10	15
National Association of Private Schools for Exceptional Children	2	0	27	29	2	0	16	18
Military	3	0	9	12	0	1	0	1
Montessori	19	0	16	35	18	0	2	20
National Association of Independent Schools	10	9	45	64	10	10	25	45
Other	41	10	119	170	40	10	40	90

Table 6b.--Private School Sample Sizes by Association and School Level

Association	Elementary	Secondary	Combined	Total
Total	1355	423	892	2670
Catholic	471	200	45	716
Friends	29	9	29	67
Episcopal	37	12	49	98
National Hebrew Day	50	24	27	101
Solomon Schechter	49	1	5	55
Other Jewish	36	14	50	100
Lutheran - Missouri Synod	81	11	8	100
Lutheran - Wisconsin Synod	88	7	5	100
Evangelical Lutheran Church in America	88	1	4	93
Other Lutheran	79	5	14	98
Seventh Day Adventist	40	23	40	103
Christian Schools International	54	11	36	101
American Association of Christian Schools	30	4	65	99
National Association of Private Schools for Exceptional Children	5	1	94	100
Military	3	6	11	20
Montessori	62	0	41	103
National Association of Independent Schools	40	38	140	218
Other	113	56	229	398

Table 6c.--Proportion of Private School Frame Selected in SASS Sample by Association

Association	Sample Size	Percent of Frame in Sample
Total	2670	13.1%
Catholic	716	8.3%
Friends	67	100.0%
Episcopal	98	32.9%
National Hebrew Day	101	67.6%
Solomon Schechter	55	100.0%
Other Jewish	100	34.7%
Lutheran - Missouri Synod	100	10.2%
Lutheran - Wisconsin Synod	100	35.0%
Evangelical Lutheran Church in America	93	100.0%
Other Lutheran	98	51.0%
Seventh Day Adventist	103	9.0%
Christian Schools International	101	34.6%
American Association of Christian Schools	99	11.2%
National Association of Private Schools for Exceptional Children	100	54.0%
Military	20	100.0%
Montessori	103	24.6%
National Association of Independent Schools	218	25.6%
Other	398	7.2%

Table 6d.--Proportion of Private School Frame Selected in SASS Sample by School Level

School Level	Sample Size	Percent of Frame in Sample
Total	2670	13.1%
Combined	892	15.0%
Elementary	1355	10.8%
Secondary	423	22.7%

Table 6e.--Proportion of Private School Frame Selected in SASS Sample by Census Region

Census Region	Sample Size	Percent of Frame in Sample
Total	2670	13.1%
Northeast	759	14.7%
North Central	731	11.8%
South	710	13.6%
West	470	12.3%

3.3 Private School Allocation for the Area Sample (See section 5.3)

The area sample is designed to represent the private schools missing from the list frame. A search for schools missing from the list frame is made within 123 selected counties (area frame). 600 out of almost 900 total schools in the area frame were allocated to the area sample. This number is large because the area frame schools represent 21 percent of the total private schools. All schools in the noncertainty areas (453 schools) were selected with certainty. The remaining 147 sample cases were selected in the certainty areas. They were allocated proportional to the number of schools in the PSU/level strata.

3.4 Teacher Allocation (See section 6)

The teacher sample was allocated among five strata (or teacher types). The strata are as follows: 1) American Indian, Aleut or Eskimo; 2) Asian or Pacific Islander; 3) Bilingual/ESL; 4) New; and 5) Experienced. The approximate allocation was 1,500 Asian or Pacific Islander teachers, 1,500 American Indian, Aleut, or Eskimo teachers,

and 2,000 Bilingual teachers. The remaining 60,088 sample teachers were allocated among new and experienced teachers. New teachers in the private sector had an oversampling factor applied so that a large enough sample was selected to ensure the ability to analyze new private school teachers for both the SASS and the Teacher Follow-up Survey (TFS) surveys.

3.4.1 General Remarks

The allocated sample size often differed from the number of sample cases selected. The reason for this is because the school's conditional probability of selection was used instead of its measure of size when selecting the sample.

3.4.2 Within School Teacher Allocation

For new/experienced teachers in public schools, oversampling was not required due to the large number of sample schools with new teachers. Therefore, teachers were allocated to the new and experienced categories proportional to their numbers in the school. However, for private teachers, new teachers were oversampled to ensure that there would be enough new teachers in both 1991 SASS and the Teacher Follow-up Survey (TFS).

Before teachers were allocated to the new/experienced strata, schools were first allocated an overall number of teachers to be selected. This overall sample size was chosen so as to equalize the teacher weights with the school stratification (state/level and association/level/region for public and private sectors, respectively), assuming no further teacher stratification and zero Asian Pacific Islander (API), American Indian, Aleut and Eskimo (AIAE), and Bilingual teachers.

Table 7 provides the average number of new and experienced teachers to be selected within each public and private school by school level.

Table 7.--Average number of new and experienced teachers selected per school by school level and sector

	School Level		
	Elementary	Secondary	Combined
Public Schools	3.49	6.98	5.23
Private Schools	3.78	4.72	2.83

Given the numbers in Table 7, the new/experienced teacher sample size was chosen to equalize the teacher weights within a school stratum. Since the school sample was selected proportional to the square root of the number of teachers in the school, an equally weighted teacher sample within a school stratum was obtained by selecting t_i new or experienced teachers in school i.

$$t_i = W_i * T_i(C/Y)$$

where: W_i is the school weight for school i (the inverse of the school selection probability).

T_i is the number of new and experienced teachers in school i, as reported on the teacher listing form.

C is the average number of teachers selected per school (See Table 7).

Y is the simple average of the school's weighted measure of size over all schools in the school stratum.

Where, for noncertainty schools, the weighted measure of size equals the school sampling interval times the square root of the number of teachers in the school. The measure of size for certainty schools is the square root of the CCD number of teachers in the school.

The maximum number of new/experienced teachers per school was set at twice the average number of teachers selected per school from Table 7. At least one teacher was selected in each school.

Given the allocation of teachers, t_i, teachers were allocated to the new/experienced strata, t_{ni} and t_{ei}, respectively, in the following manner.

$$t_{ni} = (A * T_{ni} * t_i)/(T_{ei} + A * T_{ni})$$

and

$$t_{ei} = (T_{ei} * t_i)/(T_{ei} + A * T_{ni})$$

where: A is the oversampling factor (A = 1.0 for public teachers and A = 1.8 for private teachers).

T_{ni} is the number of new teachers in school i.

T_{ei} is the number of experienced teachers in school i.

The Asian Pacific Islander (API), American Indian, Aleut, Eskimo (AIAE), and Bilingual teachers were allocated in the following manner:

$$t_{pi} = (W_i * T_{pi})/R$$

$$t_{ai} = (W_i * T_{ai})/H$$

$$t_{bi} = (W_i * T_{bi})/Q$$

where:

T_{pi} is the number of API teachers in school i.

T_{ai} is the number of AIAE teachers in school i.

T_{bi} is the number of bilingual teachers in school i.

R is the national sampling interval to ensure that at least 1500. API teachers are selected nationwide (R = 11).

H is the national sampling interval to ensure that at least 1500 AIAE teachers are selected nationwide (H = 13).

Q is the national sampling interval to ensure that at least 2000 bilingual teachers are selected nationwide (Q = 16).

To make sure a school was not overburdened, the maximum number of teachers per school was set at 20. When the number of sample teachers exceeded 20 in a school, the API, AIAE, and bilingual teachers were proportionally reduced to meet the maximum requirement.

Table 8 provides the number of teachers selected from the selection process described above. The designated number of teachers may differ from the actual number selected for the following reasons:

1) Native American, Asian/Pacific Islander and Bilingual/ESL sampling rates were approximations, so the exact sample sizes were also approximations.

2) The within school teacher allocations were determined using school teacher estimates from the frame. To the extent that the actual teacher counts differed from the estimates, the actual number selected might be higher or lower than expected.

Table 8.--Number of Selected Teachers in SASS Sample by Teacher Type and Sector

Teacher Type	Public	Private	Total
Native American	1,259	270	1,529
Asian/Pacific Islander	1,475	36	1,511
Bilingual/ESL	1,957	164	2,121
New	5,970	2,002	7,972
Experienced	45,390	6,694	52,084
Total	56,051	9,166	65,217

4. Overlapping the 1988 and 1991 SASS School Samples

4.1 Public Schools

One of the goals for the 1991 SASS was to measure change between 1988 and 1991 for various characteristics. To improve such estimates, the sample selection process controlled the amount of overlap between the 1988 and 1991 school samples. Appendix 3 describes how this was done.

The 1991 SASS pretest measured the impact of collecting data from the same school several times. For public schools, the effect on response rates was minimal - 92% for nonoverlap schools and 87% for overlap schools. (To account for overlap schools being selected only from 1988 SASS respondents, overlap pretest sample schools were adjusted for the 1988 SASS nonresponse.) This suggests that the school overlap rate can be high, since the increased precision resulting for estimates of change produces little degradation of response rates.

However, increased overlap for schools implies increased overlap for LEAs. The LEA pretest response rates were 95% for nonoverlap LEAs and significantly less (84%) for overlap LEAs. This seems to indicate some reluctance on the part of the LEAs to participate multiple times.

An estimate for the number of LEAs that would be overlapped from independent samples was 47% (obtained by summing the 1988 selection probabilities for 1988 sampled LEAs). This implies a sizable LEA overlap even if the school overlap isn't increased; thus some reduction in LEA response rates was expected in the 1991 SASS, maybe 5 percentage points. Any control to increase the school overlap would increase the LEA overlap rate and likely reduce the overall LEA response rates even more.

To minimize the impact on the 1991 LEA response rates, the school overlap was set at 30%. With a controlled 30% school overlap, the expected LEA overlap rate was 58%, which from the SASS pretest translates into an expected 6 percentage point drop in response rates if there were no overlap at all. The predicted drop in the LEA response rate did not occur. The simplification in the LEA questionnaire is a contributing factor for the actual increase in response rate.

4.2 Private Schools

From the pretest, we learned that overlapping samples reduces response rates among private schools. Since the overall 1988 SASS private school response rate was only 79%, it is important to minimize the impact overlapping samples will have on the response rate. To do this, we designed a sampling scheme which controlled the expected overlap. This sampling scheme, used in the list frame, provided a 30% overlap for associations with a high 1988 response rate and minimized the overlap for associations with a low response rate.

The 1991 SASS area frame sample was selected independently of the 1988 area frame sample. There was no need to control overlap since half the PSUs selected from the 1991 frame were in the 1988 sample as well. One would assume from this that many schools in these PSUs would be eligible again in 1991. Also, many schools in this frame were in the certainty strata. Taken together, these two designs aspects produced an overlap of 24.5 percent over the entire area sample.

Table 9 shows the anticipated expected overlap for each association for the list frame.

Table 9.--1988 Response Rates and Expected Overlap in Percent for Associations in the List Frame

Association	1988 Response Rate (%) (using unweighted data)	Expected Overlap (%)
Catholic	89.5	30
Friends	84.9	100[7]
Episcopal	83.1	15-20
National Hebrew Day	71.8[8]	minimize overlap
Solomon Schechter	71.8[7]	100[6]
Other Jewish	71.8[7]	minimize overlap
Lutheran - Missouri Synod	89.6[9]	30
Lutheran - Wisconsin Synod	89.6[8]	30
Evangelical Lutheran Church in America	89.6[8]	100[6]
Other Lutheran	89.6[8]	30
Seventh Day Adventist	88.7	30
Christian Schools International	95.8	30
American Association of Christian Schools	54.1	minimize overlap
National Association of Private Schools for Exceptional Children	84.0	20-25
Military	87.5	100[6]
Montessori	78.5	minimize overlap
National Association of Independent Schools	74.2	minimize overlap
Other	70.5	minimize overlap

[7]The overlap is 100% because all schools in the association are in the sample.

[8]These response rates are from the one Jewish stratum that existed in 1988 (i.e., one Jewish stratum split into three in 1991).

[9]These response rates are from the one Lutheran stratum that existed in 1988 (i.e., one Lutheran stratum split into four in 1991).

Table 10 shows the expected and actual overlap sample sizes for each private school affiliation in the list frame.

Table 10.--Private School Expected and Actual Overlap Sample Sizes for Associations in the List Frame

Association	Expected Overlap Sample Size	Actual Overlap Sample Size
Catholic	215	221
Friends	49	49
Episcopal	10	12
National Hebrew Day	6	11
Solomon Schechter	12	12
Other Jewish	2	3
Lutheran - Missouri Synod	30	32
Lutheran - Wisconsin Synod	29	30
Evangelical Lutheran Church in America	9	9
Other Lutheran	27	27
Seventh Day Adventist	31	31
Christian Schools International	26	33
American Association of Christian Schools	0	1
National Association of Private Schools for Exceptional Children	22	26
Military	18	18
Montessori	4	6
National Association of Independent Schools	3	1
Other	3	5
TOTAL	496	527

5. Public School, Private School, and LEA Sample Selection

5.1 Public School Sample

This section describes the frame, stratification, sorting and sample selection. The school allocation is described earlier in the School and Teacher Allocation section (see section 3). In total, 9545 public schools were selected. This is 209 more than originally designated. This difference is due to the fact that: 1) the allocation was based on the CCD school definition, while the sample selection was based on the QED school definition (see section 5.1.2); and 2) the final sample size is random given that the school overlap rate was controlled.

The SASS public school sample was selected so that a maximum of 30% of the schools in the 1988 sample were also in the 1991 sample. See Appendix 3 for a description of that process.

5.1.1 Public School Frames

The primary public school frame for the 1991 SASS was the 1988-89 school year Common Core of Data (CCD) file. The CCD is based on survey data collected annually by NCES from all state education agencies, providing data from their administrative records. NCES and the state education agencies work cooperatively to assure comparability between data elements reported. The CCD is believed to be the most complete public school listing available. The frame includes regular public schools and Department of Defense schools. Nonregular schools such as special education, vocational or technical schools are also included in the sample frame. Before sampling, duplicate schools and schools outside of the United States were removed from the frame. Schools that only teach prekindergarten, kindergarten or adult education were also removed. A total of 83,165 schools remained on the 1988-89 public school frame.

A list of BIA schools was obtained from the Bureau of Indian Affairs. This constituted the other public school sampling frame.

5.1.2 QED and CCD Definitional Differences and Matching Operation

As mentioned in section 4, the overlap between the 1988 and 1991 SASS school sample was controlled. To do this, the 1988 and 1991 frames had to be matched. For public schools matching was complicated by a school definition difference between the frames.

For the 1988 SASS, the QED (sampling frame for the 1988 SASS) and the CCD defined schools differently. The QED defined a school as a physical location, while the CCD defined it as an administrative unit (e.g., a consolidated school district may have a high school meeting in two buildings at two locations

but is administered as one high school. The QED would count this as two schools, the CCD as one. Also, an elementary school and a high school meeting at the same location with two principals would be counted as one school by the QED and as two schools by the CCD.).

To measure the impact this had on 1988 SASS estimates, the following design features were introduced into the 1991 design:

1) When the CCD and QED frames were matched, multiple CCD schools matching in one QED school were identified (many-to-one schools).

2) The many-to-one schools were collapsed together for sample selection.

3) If any of these many-to-one schools were selected for the 1991 sample, the individual CCD schools that comprise the QED school were separated out and collected individually.

The 1991 SASS interviewing collected school data using the CCD definition. However, if we combine data collected from the selected many-to-one schools into one school questionnaire, estimates can be produced using the QED school definition. Looking at the differences between these estimates measures the impact of QED and CCD definitions on the estimates in SASS. This information can be used to adjust 1988 SASS estimates to the CCD school definition. This work has not been started, but specifications have been written to compute the estimates described in this paragraph[10].

5.1.3 Stratification

The first level of stratification was four types of schools: (A) BIA (Bureau of Indian Affairs) schools; (B) Native American schools (schools with 25% or more Native American students); (C) schools in Delaware, Nevada and West Virginia (where it was necessary to implement a different sampling methodology to select at least one school from each LEA in the state - see section 5.2.3); and (D) all other schools (all schools not included in A, B, or C).

The second level of stratification: The type A schools were stratified by Arizona, New Mexico, South Dakota, and all other states. The type B schools were stratified by Arizona, North Dakota, Oklahoma and all other states (except Alaska, since most Alaskan schools have high Native American enrollment). The

[10] 1988 Schools and Staffing Survey Sample Design and Estimation, NCES 91-127, Frame Evaluation, describes the magnitude of definitional difference with respect to numbers of schools.

type C schools were stratified first by state and then by LEA. The type D schools were stratified by state (all states and the District of Columbia except Delaware, Nevada and West Virginia).

Within each second level there were 3 grade level strata (elementary, secondary, and combined schools), defined as follows:

Regular Schools:

Elementary: Lowest grade $\leq$ 6 and Highest grade $\leq$ 8

Secondary: Lowest grade $\geq$ 7 and Highest grade $\leq$ 12

Combined: Lowest grade $\leq$ 6 and Highest grade > 8

Nonregular schools, which include special education, vocational, technical, adult education (if part of an in-scope school) or alternative/continuation grades were classified as combined schools.

5.1.4 School Sorting

To facilitate the calculation of LEA weights, it was important that within a stratum all schools belonging to the same LEA be together. This can be achieved by sorting by LEA ID first. However, to get additional efficiencies into the sample design, it would be better to sort by other variables before sorting by LEA ID (see below). To achieve both of these goals, some of the sort variables' values were recoded to make them the same for every school within a stratum/LEA. They were changed in the following manner:

1) All schools within a stratum/LEA had the first three digits of the ZIP code set equal to the ZIP code of the first school in the stratum/LEA.

2) All schools within a stratum/LEA had the urbanicity code changed to the urbanicity code most prevalent among all schools within the stratum/LEA. If a tie occurred, the lower value was used.

After these fields were changed the nonBIA schools within a stratum were sorted by the following variables:

1) State;

2) LEA urbanicity
0 - unclassified
1 - central city of a Metropolitan Statistical Area (MSA)
2 - MSA, not central city
3 - outside MSA;

3) LEA Zip code — The first three digits were used;

4) CCD LEA ID number[11];

5) LEA percent minority (obtained by summing Number of Black, Hispanic, Asian/Pacific Islander and American Indian/Alaskan students and dividing by total enrollment
1 - (0-5%)
2 - (6-20%)
3 - (21-50%)
4 - (51% or more);

6) Highest grade in school;

7) School enrollment; and

8) CCD School ID[12] - for collapsed schools, the CCD ID of the last school was used.

BIA schools were sorted on the following variables:

1) BIA operated
1 - operated by the BIA
0 - not operated by the BIA, but rather operated by the tribe

2) School enrollment.

[11]CCD LEA ID is a unique number assigned to each school district by NCES.

[12] CCD School ID is a unique number assigned to each school.

5.1.5 Sample Selection

Within each stratum, schools were systematically selected using a probability proportionate to size algorithm. The measure of size used for the schools on CCD was the square root of the number of teachers in the school as reported on the CCD file. For BIA schools, the measure of size used was the square root of enrollment. 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.

5.2 LEA Sample for Public Schools

5.2.1 LEAs with Schools

During the initial design development of the SASS, consideration was given to selecting the LEAs first and then selecting schools within LEAs. It was hypothesized that doing this would reduce the reliability of both school and teacher estimates, but might be offset by the improvement in reliability of LEA estimates. Simulations done on the reliability of LEA estimates when the LEAs were selected first confirmed the loss in reliability for school and teacher estimates.[13] The simulations also showed that selecting school "first" would produce only slightly less accurate LEA estimates. For these reasons the SASS sample design selected schools first.

Hence, the LEA sample consists of the set of LEAs that were associated with the SASS public school sample. This provides the linkage between the LEA and the school. Table 11 provides the number of LEAs selected by state. Each Bureau of Indian Affairs school and Department of Defense school was defined to be an LEA. This portion of the LEA sample represented the set of LEAs associated with schools.

5.2.2 LEAs without Schools

Some LEAs were not associated with schools. Such LEAs may hire teachers who teach in schools from other LEAs. For SASS to represent teachers in these LEAs, a sample of these LEAs was also selected. The frame for this sample consisted of all LEAs on the 1988-89 CCD file that were not associated with schools. 1,352 LEAs were on this frame.

A 1 in 10 sample was taken. The sample was selected using a systematic equal probability algorithm. The sort variables were:

---

[13]"Early Elementary and Secondary Redesign Research - The Schools and Staffing Survey" by Doug Wright, 1988, an internal NCES paper.

1) State (50): one for each state;

2) Metro Status: central city of MSA (Metropolitan Statistical Area), outside central city of MSA, non-MSA;

3) First three digits of Zip Code; and

4) LEA ID.

Some 135 LEAs were selected and only 14 of the 135 sampled LEAs were actually in-scope (an operating public school agency that reported hiring teachers in SASS).

5.2.3 Delaware, Nevada and West Virginia LEAs

For each state, a simulation study was done to assess the reliability of SASS LEA estimates. The study showed that standard errors from Delaware, Nevada and West Virginia were very high relative to the LEA sampling rate (i.e., CVs of 5 to 20 percent with 90 percent of LEAs in sample). To reduce the standard error, all LEAs were defined as school sampling strata, placing all LEAs in the LEA sample, and reducing the standard error to zero.

Table 11.--Number of sampled public LEAs by State

State	LEAs	State	LEAs
Total	5424	Missouri	148
Alabama	103	Montana	147
Alaska	41	Nebraska	119
Arizona	101	Nevada	17
Arkansas	119	New Hampshire	81
California	265	New Jersey	155
Colorado	78	New Mexico	56
Connecticut	94	New York	175
Delaware	19	North Carolina	93
District of Columbia	1	North Dakota	114
Florida	55	Ohio	178
Georgia	101	Oklahoma	196
Hawaii	1	Oregon	106
Idaho	78	Pennsylvania	198
Illinois	179	Rhode Island	35
Indiana	134	South Carolina	70
Iowa	127	South Dakota	94
Kansas	113	Tennessee	94
Kentucky	101	Texas	270
Louisiana	66	Utah	35
Maine	107	Vermont	102
Maryland	23	Virginia	95
Massachusetts	131	Washington	104
Michigan	171	West Virginia	55
Minnesota	184	Wisconsin	131
Mississippi	119	Wyoming	45

5.3 Private School Sample

This section describes the frame, stratification, sorting and private school sample selection. The private school allocation is described in the School and Teacher Allocation section (See section 3).

5.3.1 Frames

3,271 private schools were selected using a dual frame approach. A list frame was the primary private school frame, and an area frame was used to find schools missing from the list frame and thereby compensating for the coverage problems of the list frame.

The 3,271 schools mentioned above include one more school than originally designated. This difference is due to the randomness of the sample sizes introduced by the school overlapping procedures described in section 4.

5.3.2 List Frame

The list frame used for private schools was the 1989 Private School Survey (PSS) list frame. NCES initiated PSS to build a universe frame of private schools. The PSS list frame universe is based on the 1988-89 QED private school list updated with private school association lists given to the Census Bureau in the spring of 1989. Various private school associations were asked to supply lists of their schools. Twenty such lists were received. These lists were matched with the QED list and any association list school not found on the QED file was added to the frame. Before sampling, duplicate schools were excluded from the frame. Schools that only teach prekindergarten, kindergarten or adult education were also removed. The list frame consisted of approximately 20,600 schools.

In section 5.1.2, a description is given about the differences between the way CCD and QED define a public school. For private schools, this is not as big a concern, since the NCES does not publish private school estimates from any other source that would be inconsistent with the PSS.

5.3.3 Area Frame

The area frame sample consisted of two sets of sample PSUs: 1) a subsample of the 1988 SASS area frame sample PSUs; and 2) sample PSUs selected independently from the 1988 SASS sample. The 1988 SASS sample PSUs were selected systematically with probabilities proportional to the square root of enrollment from each of sixteen strata defined by Census region, metro/nonmetro status, and high or low percent of enrollment in private schools. By maintaining a

fifty percent overlap of PSUs, the reliability of estimates of change was maintained at a reasonable level, while reducing respondent burden.

The eight certainty PSUs in the 1988 SASS remained in the 1991 SASS sample with certainty. The 67 non-certainty PSUs in the 1988 SASS were first sorted by PSU stratum code and then a subsample of 52 PSUs was systemically selected with equal probability. The total of 60 sample PSUs from the 1988 SASS formed the first set of sample PSUs for the 1991 SASS area frame. They were the overlap PSUs with the 1988 SASS.[14]

An additional 64 PSUs were selected independently. The United States was divided up into primary sampling units (PSUs). Each PSU consisted of a single county, independent city or cluster of geographically contiguous areas defined so that each PSU had a minimum population of 20,000 according to population projections for 1988. To avoid having PSUs covering too large a geographic area some PSUs had less than 20,000 in population. Counties of Alaska were excluded until the 1989 PSU Sample was selected. In other words, there were no Alaska counties in the overlap sample but there was a county in Alaska in the non-overlap sample. The eight certainty PSUs in 1988 were also excluded from the independent PSU sampling operation.

The strata were defined the same way as in the 1988 area frame design: a) Census region (4 levels - See Section 3.2 for a description), b) metro/nonmetro status (2 levels) and c) whether the PSU's percent private school enrollment exceeded the median percent private enrollment of the other PSUs in the census region/metro status strata (2 levels - using 1980 Census data).

A minimum of two PSUs were allocated to each of the 16 strata (32 PSUs). 32 additional PSUs were allocated to the 16 strata to more nearly approximate a uniform sampling fraction of PSUs from each stratum.

The PSUs were selected as a systematic sample with probability proportionate to the square root of the 1988 projected PSU population. A total of 123 distinct PSUs were in sample since one PSU was selected for both sets of samples. Its weight was adjusted to appropriately reflect the duplication.

The total private school sample size was 3,500 in 1988. This was the base for the 1991 sample. The 3,500 was reduced by 230 schools to pay for the extra 48 area PSUs introduced in 1991 (3,270 schools). Eighteen percent of these schools

[14]For details of how the original 67 non-certainty PSUs in the 1988 SASS were selected refer to pages 28-29 of the 1988 Schools and Staffing Survey Sample Design and Estimation, Technical Report NCES 91-127, dated May 1991.

(600 schools) made up the area sample. The 1991 total list frame sample was then 2670 schools, with 600 schools for the area sample.

5.3.4 Area Sample Frame Building

Within each of the PSUs, the Census Bureau attempted to find all eligible private schools (i.e., nonpublic schools providing the following: instruction for any grades 1-12, instruction not provided exclusively in the home, a school year at least 160 days long, and a normal school day at least 4 hours long). An area canvas was not attempted. However, regional field staff were used to create the frame using such sources as: yellow pages, non-Roman Catholic religious institutions, local education agencies, Chamber of Commerce, and local government offices. Roman Catholic religious institutions were not contacted because QED calls each Catholic diocese during its annual list update. Once these lists of schools were constructed, they were matched with the updated list frame school file. Schools that did not match the list were contacted to make sure they were eligible schools.

5.3.5 Private School List Frame Sample

5.3.5.1 Stratification

For private schools, the list frame was partitioned into an initial set of 216 cells. The first level of stratification was school association membership (18):

1) Military - membership in the Association of American Military Colleges and Schools;

2) Catholic - affiliation as Catholic or membership in the National Catholic Education Association or the Jesuit Secondary Education Association;

3) Friends - affiliation as Friends or membership in the Friends Council on Education;

4) Episcopal - affiliation as Episcopal or membership in the National Association of Episcopal Schools association;

5) Hebrew Day - membership in the National Society for Hebrew Day Schools association;

6) Solomon Schechter - membership in the Solomon Schechter Day Schools;

7) Other Jewish - other Jewish affiliation;

8) Missouri Synod - membership in the Lutheran Church, Missouri Synod school association;

9) Wisconsin Synod - membership in the Evangelical Lutheran Church - Wisconsin Synod school association or affiliation as Evangelical Lutheran - Wisconsin Synod;

10) Evangelical Lutheran - membership in the Association of Evangelical Lutheran Churches school association or affiliation as Evangelical Lutheran Church in America;

11) Other Lutheran - other Lutheran affiliation;

12) Seventh-Day Adventist - affiliation as Seventh-Day Adventist or membership in the General Conference of Seventh-Day Adventists;

13) Christian Schools International - membership in Christian Schools International;

14) Association of Christian Schools International - membership in the Association of Christian Schools International;

15) National Association of Private Schools for Exceptional Children - membership in the National Association of Private Schools for Exceptional Children;

16) Montessori - membership in the American Montessori Society or other Montessori associations;

17) National Association of Independent Schools - member of the National Association of Independent Schools;

18) All else - member of any other association specified in the PSS or affiliated with a group not listed above or not a member of any association.

Within each association membership, schools were stratified by grade level (elementary, secondary, and combined schools). The definitions are provided below:

Regular Schools:

Elementary: Lowest grade $\leq$ 6 and Highest grade $\leq$ 8

Secondary: Lowest grade $\geq$ 7 and Highest grade $\leq$ 12

Combined: Lowest grade $\leq$ 6 and Highest grade > 8

Nonregular School:

Nonregular schools, which include special education, vocational, technical, adult education (if part of in-scope school) or alternative/continuation grades were classified as combined schools.

Within association/grade level, schools were stratified by four Census regions: Northeast, Midwest, South, and West. For a definition of the four Census Regions, see Section 3.2.

5.3.5.2 School Sorting

Within each stratum, sorting took place on the following variables:

1) State (51): 1 for each state and the District of Columbia;

2) Urbanicity: 0 - unclassified
1 - urban
2 - suburban
3 - rural
4 - affiliation adds with no classification;

3) Zip code: The first two digits were used;

4) Highest grade in the school;

5) 1989 PSS Enrollment;

6) PIN number: The PIN number is a unique number assigned by QED which identifies the school on PSS. PIN numbers for schools added from private affiliation list adds were assigned by the Census Bureau.

5.3.5.3 Sample Selection

Within each stratum, schools were systematically selected using a probability proportionate to size algorithm. The measure of size used was the square root of the 1989 PSS number of teachers in the school. Any school with a measure of size larger than the sampling interval was excluded from the probability sampling process and included in the sample with certainty.

5.3.6 Area Frame Sample

5.3.6.1 Stratification

Schools were stratified by the following variables:

1) PSU;

2) Grade level - elementary, secondary, and combined. In some cases, when the grade level was unknown, it was imputed based on the school name.

5.3.6.2 Sort Variables

Eligible schools were sorted using the following variables:

1) Affiliation - Catholic, other religious, and nonsectarian. In some cases, when the affiliation was unknown, it was imputed based on the school name;

2) 1989 PSS enrollment; and

3) School name (in alphabetical order).

5.3.6.3 Measure of Size

For eligible schools, the measure of size was the square root of the number of reported teachers.

5.3.6.4 Sample Selection

Within each stratum, eligible schools were systematically selected using a probability proportionate to size algorithm. The measure of size used was the square root of the reported number of teachers. 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.

6. Public and Private Teacher Sample

This section describes the frame, stratification, sorting, and the sample selection for the public and private teacher sample.

Selecting the teacher sample involved the following steps. First, the selected schools were asked to provide teacher lists for their schools. From the teacher lists, 56,051 public school teachers and 9,166 private school teachers were selected.

The public and private school teacher samples will be described together because they were selected using identical methodology. The only differences were in the average number of teachers selected within a school (See section 3.4.2, table 7).

The details of the teacher selection are provided below.

6.1 Teacher Frame

Each selected school was asked to provide a list of their teachers with the following information:

1) *New/experienced.* Teachers in their first, second, or third year of teaching are classified as new teachers.

2) *Race/Ethnicity.* 1. White (non-Hispanic); 2. Black (non-Hispanic);3. Hispanic; 4. Asian or Pacific Islander (API); and 5. American Indian, Aleut, or Eskimo (AIAE).

3) *Bilingual/ESL.* Teachers who use NATIVE LANGUAGE to instruct students with limited English proficiency (bilingual); or Teachers providing students with limited English proficiency with intensive instruction in English (English as a Second Language).

4) *Field of Teaching.* Elementary teachers were classified as: general elementary, special education or other.

   Secondary teachers depending on their primary subject taught were classified as: math, science, English, social studies, vocational education or other.

The above information for each teacher in a selected SASS school comprises the school teacher frame.

Eleven percent of the in-scope private schools and five percent of the in-scope public schools did not provide teacher lists. For these schools no teachers were selected.

A factor in the teacher weighting is used to adjust the weights to reflect the fact that some schools did not provide teacher lists.

6.2 Teacher Stratification

Within each selected school, teachers were stratified into one of five teacher types in the following hierarchical order:

1) Asian or Pacific Islander (API)
2) American Indian, Aleut, or Eskimo (AIAE)
3) Bilingual/ESL
4) New (less than 3 years completed in the teaching profession)
5) Experienced (3 or more years completed teaching)

6.3 Teacher Sorting

Within each teacher stratum, teachers were sorted by primary field of teaching. Elementary teachers were sorted by general elementary, special education or other teaching assignment. Secondary teachers were sorted by math, science, English, social studies, vocational education or other teaching assignment. When combined schools had both elementary and secondary teachers, the teachers were sorted by grade level/primary field of teaching. This was done to assure a good distribution of teachers by field of teaching.

6.4 Teacher Selection

Within each school and teacher stratum, teachers were selected systematically with equal probability.

65,088 teachers were designated for selection (approximately 6,260 new and 53,828 experienced; 1,500 API; 1,500 American Indian, Aleutian, or Eskimo, and 2,000 bilingual/ESL), while 65,217 were actually selected (approximately 7,972 new and 52,084 experienced; 1,511 Asian Pacific Islander; 1,529 American Indian, Aleutian, or Eskimo and 2,121 bilingual/ESL). This slight difference was due to the fact that in allocating the sample, Y, the average of the school's weighted measure of size over all schools in the school stratum, was based on universe files of teacher counts for the previous year (CCD for public, PSS for private) instead of reported teacher counts from the school just prior to data collection. This caused the overall average number of teachers per school to be slightly different than the target numbers in Table 7.

To reduce the variance of teacher estimates, one goal of the teacher selection was to make the teacher sample self-weighting (i.e., equal probabilities of selection). The goal was generally met within teacher stratum within school stratum. However, since the school sample size of teachers was altered due to the minimum constraint (i.e., at least 1 teacher/school) or maximum constraint (i.e., no more than either twice the average stratum allocation or 20 teachers/school), the goal of self-weighted teachers was lost for some schools.

Census estimated the Q, R, and H factors (mentioned in the Allocation section 3.4.2) conservatively so that there would be more than the designated number of API, AIAE, and bilingual/ESL teachers in sample. These groups of teachers were randomly assigned to 101 reduction groups. After sampling was completed certain reduction groups from each of these groups of teachers were eliminated to sample back to the designated number of API, AIAE, and bilingual/ESL teachers. The reduction groups were eliminated at different rates among these strata.

7. Estimation

7.1 Weighting

This section describes the weighting processes for the different SASS samples. The general purpose of the weighting steps is to adjust for nonresponse using respondents' data; and also to adjust the sample totals to the frame total to reduce sampling variability. For each questionnaire, the formula for the weight will be presented, along with a brief description of each component. When computations are done within cells, such as nonresponse adjustments, the cells will be described. Sometimes a cell did not have enough data to produce a reliable estimate; in such cases, cells were collapsed. The least important variables were always collapsed first. The collapsing criteria are also described.

First, the school weight will be described. Since the public and private school weights have the same structure, they will be presented together. They differ only in the definition of the cells used to compute the nonresponse adjustment factor and the first-stage ratio adjustment factor. These cells will be described separately within the school weight section. Since the public and private administrator weights are similar to the school weights, they will be described next. In the third section, the public teacher demand and shortage weights will be described. In the last weighting section, the teacher weights will be described. Since the public and private school teacher weights have the same structure, they will be presented together. They differ only in the definition of the cells used to compute the various weighting factors. These cells will be described separately within the teacher weight section.

7.1.1 School Weight (SASS Questionnaire Forms 3A, 3B, and 3C)

The final weight for the public and private school data is:

(Basic Weight) X (Sampling Adjustment Factor) X (School Noninterview Factor) X (Frame Ratio Adjustment Factor)

Where:

Basic Weight is the inverse of the probability of selection of the school.

Sampling Adjustment Factor is an adjustment that accounts for unusual circumstances that affect the school's probability of selection, such as a merger, split, or duplication(e.g., a Junior/Senior High which splits into separate Junior and Senior High Schools).

Noninterview Adjustment Factor is an adjustment that accounts for total school nonresponse. It is the weighted (basic weight*sampling adjustment factor) ratio of the total eligible in-scope schools to the total responding in-scope schools within cells.

Frame Ratio Adjustment Factor. is a factor that adjusts the sample estimates to known frame totals. For public schools, it is equal to the ratio of the total number of SASS frame noncertainty schools to the weighted sample estimate of the total number of noncertainty schools within each cell in the frame. For private schools, the adjustment is the same, except for the area frame. For the area frame, only sample schools in certainty PSUs were adjusted to the PSU totals from the area frame since all schools in the non-certainty PSUs were in sample and we did not have universe counts for all non-certainty PSUs. Certainty schools were excluded from the numerator and denominator and their factor set equal to 1. Also, schools from noncertainty PSUs had their factors set equal to 1, since no subsampling took place.

7.1.2 School Weighting Adjustment Cells

School noninterview and frame ratio adjustments are computed within cells.

7.1.2.1 Public School Adjustment Cells

For public schools, (except BIA and Native American schools) the noninterview adjustment cells were: state by school grade level by enrollment size class by urbanicity. If the factor was less than 1.5 and there were at least 15 schools in the cell, no collapsing was done. Otherwise, cells were collapsed (enrollment size class first, urbanicity second, and grade level third).

For BIA elementary schools, the noninterview adjustment cells were grade level by enrollment size class; while BIA secondary schools cells were by grade level. Cells for Native American elementary schools were grade level by state (4 levels) by enrollment size class; while secondary school cells were grade level by state (4 levels). If the factor was less than 2.0 and there were at least 10 schools in the cell, no collapsing was done. Otherwise, cells were collapsed in the same sequence as in other public schools.

The frame ratio adjustment cells for public schools (except BIA and American Indian schools) were state by grade level by urbanicity; for BIA schools, they were grade level; and for Native American schools, they were state (4 groups) by grade level. If the factor was between 0.667 and 1.5 and there were at least 15 (10 for BIA and American Indian Schools) non-certainty schools in the cell, no collapsing was done. Otherwise, cells were collapsed (urbanicity first, grade level second, and state third).

7.1.2.2 Private School Adjustment Cells

For private list frame schools, the noninterview adjustment cells were: 18 associations by grade level. The Catholic and All Else associations additionally used urbanicity to define the cells. If the factor was less than 2.0 and there were at least 15 schools in the cell, no collapsing was done. If collapsing was done, urbanicity was collapsed first (for Catholic and ALL ELSE associations), grade level second and association last. The frame ratio factor cells were the same as the noninterview adjustment cells. If the factor was between 0.667 and 1.5 and there were at least 15 non-certainty schools in the cell, no collapsing was done. Otherwise, cells were collapsed (urbanicity first, grade level second, and association last).

For private area frame schools, the noninterview adjustment cells were: affiliation (Catholic, other religious, nonsectarian, and unknown) by grade level by enrollment size class. If the factor was less than 2.0 and there were at least 15 schools in the cell, no collapsing was necessary. If collapsing was necessary, the enrollment size class was collapsed first, grade level was second, and affiliation was collapsed last.

The frame ratio factor cells were: grade level by certainty PSU. If the factor was between 0.667 and 1.5 and there were at least 15 noncertainty schools in a certainty PSU in the cell, no collapsing was done. Otherwise, cells were collapsed: grade level first and PSU second.

7.1.3 Administrator Weight (SASS Questionnaire Forms 2A and 2B)

The public and private administrator weighting was done the same way as the school questionnaire weighting described above. Since the respondents for each of the administrator surveys and the corresponding school surveys could be different, the weighting process was done separately for each questionnaire. The sum of the administrator weights may not equal the sum of the school weights because some schools do not have administrators.

7.1.4 Teacher Demand and Shortage for Public Districts (SASS Questionnaire Form 1A)

The final weight for the public school district data is:

(Basic Weight) X (Sampling Adjustment Factor) X (LEA Noninterview Factor) X (Frame Ratio Adjustment Factor)

where:

Basic Weight is the inverse of the probability of selection of the LEA. Note that LEAs were not selected directly, so the computation of this probability is rather complex. See section 7.1.5 for more details.

Sampling Adjustment Factor is an adjustment that accounts for unusual circumstances that affect the LEA's probability of selection, such as a merger, split or duplication. For example, if two LEAs consolidated into one.

Noninterview Adjustment Factor is an adjustment that accounts for total LEA nonresponse. It is the weighted (basic weight*sampling adjustment factor) ratio of total eligible in-scope LEAs to the total responding in-scope LEAs, computed within cells.

Frame Ratio Adjustment Factor is a factor that adjusts the sample estimates to known frame totals. It is the ratio of the total number of noncertainty LEAs in the frame to the weighted sample estimate of the total number of noncertainty LEAs in the frame, computed within cells. Certainty LEAs were assigned a factor of 1.

Noninterview and frame ratio adjustments are computed within cells. The noninterview adjustment cells were: state by LEA enrollment size class by metro status (central city of MSA, outside central city of MSA, outside MSA) for LEAs with schools, and metro status only for LEAs without schools. If the factor was less than 1.5 and there were at least 10 LEAs in the cell, no collapsing was done. Otherwise, cells were collapsed (LEA enrollment size class first and metro status second).

The frame adjustment cells were the same as the noninterview adjustment cells. If the factor was between 0.667 and 1.5 and there were at least 10 noncertainty LEAs in the cell, no collapsing was done. Otherwise, cells were collapsed: LEA enrollment size class first and metro status second.

7.1.5 LEA Basic Weights

Given the complexity of the sampling scheme, the calculation of the LEA basic weights is not straightforward. There are three situations that need discussion: LEAs with schools, LEAs without schools and LEAs in Delaware, Nevada and West Virginia which are all certainty LEAs.

7.1.5.1 LEAs with Schools

The LEA sample was not selected directly through an LEA frame. Instead, the LEAs were selected through the school (i.e., the LEAs associated with the school sample comprised the LEA sample). The basic weight, therefore, is more complicated than normal.

Since schools were stratified by grade level (elementary, secondary, and combined), the probability of selection for LEA k, (P_k(sel)) can be written as follows:

$$P_k(Sel) = 1 - (1 - P_k(Nam,El))(1 - P_k(Nam,Sec))(1 - P_k(Nam,Com))(1 - P_k(Pub,El))(1 - P_k(Pub,Sec))(1 - P_k(Pub,Com))$$

where: P_k(Nam,El) is the probability of selecting LEA k in elementary schools which are classified as Native American schools. This equals the sum of the school selection probabilities for the schools which are Native American, elementary, and in LEA k. If the sum is greater than one then P_k(Nam,El) is set equal to one.

P_k(Nam,Sec) is the probability of selecting LEA k in secondary schools which are classified as Native American schools. This equals the sum of the school selection probabilities for the schools which are Native American, secondary, and in LEA k. If the sum is greater than one, then P_k(Nam,Sec) is set equal to one.

P_k(Nam,Com) is the probability of selecting LEA k in combined schools which are classified as Native American schools. This equals the sum of the school selection probabilities for the schools which are Native American, combined, and in LEA k. If the sum is greater than one P_k(Nam,Com) is set equal to one.

P_k(Pub,El) is the probability of selecting LEA k in elementary schools which are not Native American. This equals the sum of the school selection probabilities for the schools which are not Native American, are elementary and in

LEA k. If the sum is greater than one then P_k(Pub,El) is set equal to one.

P_k(Pub,Sec) is the probability of selecting LEA k in secondary schools which are not Native American. This equals the sum of the school selection probabilities for the schools which are not Native American, are secondary and in LEA k. If the sum is greater than one then P_k(Pub,Sec) is set equal to one.

P_k(Pub,Com) is the probability of selecting LEA k in combined schools which are not Native American. This equals the sum of the school selection probabilities for the schools which are not Native American, are combined and in LEA k. If the sum is greater than one then P_k(Pub,Com) is set equal to one.

7.1.5.2 LEAs Without Schools

The basic weight for LEAs that have no associated schools was 10, since these LEAs were selected with equal probability at a rate of 1 in 10.

7.1.5.3 LEA Basic Weights for Delaware, Nevada and West Virginia

The basic weight is 1 for all LEAs in Delaware, Nevada and West Virginia since all LEAs in these three states were guaranteed being selected for sample.

7.1.6 Teacher Weights (SASS Questionnaire Forms 4A and 4B)

The final weight for public and private school teachers is:

(Basic Weight) X (Sampling Adjustment Factor) X (School Nonresponse Factor) X (Teacher-Within-School Noninterview Adjustment Factor) X (Frame Ratio Adjustment Factor) X (Teacher Adjustment Factor)

where:

Basic Weight is the inverse of the probability of selection of the teacher.

School Sampling Adjustment Factor is an adjustment that accounts for unusual circumstances that affect the school's probability of selection, such as a merger, split or duplication. It is the same factor used in the school weight.

School Nonresponse Adjustment Factor is an adjustment that accounts for schools that did not have teachers selected because teacher lists were not provided by the school. It is the weighted (school basic weight*school sampling adjustment factor) ratio of total eligible in-scope schools to the total in-scope schools providing teacher lists, computed within cells.

Teacher-within-school noninterview adjustment factor is an adjustment that accounts for sampled teachers that did not respond to the survey. It is the weighted (product of all previously defined components) ratio of the total eligible teachers to the total eligible responding teachers computed within cells.

Frame Ratio Adjustment Factor. is a factor that adjusts the sample estimates to known frame totals of number of teachers. For the set of noncertainty schools, the factor is the ratio of the frame estimate of the total number of teachers to the weighted (all previously defined components) sample estimate of the total number of teachers. These factors are computed within cells. The sample estimate uses the frame count of the number of teachers in the school. For public schools, the 1990 CCD was used as the frame and the teacher counts were in terms of FTEs. For private schools, the 1989 PSS was used as the frame and teachers were in terms of headcounts.

For the set of certainty schools, the factor is 1.

Teacher Adjustment Factor. is a factor that adjusts the inconsistency between the estimated number of teachers from the SASS school data files and the SASS teacher sample files. It is the ratio of weighted number of teachers from the school data file for a cell to the weighted number of teachers on the teacher data file for a cell. The weight is the product of all previously defined components. This factor ensures that teacher aggregates from the school file (after imputation) will agree with the corresponding teacher estimates from the teacher file.

The school nonresponse adjustments, the teacher within-school noninterview adjustments, the frame ratio adjustments, and the teacher adjustments are computed within cells. The cells for the frame ratio adjustments are the same as those used in the school weight except for BIA schools where no frame ratio adjustment was done for the teacher weight

because no teacher data existed on the BIA school sample frame. The cells for the frame adjustments are described in the school weight section.

7.1.6.1 Public Adjustment Cells

For public schools, the school nonresponse adjustment cells were: the same as those used for the school noninterview adjustment cells in the school weight except that enrollment size classes were replaced by teacher size classes for Native American schools and other public schools. The collapsing criteria were also the same as those used in the school noninterview adjustment in the school weight.

The teacher within-school noninterview adjustment cells were: state by field of teaching by teacher type (new, experienced, bilingual, Asian, American Indian) by school urbanicity (only for experienced teachers). If the factor was less than 1.5 and there were at least 15 teachers in the cell, no collapsing was done. Otherwise, cells were collapsed (urbanicity first, teacher type second, and field of teaching third).

The teacher adjustment cells and the collapsing criteria were the same as those used for the frame ratio adjustments in the school weight.

7.1.6.2 Private Adjustment Cells

For private list frame schools, the school nonresponse adjustment cells were the same as those used for the school noninterview adjustment cells in the school weight, and the collapsing criteria were also the same.

The teacher within-school noninterview adjustment cells were: association membership (18 levels) by field of teaching by experience level (new/experienced). Urbanicity was additionally used to define cells in the Catholic and ALL ELSE associations. If the factor was less than 1.5 and there were at least 15 teachers in the cell, no collapsing was done. If collapsing occurred, urbanicity was collapsed first (for Catholic and ALL ELSE associations), teaching experience was collapsed second, field of teaching was collapsed third, and association was collapsed last.

The teacher within-school noninterview adjustment cells were: association membership (18 levels) by field of teaching by experience level (new/experienced) by urbanicity for (Catholic schools and All Else schools only). If the factor was less than 1.5 and there were at least 15 teachers in the cell, no collapsing was done. If collapsing occurred, urbanicity was collapsed first, teaching experience was collapsed second, field of teaching was collapsed third, and association was collapsed last. The teacher

adjustment cells and the collapsing criteria were the same as those used for the frame ratio adjustments in the school weight.

For private schools found on the area frame, the school noninterview adjustment cells were: affiliation (three levels plus Don't Know) by grade level by teacher size class. If the factor was less than 2.0 and there were at least 15 schools in the cell, no collapsing was done. If collapsing occurred, teacher size class was collapsed first, grade level was collapsed second, and affiliation was collapsed last.

The teacher within-school noninterview adjustment cells were: affiliation (four levels) by field of teaching by teaching experience (new/experienced). If the factor was less than 1.5 and there was at least 15 teachers in the cell, no collapsing was done. If collapsing was done, teaching experience was collapsed first, field of teaching was collapsed second, and affiliation was collapsed last.

The teacher adjustment cells and the collapsing criteria were the same as those used for the frame ratio adjustments in the school weight.

8. Item Response Rates and Imputation

8.1 Item Response Rates

The unweighted item response rates (i.e., the number of sampled units responding to an item divided by the number of responding sampled units) for the SASS surveys ranged from 25 percent to 100 percent. Tables 12 and 13 provide a brief summary of the item response rates. The item response rates in these tables are unweighted, and do not reflect additional response loss due to complete questionnaire refusal.

Table 12.--Summary of Item response rates by Questionnaire (unweighted)

Survey	Range of item response rates	Percent of items with a response rate of 90% or more	Percent of items with a response rate less than 75%
LEA Survey			
Public	85-100%	90%	0%
Administrator Survey			
Public	90-100%	100%	0%
Private	80-100%	98%	0%
School Survey			
Public	56-100%	77%	1%
Private	67-100%	77%	5%
Indian	60-100%	87%	4%
Teacher Survey			
Public	76-100%	84%	0%
Private	71-100%	79%	1%

Table 13.--Items with response rates less than 75 percent[15]

LEA Survey	
Public	None
Administrator Survey	
Public	None
Private	None
School Survey	
Public	31 (part-time), 2, 33d(3)-33d(9)
Private	18b(1), 35d(2), 35d(7), 35d(9), 43a-d,f-i (parttime)
Indian	42a-d,f-i (part-time)
Teacher Survey	
Public	None
Private	20b(undergraduates), 20b(graduates)

8.2 Imputation Procedures

Values were imputed for items with missing data for records that had been classified as interviews (ISR=1). Noninterview adjustment factors were used during the weighting process to compensate for data that were missing because the sample case was a noninterview (ISR=2).

For questionnaire items that should have been answered but were not, values were imputed either in the first or second stage imputation procedures. The first stage imputation process is done by: (1) using data from other items on the questionnaire; (2) extracting data from a related component of the Schools and Staffing Survey (for example, using data from a school record to impute missing values on that school's LEA questionnaire); or (3) extracting data from the sample file (information about the sample case from other sources; for example, the Private Schools Survey or the Common Core of Data, collected in the 1988-89 school year). If a value is not imputed in the first stage processing, it is imputed during the second stage processing.

The second stage imputation process is done using a nearest neighbor hot deck imputation methodology, extracting data from the respondent (donor) closest to the

[15]The questionnaire wording for these items can be found in The Schools and Staffing Surveys: 1990-1991, Data File User's Manual, an NCES publication

nonrespondent, with similar characteristics. Matching variables are used to define cells of respondents which are assumed to be similar to the nonrespondent. All units (donors and nonrespondents) within each cell are further sorted by the sorting variables. The last respondent encountered before reaching the nonrespondent (nearest neighbor) is used as the donor for the missing item.

For some incomplete items, the entry from another part of the questionnaire or information from the sample file was directly imputed to complete the item; for others the entry was used as part of an adjustment factor with other data on the incomplete record. For example, if a respondent did not report whether a school offered remedial reading in item 10c of the public school questionnaire, the response (1 = Yes or 2 = No) for a similar school was imputed to item 10c of the incomplete record. However, if a respondent had answered "Yes" to item 10c but had not reported the number of students in the program, the ratio of number of students in remedial reading to the total enrollment for a similar school was used with the enrollment at the school for which item 10c was incomplete to impute an entry to item 10c (i.e., SCHOOL A item 10c = SCHOOL A ENROLLMENT multiplied by the ratio of SCHOOL B item 10c to SCHOOL B ENROLLMENT).

8.2.1 Imputation Procedures: Teacher Demand and Shortage (TDS) Survey (SASS-1A)

Data were imputed to items with missing values in two stages.

First Stage - TDS Imputation

In the first stage, information about the same LEA but obtained from other sources (other questionnaire items on the same record, records for schools in the LEA, and the Common Core of Data) was used.

If the LEA with missing data had only one school in its jurisdiction and data for that school were collected in the 1991 SASS, information from the school record was used to fill items with missing values on the LEA record whenever possible. For one-school LEAs, the following items were imputed with school data when available:

TDS Items (SASS-1A)	School Source Items (SASS-3A)
1a-g ("This year's enrollment" column)	17
9c	31e (if values equal 0)
28	9
29	25

For LEAs with missing values that could not be imputed with school data, the next step was to use data from other items on the LEA's record. For example, if an LEA reported the 1990 enrollment (in column labeled "This year's enrollment") by the requested instructional levels but reported only the total enrollment for 1989 (column labeled "Last year's enrollment"), the 1989 total enrollment was allocated to the instructional levels by using the same proportions as 1990; e.g., 1989 ungraded enrollment equals 1989 total enrollment multiplied by the ratio of 1990 ungraded enrollment to 1990 total enrollment.

Items in Figure 1 were imputed using ratios or proportions from other items on the record for the same LEA. Note: All figures are located at the end of this section.

If the LEA's total enrollment for 1989 was not reported in item 1 (total for "Last year's enrollment" column), the total enrollment reported for the 1988 Common Core of Data was extracted from the LEA's sample data and imputed to item 1.

Second Stage - TDS Imputation

For items that still had missing values after stage 1 of the imputation was completed, a nearest neighbor hot deck method was used. Variables (SASS-1A imputation variables) which described certain characteristics of the LEAs (e.g., size, instructional levels, percent minority students, etc.) were created and used to sort the file and to match incomplete records to those with complete entries (donors). The imputation variables are defined in figure 2. The sort order is described below and the matching variables and collapse orderings are described in Figures 3 and 4.

During the second stage of imputation, items on the TDS (SASS-1A) questionnaire were grouped according to the relevance of the SASS-1A imputation variables to the data collected by the item. For example, LEALEVEL was the most important variable for matching incomplete records and donors to fill item 1 (students by grade level) but LEALEVEL was not used to match LEAs to impute item 18 (merit pay plan).

Items 1-6, 9-12, and 27 - The records were sorted by GROUP/State/LEALEVEL/MSC88/LEAENR. The matching variables and their order of collapse are given in Figure 3.

Items 7, 8, 13-26, 28, 29 - The records were sorted by GROUP/state/MSC88/ LEAENR. The matching variables and their order of collapse are given in Figure 4.

8.2.2 Imputation Procedures: Public School Administrators (SASS-2A) and Private School Administrators (SASS-2B)

Data were imputed in two stages:

First Stage - Administrator Imputation

During the first stage, items with missing values were filled by using other data from the same record or by making some assumptions about the respondent's intended answer (e.g., not answering means "No" or "None"). Values were imputed to the following items during the first stage: 1, 2a 2d, 3a, 4a, 5a, 5b, 8b, 9b-9e, and 19a.

Also during the first stage, imputation variables (SASS-2A/2B imputation variables) (Figure 5) were created from questionnaire data or copied from the matching school record. These variables were used during the second stage of imputation.

Second Stage - Administrator Imputation

The second stage imputation variables for the SASS-2A/2B nearest neighbor hot deck imputations are defined in Figure 5. The sort orderings for the administrator records are described below The matching variables and collapse orderings are provided in figures 6 and 7.

Public school administrators - The records were sorted by STATE/ GRDLEVEL/PUBURB/LEANUMBR (a code on the sample file record which identified the LEA with jurisdiction over the administrator's school)/MINEN/PUBENR. The matching variables and their order of collapse for items imputed in the second stage are given in Figure 6. Items are listed in the order of imputation.

Private school administrators - The records were sorted by AFFLG/AFFILS/ GRDLEVEL/PRVURB/MINEN/PRVENR. The matching variables and their order of collapse are given in Figure 7. Items are listed in the order of their imputation.

8.2.3 Imputation Procedures: Public Schools (SASS-3A)

Data were imputed to items with missing values in two stages.

First Stage - Public School Imputation

In the first stage, information about the same school but obtained from other sources (other questionnaire items on the same record, the record for the LEA with jurisdiction over the school, and the Common Core of Data) was used.

Values for a few items with missing values were imputed by using data for the LEA (SASS-1A) with jurisdiction over the school; for example, if the school did not report whether or not they had postsecondary students, but the LEA record indicated there were no postsecondary students in the district, then entries were imputed to item 17 to indicate that there were no postsecondary students in the school.

For schools in one-school LEAs, more data were extracted from the LEA record to impute values to the school record.

In Figure 8, SASS-3A items were imputed by using data from SASS-1A records:

During the first stage, items with missing values were also filled by using information from other items. The following items were imputed by this method: 3, 10c (yes/no), 11a, 12a, 13, 14a, 15a, 16a, 21, 33a.

If an item could not be filled by using LEA data or data from other items on the school record, information from the 1988 Common Core of Data (included in the school's sample file record) was used whenever possible. These items were filled by using the CCD data in the sample file: 1b, 2, 3, 4, 9a-9e, 15, 17, 23a.

Second Stage - Public School Imputation

For items that still had missing values after the first stage of imputation, a hot deck imputation method was used. Variables (SASS-3A imputation variables) (Figure 9) that described certain characteristics of the school (e.g., size, urbanicity, instructional level, etc.) were created and used to sort the records and to match incomplete records to those with complete data (donors). The sort orderings are described below. The matching variables and collapse orderings are described in Figures 10-12.

During the second stage of imputation, items on the SASS-3A questionnaire were grouped according to the relevance of the SASS-3A imputation variables to the data collected by the item. For example, TYPE was used for matching incomplete records and donors to fill item 10

(school programs and services) but was not used for item 8 (number of students absent).

Items are listed in the order in which they were imputed.

Items 17, 1b, 4, 6, 7, 10-13, 16, 18, and 20-23 - The records were sorted by STATE/GRDLEVEL/TYPE/MEAS (the square root of the number of teachers reported in the CCD). The matching variables and their order of collapse are given in Figure 10.

Items 8, 9, 14-15, 25, 26 - The records were sorted by STATE/GRDLEVEL/MINEN/PUBURB/CNTY (the county where the school is located)/MEAS. The matching variables and their order of collapse are given in Figure 11.

Items 24, 27-35 - The records were sorted by STATE/TYPE/ GRDLEVEL/PUBURB/CNTY/MEAS. The matching variables and their order of collapse are given in Figure 12.

Items listed in the text, but not shown in the figures, did not require the selection of a separate donor. For example, item 1a was imputed from data in item 17, after item 17 was imputed, if necessary.

8.2.4 Imputation Procedures: Private Schools (SASS-3B)

Data were imputed to items with missing values in two stages.

First stage - Private School Imputation

In the first stage of imputation, values for missing items were imputed by using other information on the questionnaire and information collected for the sample school in the 1988-89 Private Schools Survey.

The items were filled by using information from other questionnaire items: 14, 16, 17, 20, 24, 27, 32a, 33, 34, 35, 41a.

If an item could not be filled by using data from other questionnaire items, information from the 1988-89 Private Schools Survey or other sources included in the sample file record (or added to the school record) was used whenever possible. The following items were filled by using PSS or sample file information: 2, 4, 12a-12d, 13, 14, 16a, 26a, 27.

Second stage - Private School Imputation

For items that still had missing values after the first stage of imputation, a hot deck imputation method was used. Variables (SASS-3B imputation variables) (Figure 13) that described certain characteristics of the schools (e.g., size, urbanicity, instructional level, etc.) were created and used to sort the records and to match incomplete to those with complete data (donors). The sort orderings are provided below. The matching variables and collapse orderings are provided in Figures 14-17.

During the second stage of imputation, items on the SASS-3B questionnaire were grouped according to the relevance of the imputation variables to the data collected by the item. For example, urbanicity (PRVURB) was used for matching incomplete records and donors to fill item 7 (students by racial categories) but was not used for item 32 (number of newly hired teachers in FTEs). The items below are listed in the order in which they were imputed.

Items 2, 27, 28, 5, 15, 16, 22-26, 36, 39, 43, 44, 35, 41, 42, and 32: The records were sorted by GRDLEVEL/AFFILG/AFFILS/TYPE/ HIGHGRADE/MEAS (the square root of the number of teachers reported in the 1989 Private Schools Survey). The matching variables and their order of collapse are given in Figure 14.

Items 3, 6, 10, 12b, 13, 19, 20, 45-48, 58: The records were sorted by AFFLG/AFFILS/AFFILR/TYPE/PRVURB/REGION/STATE/MEAS. The matching variables and their order of collapse are given in Figure 15.

Items 4, 7, 9, 11, 18, 29, 30, 33, 37, 38, 49-57: The records were sorted by AFFLG/AFFILS/PRVURB/MINEN/REGION/MEAS. The matching variables and their order of collapse are given in Figure 16.

Items 8, 17, 34, 40: The records were sorted by AFFLG/AFFILS/ GRDLEVEL/HIGHGRADE/MEAS. The matching variables and their order of collapse are given in Figure 17.

8.2.5 Imputation Procedures: Bureau of Indian Affairs Schools (SASS-3C)

Because there were only 97 Bureau of Indian Affairs school records and the item response rates were very high for all items, imputation was done clerically. The computer records were sorted by BIA status (whether school was operated by the Bureau of Indian Affairs), state, and size so that records for similar schools were close together. The actual questionnaires were also reviewed for notes and other entries which were

useful in deciding the entries to be imputed. If an item could not be filled by using information on the questionnaire, entries from the record for a similar school were used.

8.2.6 Imputation Procedures: Public School Teachers (SASS-4A) and Private School Teachers (SASS-4B)

Data were imputed in two stages:

First stage - Teacher Imputation

During the first stage, items with missing values were filled by using other data from the same record or by making some assumptions about the respondent's intended answer (i.e., not answering a question implies a "No" response).

Values were imputed to the following items during the first stage if enough information were available: 3c, 6, 10, 11, 15, 16, 17, 18b, (Yes/No), 19b, 20b, 23, 27, 31b, 32, 34b, 45, 46, 50.

To impute missing values for item 32c, the average was calculated for the number of times per week (item 32c) for each subject-grade level combination (items 32b and 32d) on the file, and these averages were used to impute item 32c.

Also, during the first stage, imputation variables were created from questionnaire data or copied from the matching school record. These variables (SASS-4A/4B imputation variables) were used during the second stage of imputation. They are given in Figure 18.

Second stage - Teacher Imputation

During the second stage, a hot deck method of imputation was used to fill items that still had missing values. The variables listed in Figure 18 (except REGION) were used to sort the teacher records and to match incomplete records to records with complete data (donors). The sort orderings are provided below. The matching variables and collapse orderings are provided in Figures 19-20.

Items on the teacher questionnaire were grouped according to the relevance of the imputation variables to the data collected by the item.

Public school teachers

Items 2-4, 6-28, 30, 31, 32f-i, 33-56. The records were sorted by STGROUP/STATE/TEALEVEL/GRADELEV/PUBURB/TEAFIELD/PUBENR. The matching variables and their order of collapse for items imputed in the second stage are given in Figure 19. Items are listed in the order in which they were imputed.

Item 32e. The records were sorted by STGROUP/STATE/TEALEVEL/school enrollment. The matching variables and their order of collapse for items imputed in the second stage are given in Figure 19.

Items 17d, 32a-d. These items were all imputed during the internal imputations.

Items 1, 5, and 29. These items are require item for all responding teachers, thus do not require imputation.

Private school teachers

Items 2-4, 6-28, 31, 32f-i, 33-56. The records were sorted by AFFLG/AFFILS/TEALEVEL/GRADELEV/PRVURB/TEAFIELD/PRVENR. The matching variables and their order of collapse for items imputed in the second stage are given in Figure 20. Items are listed in the order in which they were imputed.

Item 32e. The records were sorted by TEALEVEL/school enrollment. The matching variables and their order of collapse for items imputed in the second stage are given in Figure 20.

Items 17d, 32a-d. These items were all imputed during the internal imputations.

Items 1, 5, and 29. These items are require item for all responding teachers, thus do not require imputation.

Figure 1.--SASS-1A items imputed using other data on record

Imputed Item	Source Items
1a-g (This year's enrollment)	1a-g (Proportions from last year's enrollment); or 2a-g (Proportions from this year's FTE teachers); or 2a-g (Proportions from last year's FTE teachers or Last year's student/teacher ratios)
1a-g (Last year's enrollment)	1a-g (Proportions from this year's enrollment); or 2a-g (Proportions from last year's FTE teachers); or 2a-g (Proportions from this year's FTE teachers); or This year's student/teacher ratios
2a-g (This year's FTE teachers)	2a-g (Proportions from last year's FTE teachers); or 1a-g (Proportions from this year's enrollment); or 1a-g (Proportions from last year's enrollment); or Last year's student/teacher ratios
2a-g (Last year's FTE teachers)	2a-g (Proportions from this year's FTE teachers); or 1a-g (Proportions from last year's enrollment); or 1a-g (Proportions from this year's enrollment); or This year's student/teacher ratios.

Figure 2.--SASS-1A Imputation Variables

Variable Name	Description	Values
LEAENR	Number of students in LEA	0 - 990,000
SIZE	Number of students by categories	1 = None 2 = 1-999 3 = 1,000-9,999 4 = 10,000-990,000 5 = Unknown
LEALEVEL	Instructional levels in LEA	1 = Elementary only 2 = Combined, mostly elementary 3 = Combined, comparable elementary and secondary students counts 4 = Combined, mostly secondary 5 = Secondary only 6 = Ungraded
MSC88	Urbanicity (as reported on the 1988 Common Core of Data)	1 = Large central city of SMSA 2 = Medium city of SMSA 3 = Urban fringe of large city 4 = Urban fringe of medium city 5 = Large town, not in SMSA 6 = Small town, not in SMSA 7 = Rural 8 = Unknown
GROUP*	Groups of states with similar LEAs	1 = Connecticut, Rhode Island 2 = Delaware, District of Columbia, Maryland 3 = Maine, New Hampshire, Vermont 4 = Massachusetts, New York 5 = New Jersey, Pennsylvania 6 = Illinois, Indiana 7 = Iowa, Nebraska 8 = Kansas, Oklahoma 9 = Michigan, Ohio 10 = Minnesota, Missouri, Wisconsin 11 = North Dakota, South Dakota 12 = Alabama, Louisiana 13 = Arkansas, Mississippi, West Virginia 14 = Florida, Texas 15 = Georgia, Virginia 16 = Kentucky, South Carolina 17 = North Carolina, Tennessee 18 = Alaska, Wyoming 19 = Arizona, Nevada, Utah 20 = California, Hawaii 21 = Colorado, Washington 22 = Idaho, Montana 23 = New Mexico, Oregon

* The variable GROUP was created because some states (e.g., Hawaii, District of Columbia, Delaware) have few LEAs; combining states made more records of LEAs with similar characteristics available as donor sources.

Figure 3.--SASS-1A Matching Variables and Collapse Order

Items	Matching Variables	Order of Collapse
1-6, 9-12	GROUP, LEALEVEL, MSC88, SIZE	SIZE, MSC88
27	GROUP, LEALEVEL, STATE, MSC88	MSC88, STATE, LEALEVEL

Figure 4.--SASS-1A Matching Variables and Collapse Order

Items	Matching Variables	Order of Collapse
7, 8	GROUP, MSC88, SIZE	SIZE, MSC88
13-17	GROUP, STATE, MSC88, SIZE	SIZE. MSC88, STATE
18-22	GROUP, MSC88, SIZE	SIZE, MSC88
23-26	GROUP, STATE, MSC88, SIZE	SIZE, MSC88, STATE
28, 29	GROUP, MSC88, SIZE	SIZE, MSC88

Figure 5.--SASS-2A/2B Imputation Variables

Variable name	Description	Values
GRDLEVEL	Instructional level of school	1 = Elementary 2 = Combined 3 = Secondary 4 = Not known
COMBSTU	For principals of combined schools, whether school is predominantly elementary or secondary	1 = Elementary 3 = Secondary
PUBURB (SASS-2A only)	Urbanicity of community where public school is located	1 = Large central city of SMSA 2 = Medium city of SMSA 3 = Urban fringe of SMSA 4 = Urban fringe of medium city 5 = Large town, not in SMSA 6 = Small town, not in SMSA 7 = Rural 8 = Not known
PRVURB (SASS-2B only)	Urbanicity of community where private school is located	1 = Urban 2 = Suburban 3 = Rural 4 = Not known
MINEN	Schools' percent minority enrollment	1 = 0-5% 2 = 6-20% 3 = 21-50% 5 = 51-100% 4 = Not known
PUBENR (SASS-2A only)	Public school enrollment size code	1 = 1-299 2 = 300-599 3 = 600-9,000 4 = Not known
PRVENR (SASS-2B only)	Private school enrollment size code	1 = 1-149 2 = 150-399 3 = 400-5,000 4 = Not known

Figure 5.--SASS-2A/2B Imputation Variables (Continued)

Variable name	Description	Values
AFFILS (SASS-2B only)	Schools' religious affiliation and/or association membership	1 = Catholic, parochial 2 = Catholic, diocesan 3 = Catholic, private 4 = Catholic, unclassified 5 = Member of conservative Christian school association 6 = Other schools with religious affiliation or orientation not included in categories 1-5 7 = Religious schools, unknown affiliation/ association 8 = Secular school - regular program 9 = Secular school - special program, vocational or alternative 10 = Secular school - special education 11 = Secular school - unknown program 12 = Unclassified
AFFLG (SASS-2B only)	General affiliation code	1 = Catholic 2 = Other religious affiliation or orientation 3 = Secular 4 = Unclassified
PRVANNSAL (SASS-2B only)	Private school principal's annual salary	1 = 0-12,999 2 = 13,000-17,999 3 = 18,000-21,999 4 = 22,000-28,999 5 = 29,000-32,999 6 = 33,000+ 7 = Not reported
PUBANNSAL (SASS-2A only)	Public school principal's annual salary	1 = 0-35,299 2 = 35,300-38,599 3 = 38,600-41,999 4 = 42,000-46,999 5 = 47,000-53,799
ANNSAL (SASS-2B only)	Principal's annual salary	6 = 53,800+ 7 = Not reported
HIGHDEG	Highest degree received by respondent	1 = No degree reported 2 = Bachelor's 3 = Master's or higher
AGE	Age of respondent	1 = 21-29 2 = 30-45 3 = 45-60 4 = 61-90 5 = Not reported

Figure 5.--SASS-2A/2B Imputation Variables (Continued)

Variable name	Description	Values
YRPRINSC	Years as principal of this school	1 = 0-3 2 = 4-15 3 = 16-30 4 = 31-70
YEARPRIN	Years as principal in all schools	1 = 0-3 2 = 4-15 3 = 16-30 4 = 31+

Figure 6.--SASS-2A Matching Variables and Collapse Order

Items	Matching Variables	Order of Collapse
5a and 5b, where respondent's age was known	STATE, GRDLEVEL, HIGHDEG, AGE	AGE, HIGHDEG, GRDLEVEL
5a and b, where respondent's age was not known	STATE, GRDLEVEL, YEARPRIN	YEARPRIN, HIGHDEG, GRDLEVEL
4b	STATE, GRDLEVEL, YEARPRIN	YEARPRIN, GRDLEVEL
2b, 2c, 3b, 4c, 6	STATE, GRDLEVEL, PUBURB	PUBURB, GRDLEVEL
7, 8, 10	STATE, GRDLEVEL, HIGHDEG, AGE	AGE, HIGHDEG, GRDLEVEL
11	STATE, GRDLEVEL, AGE, PUBURB	PUBURB, AGE, GRDLEVEL
18	STATE, GRDLEVEL, PUBURB, AGE	AGE, PUBURB, GRDLEVEL
19, 20	STATE, GRDLEVEL, PUBURB, MINEN	MINEN, PUBURB, GRDLEVEL
14-17	STATE, GRDLEVEL, PUBURB, YRPRINSC	YRPRINSC, PUBURB
13	STATE, GRDLEVEL, PUBURB, YRPRINSC	YRPRINSC, PUBURB, GRDLEVEL
12a,b (if salary was not known)	STATE, GRDLEVEL, PUBURB, HIGHDEG, YEARPRIN	YEARPRIN, HIGHDEG, PUBURB, GRDLEVEL
12b (if salary was known)	STATE, GRDLEVEL, PUBURB, PUBANNSAL	PUBANNSAL, GRDLEVEL

Figure 7.--SASS-2B Matching Variables and Collapse Order

Items	Matching Variables	Order of Collapse
5a and 5b, where respondent's age was known	AFFILS, GRDLEVEL, HIGHDEG, AGE	AGE, HIGHDEG, GRDLEVEL, AFFILS
5a and 5b, where respondent's age is not known	AFFILS, GRDLEVEL, HIGHDEG, YEARPRIN	YEARPRIN, HIGHDEG, GRDLEVEL, AFFILS
4b	AFFILS, GRDLEVEL, YEARPRIN	YEARPRIN, GRDLEVEL, AFFILS
2b, 2e, 3b, 4c, 6	AFFILS, GRDLEVEL, PRVURB	PRVURB, GRDLEVEL, AFFILS
7, 8, 10	AFFILS, GRDLEVEL, HIGHDEG, AGE	AGE, HIGHDEG, GRDLEVEL, AFFILS
11	AFFILS, GRDLEVEL, AGE, PRVURB	PRVURB, AGE, GRDLEVEL, AFFILS
18	AFFILS, GRDLEVEL, PRVURB, AGE	AGE, PRVURB, GRDLEVEL, AFFILS
19, 20	AFFILS, GRDLEVEL, PRVURB, MINEN	MINEN, PRVURB, GRDLEVEL, AFFILS
14-17	AFFILS, LEVEL, URB, YRSPRINSC	YRSPRINSC, PRVURB, AFFILS
13	AFFILS, GRDLEVEL, PRVURB, YRPRINSC	YRSPRINSC, PRVURB, GRDLEVEL, AFFILS
12 (If salary was not known)	AFFILS, GRDLEVEL, PRVURB, HIGHDEG, YEARPRIN	YEARPRIN, HIGHDEG, PRVURB, GRDLEVEL, AFFILS
12b (if salary was known)	AFFILS, GRDLEVEL, PRVURB, ANNSAL	ANNSAL, PRVURB, GRDLEVEL, AFFILS

Figure 8.--SASS-3A Items Imputed Using The SASS-1A Record

SASS-3A Items	SASS-1A Items
1a	1c-e ("This year's enrollment" column)*
1b	1c-e ("Last year's enrollment" column)*
9a-e	28a-e*
17	1a-g (This years' enrollment)
25a-e	29a-e*

* Indicates LEA data were used only when sample school was the only school in the LEA.

Figure 9.--SASS-3A Imputation Variables

Variable Name	Description	Values
GRDLEVEL	Instructional level of school	1 = Elementary 2 = Combined 3 = Secondary 4 = Unknown
TYPE	Type of school	1 = Regular 2 = Special education 3 = Vocational education 4 = Alternative 5 = Unclassified
PUBURB	Urbanicity of community where school is located	1 = Large central city where school is located of an SMSA 2 = Medium city of an SMSA 3 = Urban fringe of a large city 4 = Urban fringe of a medium city 5 = Large town, not in an SMSA 6 = Small town, not in an SMSA 7 = Rural 8 = Unknown
PUBENR	School enrollment size code	1 = 1-299 2 = 300-599 3 = 600-9,000 4 = Unknown
MINEN	Percent minority enrollment	1 = 0-5% 2 = 6-20% 3 = 21-50% 5 = 51-100% 4 = Unknown
REGION	Census region where school is located	1 = Northeast 2 = Midwest 3 = South 4 = West

Figure 10.--SASS-3A Matching Variables and Collapse Ordering

Items	Matching Variables	Order of Collapse
17	STATE, GRDLEVEL, TYPE	TYPE
1b	STATE, GRDLEVEL	No collapsing
4	STATE, TYPE, GRDLEVEL	GRDLEVEL, TYPE
6, 7	STATE, GRDLEVEL, TYPE	TYPE
10a-h (Yes/No) and 10a, 10b, 10h (Number of students served)	STATE, GRDLEVEL, TYPE	TYPE, GRDLEVEL
10c-g (Number of students served)	STATE, TYPE, GRDLEVEL	GRDLEVEL, TYPE
11	STATE, GRDLEVEL, TYPE	TYPE, GRDLEVEL
12, 13	STATE, GRDLEVEL, TYPE	TYPE
16, 18	STATE, TYPE, GRDLEVEL	GRDLEVEL, TYPE
20-23	STATE, GRDLEVEL, TYPE	TYPE, GRDLEVEL

Figure 11.--SASS-3A Matching Variables and Collapse Ordering

Items	Matching Variables	Order of Collapse
8	STATE, GRDLEVEL, MINEN, PUBURB	PUBURB, MINEN
9	STATE, MINEN, PUBURB	PUBURB, MINEN
14, 15	STATE, MINEN	MINEN
25	STATE, MINEN, PUBURB	PUBURB, MINEN
26	STATE, PUBURB, MINEN	MINEN, PUBURB

Figure 12.--SASS-3A Matching Variables and Collapse Ordering

Items	Matching Variables	Order of Collapse
24, 27, 28	STATE, GRDLEVEL, TYPE	TYPE, GRDLEVEL
29, 30	STATE, TYPE, GRDLEVEL	GRDLEVEL, TYPE
31-33	STATE, GRDLEVEL, TYPE	TYPE, GRDLEVEL
34, 35	STATE, GRDLEVEL	GRDLEVEL

Figure 13.--SASS-3B Imputation Variables

Variable Name	Description	Values
REGION	Census region	1 = Northeast 2 = Midwest 3 = South 4 = West
AFFILR	Specific religious affiliation code	1 = Catholic 2 = Amish 3 = Assembly of God 4 = Baptist 5 = Calvinist 6 = Christian 7 = Church of Christ 8 = Church of God 9 = Disciples of Christ 10 = Episcopal 11 = Friends 12 = Greek Orthodox 13 = Islamic 14 = Jewish 15 = Latter Day Saints 16 = Lutheran 17 = Mennonite 18 = Methodist 19 = Pentecostal 20 = Presbyterian 21 = Seventh-Day Adventist 22 = Other 23 = No religious affiliation 24 = Unknown
AFFLG	General affiliation code	1 = Catholic 2 = Other religious affiliation 3 = No religious affiliation 4= Unknown

Figure 13.--SASS-3B Imputation Variables (Continued)

Variable Name	Description	Values
AFFILS	Religious affiliation and or association membership	1 = Catholic, parochial 2 = Catholic, diocesan 3 = Catholic, private 4 = Catholic, unclassified 5 = Member of conservative Christian school association 6 = Other schools with religious affiliation and/or association membership not included in codes 1-5 7 = Religious schools, unknown affiliation or association 8 = Secular - regular elementary and/or secondary 9 = Secular - special program, vocational, or alternative 10 = Secular - special education 11 = Secular - unknown program 12 = Unclassified
PRVENR	School enrollment size code	1 = 1-149 students 2 = 150-399 3 = 400-5,000 4 = Not known
PRVURB	Urbanicity of community where school is located	1 = Urban 2 = Suburban 3 = Rural 4 = Not known
GRDLEVEL	Instructional level of school	1 = Elementary 2 = Combined 3 = Secondary 4 = Not known
TYPE	School type	1 = Regular 2 = Special education 3 = Vocational education 4 = Alternative 5 = Not known
MINEN	Percent minority enrollment	1 = 0-5% 2 = 6-20% 3 = 21-50% 5 = 51-100% 4 = Not known

Figure 13.--SASS-3B Imputation Variables (Continued)

Variable Name	Description	Values
HIGRADE	Highest grade in school	1 = Ungraded, nursery, prekindergarten, or kindergarten 2 = First grade 3 = Second grade 4 = Third grade 5 = Fourth grade 6 = Fifth grade 7 = Sixth grade 8 = Seventh grade 9 = Eighth grade 10 = Ninth grade 11 = Tenth grade 12 = Eleventh grade 13 = Twelfth grade 14 = Postsecondary 15 = Unclassified

Figure 14.--SASS-3B Matching Variables and Collapse Ordering

Items	Matching Variables	Order of Collapse
2, 27, 28	GRDLEVEL, AFFLG	AFFLG, GRDLEVEL
5	GRDLEVEL, AFFLG, AFFILS	AFFILS, AFFLG
15	GRDLEVEL, TYPE, AFFLG	AFFLG, TYPE, GRDLEVEL
16	GRDLEVEL, AFFLG, AFFILS	AFFILS, AFFLG, GRDLEVEL
22	GRDLEVEL, AFFLG, TYPE, AFFILS	AFFILS, TYPE, AFFLG, GRDLEVEL
23-26	GRDLEVEL, AFFLG, AFFILS	AFFILS, AFFLG, GRDLEVEL
36	AFFLG, AFFILS	AFFILS
39	GRDLEVEL, AFFLG, AFFILS	AFFILS, AFFLG
43-44	GRDLEVEL, AFFLG, AFFILS	AFFILS, AFFLG, GRDLEVEL
35, 41, 42	GRDLEVEL, TYPE, AFFLG, AFFILS	AFFILS, AFFLG, TYPE
32	GRDLEVEL, AFFLG, AFFILS	AFFILS, AFFLG

Figure 15.--SASS-3B Matching Variables and Collapse Ordering

Item	Matching Variables	Order of Collapse
3, 6	AFFLG, AFFILS, TYPE	TYPE, AFFILS
10	AFFLG, TYPE, AFFILS	AFFILS, TYPE
12b	AFFLG, AFFILS	AFFILS
13	AFFLG, AFFILS, TYPE	TYPE
19, 20	AFFLG, TYPE, AFFILS	AFFILS, TYPE
45-48	AFFLG, TYPE, PRVURB, AFFILS	AFFILS, PRVURB
58	AFFLG, AFFILS, TYPE	TYPE, AFFILS

Figure 16.--SASS-3B Matching Variables and Collapse Ordering

Items	Matching Variables	Order of Collapse
4	AFFLG, AFFILS	AFFILS
7, 9	AFFLG, AFFILS, PRVURB	PRVURB, AFFILS
11	PRVURB, AFFLG, AFFILS	AFFILS, AFFLG
18, 29, 30, 33	PRVURB	PRVURB, AFFILS
37	AFFLG, AFFILS, PRVURB, MINEN	MINEN, PRVURB, AFFILS
38, 49-57	AFFLG, AFFILS, PRVURB	PRVURB, AFFILS

Figure 17.--SASS-3B Matching Variables and Collapse Ordering

Items	Matching Variables	Order of Collapse
8, 17, 34, 40	AFFLG, AFFILS, GRDLEVEL	GRDLEVEL, AFFILS

Figure 18.--SASS-4A/4B Imputation Variables

Variable Name	Description	Values
HIGHDEG	Highest degree received	1 = No degree reported by respondent 2 = Bachelor's 3 = Master's or higher
AGE	Age of respondent	1 = Under 30 2 = 30-45 3 = 46-60 4 = 61-91 5 = Not reported
TEAEXPER	Years teaching in all schools	1 = 0-3 2 = 4-15 3 = 16-30 4 = 31-70 5 = Not reported
TEAFIELD	Teaching assignment field	1 = Prekindergarten, kindergarten, or general elementary 2 = Special areas other than foreign language, science, vocational education, and special education 3 = Foreign language 4 = Science 5 = Vocational education 6 = Special education 7 = All others
FULPTIME	Full-time/part-time status	1 = Full-time teacher 2 = Part-time teacher 3 = Not reported
TEALEVEL	Instructional level for teacher	1 = Elementary, prekindergarten and special education 2 = All others

Figure 18.--SASS-4A/4B Imputation Variables (Continued)

Variable Name	Description	Values
STGROUP	Groups of states with similar schools	1 = Connecticut and Rhode Island 2 = Delaware, District of Columbia, Maryland 3 = Maine, New Hampshire, Vermont 4 = Massachusetts, New York 5 = New Jersey, Pennsylvania 6 = Illinois, Indiana 7 = Iowa, Nebraska 8 = Kansas, Oklahoma 9 = Michigan, Ohio 10 = Minnesota, Missouri, Wisconsin 11 = North Dakota, South Dakota 12 = Alabama, Louisiana 13 = Arkansas, Mississippi, West Virginia 14 = Florida, Texas 15 = Georgia, Virginia 16 = Kentucky, South Carolina 17 = North Carolina, Tennessee 18 = Alaska, Wyoming 19 = Arizona, Nevada, Utah 20 = California, Hawaii 21 = Colorado, Washington 22 = Idaho, Montana 23 = New Mexico, Oregon
REGION	Census geographic region where school is located	1 = Northeast 2 = Midwest 3 = South 4 = West
GRADELEV	Grade levels taught this year	1 = Prekindergarten 2 = K-6 3 = K-8 4 = 7-12 5 = Postsecondary 6 = All others
BEGINTEA	Years since beginning first teaching position	1 = 0-3 2 = 4-7 3 = 8-15 4 = 16-24 5 = 25-70
PUBURB (SASS-4A Only)	Urbanicity of community where public school is located	1 = Large central city of SMSA 2 = Medium city of SMSA 3 = Urban fringe of large city 4 = Urban fringe of medium city 5 = Large town, not in SMSA 6 = Small town, not in SMSA 7 = Rural 8 = Unknown

Figure 18.--SASS-4A/4B Imputation Variables (Continued)

Variable Name	Description	Values
PUBENR (SASS-4A Only)	Enrollment size code for public school	1 = 1-299 2 = 300-599 3 = 600-9,000 4= Not known
MINEN	Percent minority enrollment at school	1 = 0-5% 2 = 6-20% 3 = 21-50% 5 = 51-100% 4 = Not known
PRVURB (SASS-4B Only)	Urbanicity of community where private school is located	1 = Urban 2 = Suburban 3 = Rural 4 = Not known
PRVENR (SASS-4B only)	Private school enrollment size code	1 = 1-149 2 = 150-399 3 = 400-5,000 4 = Not known
AFFILS (SASS-4B Only)	Religious affiliation and/or association	1 = Catholic, parochial 2 = Catholic, diocesan 3 = Catholic, private 4 = Catholic, unclassified 5 = Member of conservative Christian school association 6 = Other schools with religious affiliation and/or association membership not included in codes 1-5 7 = Religious schools, unknown affiliation or association 8 = Secular school - regular program 9 = Secular school - special program, vocational, or alternative 10 = Secular - special education 11 = Secular - unknown program 12 = Unclassified
AFFLG (SASS-4B)	General affiliation code for school	1 = Catholic 2 = Other religious affiliation or orientation 3 = Secular 4 = Unclassified

Figure 19.--SASS-4A/B Matching Variables and Collapse Ordering

Items	Matching Variables	Order of Collapse
52	STGROUP, STATE, TEALEVEL, BEGINTEA	BEGINTEA, STATE
17b, 15b, 15e, 16b, 17c	STGROUP, STATE, TEALEVEL, PUBURB	PUBURB, STATE
2, 3, 4	STGROUP, STATE, TEALEVEL, PUBURB, PUBENR	PUBENR, PUBURB, STATE
6, 7, 8	STGROUP, STATE, TEALEVEL, AGE, HIGHDEG	HIGHDEG, AGE, STATE
9	STGROUP, STATE, TEALEVEL, GRADELEV	GRADELEV, STATE
12, 10, 11, 13, 14	STGROUP, STATE, TEALEVEL, AGE, HIGHDEG	HIGHDEG, AGE, STATE
18, 19, 20, 21, 22, 23a, 23b	STGROUP, STATE, TEALEVEL, HIGHDEG, TEAEXPER	TEAEXPER, HIGHDEG, STATE
23c	STGROUP, STATE, TEALEVEL, HIGHDEG, TEAEXPER	TEAEXPER, HIGHDEG, STATE, TEALEVEL
24	STGROUP, STATE, TEALEVEL, HIGHDEG, TEAEXPER	TEAEXPER, HIGHDEG, STATE
25, 26	STGROUP, STATE, TEALEVEL, AGE, HIGHDEG, TEAEXPER	TEAEXPER, HIGHDEG, STATE
27, 28	STGROUP, STATE, TEALEVEL, HIGHDEG, TEAEXPER	TEAEXPER, HIGHDEG, STATE
30, 31, 33, 34	STGROUP, STATE, TEALEVEL, FULPTIME, TEAEXPER	TEAEXPER, FULFPTIME, STATE
35-44	STGROUP, STATE, TEALEVEL, PUBURB, AGE, TEAEXPER	TEAEXPER, AGE, STATE
45-51	STGROUP, STATE, TEALEVEL	TEAEXPER
53-56	PUBURB, HIGHDEG, TEAEXPER	HIGHDEG, STATE
32e	STGROUP, TEALEVEL	TEALEVEL

Figure 19.--SASS-4A/4B Matching Variables and Collapse Ordering (Continued)

Items	Matching Variables	Order of Collapse
32f	STGROUP, PUBENR, TEALEVEL, PUBURB	PUBURB, TEALEVEL, PUBENR
32g	STGROUP, MINEN, PUBENR, TEALEVEL	TEALEVEL, PUBENR, MINEN
32h, 32i	STGROUP, GRADELEV	GRADELEV

Figure 20.--SASS-4A/B Matching Variables and Collapse Ordering

Items	Matching Variables	Order of Collapse
52	AFFILS, TEALEVEL, BEGINTEA	BEGINTEA, AFFILS
17b, 15b, 15e, 16b, 17c	AFFILS, TEALEVEL, PRVURB	PRVURB, AFFILS
2, 3, 4	AFFILS, TEALEVEL, PRVURB, PRVENR	PRVENR, PRVURB, AFFILS
6, 7, 8	AFFILS, TEALEVEL, AGE, HIGHDEG	HIGHDEG, AGE, AFFILS
9	AFFILS, STATE, TEALEVEL, GRADELEV	GRADELEV, AFFILS
12, 10, 11, 13, 14	AFFILS, STATE, TEALEVEL, AGE, HIGHDEG	HIGHDEG, AGE, AFFILS
18, 19, 20, 21, 22, 23a, 23b	AFFILS, STATE, TEALEVEL, HIGHDEG, TEAEXPER	TEAEXPER, HIGHDEG, AFFILS
23c	AFFILS, TEALEVEL, HIGHDEG, TEAEXPER	TEAEXPER, HIGHDEG, TEALEVEL, AFFILS
24	AFFILS, TEALEVEL, HIGHDEG, TEAEXPER	TEAEXPER, HIGHDEG, AFFILS
25, 26	AFFILS, TEALEVEL, AGE, HIGHDEG, TEAEXPER	TEAEXPER, HIGHDEG, AFFILS
27, 28	AFFILS, TEALEVEL, HIGHDEG, TEAEXPER	TEAEXPER, HIGHDEG, AFFILS
30, 31, 33, 34	AFFILS, TEALEVEL, FULPTIME, TEAEXPER	TEAEXPER, FULPTIME, AFFILS
35-44	AFFILS, TEALEVEL, PRVURB, AGE, TEAEXPER	TEAEXPER, AGE., AFFILS
45-51, 53-56	AFFILS, TEALEVEL, PRVURB, HIGHDEG, TEAEXPER	TEAEXPER, HIGDEG, AFFILS
32e	TEALEVEL	TEALEVEL
32f	AFFILS, PRVENR, TEALEVEL, PRVURB	PRVURB, TEALEVEL, PRVENR
32g	AFFILS, MINEN, PRVENR, TEALEVEL	TEALEVEL, PRVENR, MINEN
32h, 32i	AFFILS, GRADELEV	GRADELEV

9. Variance Estimation

Each SASS public use file includes a set of replicate weights designed to produce balanced half-sample replicated variance estimates. As with any replication method, balanced half sample replication involves constructing a number of subsamples (replicates) from the full sample and computing the statistic of interest for each replicate. The mean square error of the replicate estimates around the full sample estimate provides an estimate of the variance of the statistic[16]. The SASS variances are based on 48 half-sample replicates of the full sample. The balanced half-sample technique was used because software to produce such variance estimates is relatively common.

A proprietary computer program (WESVAR), available at Westat, Inc., can be used to estimate these standard errors. The software runs under IBM/OS and VAX/VMS systems. Other programs are available (See Wolter, Appendix E) The formula for the variance of a statistic Y is given below.

$$\text{Variance } (Y) = 1/n \sum_{r} (Y_r - Y)^2$$

where: Y_r is the estimate of Y using the r^{th} set of replicate weights.

n is the number of replicates.

The replicates do not properly reflect the increased precision of 1988 to 1991 change estimates, given the controlled overlap designed into the 1991 sample (see section 4). Conservative variance estimates can be produced by assuming the 1988 and 1991 samples are independent.

Below is a brief description of how the replicates were formed.

9.1 Public School Replicates

The public school file was placed into replicates by first forming 48 variance strata. Each variance stratum contained at least two schools which were alternately divided into two half-samples. To form the variance strata, certainty schools were placed in their own variance stratum where each certainty school was assigned to both half-samples. The noncertainty schools within a state/school level sampling stratum were sorted by the school's order of selection. Pairs of schools were then systematically placed into consecutive variance strata, each element of a pair being assigned to different half-samples.

---

[16] Wolter, K. M.(1985). Introduction to Variance Estimation. New York: Springer-Verlag, chapter 3.

When the 48 variance strata were exhausted, the placement of variance stratum started from the first variance stratum again and continued until all the schools in the sampling stratum have been placed into variance strata. The variance stratum numbering for the next sampling stratum started where the previous sampling stratum left off. When there was an odd number of noncertainty schools within a sampling stratum, one variance stratum was assigned an odd number of cases and adjustments were made to the replicate weights to account for the odd number of cases. After the variance strata were assigned, an orthogonal matrix was used to form the 48 replicates.

9.2 Private School Replicates

For list frame and certainty area frame PSUs, the following was done to form variance stratum half-samples:

> Within each sampling stratum, noncertainty schools were sorted by order of selection. Pairs of schools were then consecutively placed into 48 variance strata, each element of a pair being assigned to different half-samples. If a sampling stratum had an odd number of noncertainty schools, then one of the variance strata had an odd number of schools and an adjustment was made to account for this. Certainty schools were assigned to both half-samples.

For noncertainty area frame PSUs, within each subsample (overlap sample or new sample) and within each PSU stratum, PSUs were sorted by measures of size and then paired into half-samples. These pairings were consecutively assigned to variance strata.

After the variance strata were assigned, an orthogonal matrix was used to form the 48 balanced half-sample replicates.

9.3 Administrator Replicates

The administrator replicates are the same as the school replicates.

9.4 Teacher Replicates

For teachers in noncertainty schools, teacher replicates are the same as school replicates. Teachers from certainty schools were placed into the school variance stratum, but were split into two half-samples.

9.5 LEA Replicates

To reflect the fact that LEAs were selected through the school, it is important to form LEA replicates using the school replicates. An LEA was placed into an LEA

replicate if any of the schools associated with the LEA were in that particular school replicate. Certainty LEAs were placed into all replicates.

LEAs without schools were sorted by order of selection. Pairs of LEAs were then systematically placed into consecutive variance strata and each element of a variance strata was assigned to alternating half-samples. After the variance strata were assigned, an orthogonal matrix was used to form the 48 replicates.

9.6 Replicate Weights

For school, administrator, and teacher replicates, the nonzero replicate basic weights were either: 1) twice the basic weight for noncertainty sample units; or 2) the basic SASS weight for certainty sample units.

The noncertainty LEA's replicate basic weights were assigned using the same LEA weight formula described in the weighting section, the only difference being that, each school stratum's probability of selection for the LEA was divided by two. This is appropriate since the half-sample LEA selection probability within a school stratum is half of the respective full sample probability. Certainty LEA's replicate basic weights were 1.0.

After replicate basic weights were assigned, each replicate was processed through the steps of weighting for each appropriate data file (as described in Section 7). The final weights resulting from the weighting process were the replicate weights.

9.7 Cautions

Replicated variance estimates assume sampling is done with replacement. For SASS, this was not the case. Unless the sampling rate is very high, the variance estimates should be a slight overestimate. None of the public school sampling rates were high enough to provide a large overestimate of variance. However, for some of the small private school association strata, the sampling rates were high. This may lead to a large overestimate of the variance. Thus, in the 1988 SASS, consideration was given to adjusting the private school replicate weights to improve the estimate of the variance. The adjustment considered was the one appropriate for simple random sampling (i.e., 1-n/N where n and N are the sample and universe sizes, respectively).

An analysis was done to see if such an adjustment would improve the variance estimates. The results showed it was unlikely that any private school variance estimate would be greatly reduced by an appropriate finite population correction. For more information on the analysis, see Appendix 4.

10. Frame Evaluation

For private schools, the 1989 Private School Survey (PSS) was the most complete private school universe. Since it was a private census conducted by the SASS staff, there was no definitional difference between SASS and PSS. However, while the preliminary data tapes were being reviewed, some duplicate private schools were found. The schools were called to verify that the schools were duplicates. The weights were then adjusted for the duplication. Due to time constraints, no matching was done to the frame to find duplicates there.

For public schools, the Common Core of Data (CCD) contained the most complete list of public schools in the United States. Nevertheless, some school definitional differences were found between the SASS and the CCD. In some states, intermediate units between LEAs and schools are treated as schools on CCD, while SASS treats each location within each intermediate unit as a school. In California, special education programs are listed on CCD as schools. Los Angeles Special Education Program appeared on CCD as one school record. However, it had 115 locations; and 74 of the 115 were special education programs operating in regular schools listed on CCD. Other special education programs in California had similar idiosyncracies. We obtained from the state of California a universe file of all locations for all special education programs and selected a subsample of locations for each sample program. In Minnesota and Missouri, several schools selected had no corresponding LEA records on the CCD. All of these schools were operated by state agencies. No LEA questionnaires were sent out and no adjustments to the weights were made.

While reviewing the teacher weighting, it was discovered that some of CCD teacher counts used in the numerator of teacher weight first stage ratio adjustment were 10 times smaller than the SASS reported teacher count. Since the problem was worst in Iowa, the Iowa teacher weights were recomputed. This problem may exist in other states, but given the difficulty of identifying these cases and the timing no further reweighting was done. However, there is a research project that will investigate the magnitude of this problem in the other states.

Acknowledgments

The authors would like to thank Karen King and members of her staff who contributed to this paper, as well as, Sharon Fondelier and Pat Healy. We thank Carol Rohr for putting the paper in its final form. We also thank Dan Kasprzyk, Marilyn McMillen, Bob Burton, Shelly Burns, Roslyn Korb and Mark Gorsak for reviewing this document.

Appendices

Appendix 1

Descriptions of the Common Core of Data and the Private School Survey

Common Core of Data:

The Common Core of Data (CCD) is the Center's primary database on elementary and secondary public education in the United States. CCD is a comprehensive, annual, national statistical database of all public elementary and secondary schools and school districts, which contains data comparable across all states.

The objectives of the CCD are twofold. First, it provides an official listing of public elementary and secondary schools and school districts in the nation, which can be used to select samples for other NCES surveys. Second, it provides basic information and descriptive statistics on public elementary and secondary schools and schooling in general.

The Private School Survey

Because of concern about alternatives in education, the interest and need for data on private education has also increased. NCES has recognized this need and has determined that a private elementary and secondary school data collection comparable to the Common Core of Data universe survey for public schools is an NCES priority.

The purposes of this data collection activity are to:

a) build an accurate and complete list of private schools to serve as a sampling frame for NCES surveys of private schools;

b) generate periodic data on the total number of private schools, teachers, and students in the universe (the next survey took place in 1991 - 92); and

c) produce annual early estimates of private school characteristics on a fast-turnaround basis.

Appendix 2

Graphic Explanation of the Effect of Sample Reallocation on the Precision of the State-level Estimates between 1988 SASS and 1991 SASS

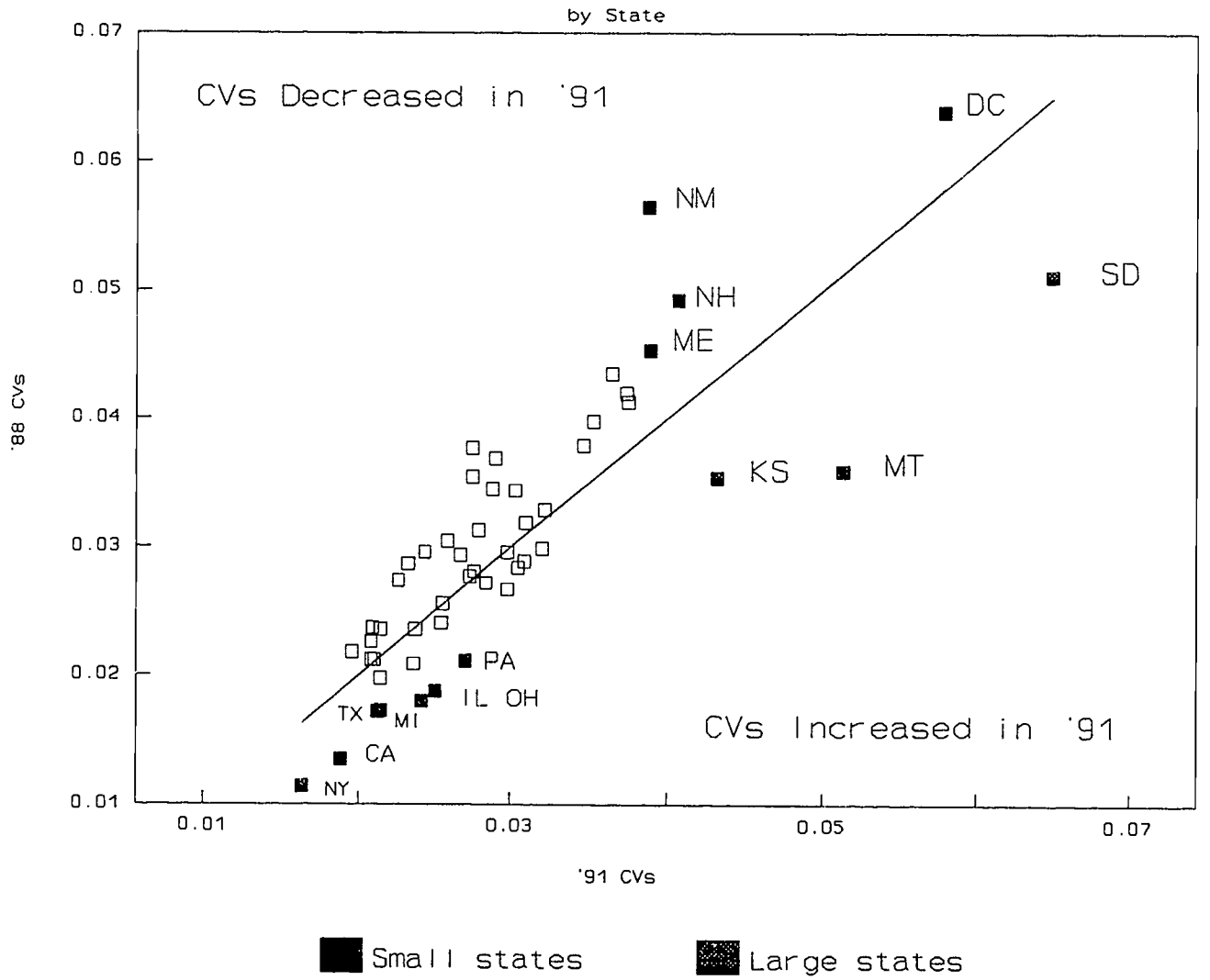
Graphs 1 - 3 and graphs 4 - 6 show graphically the result of the sample allocation with respect to a comparison between the 1988 SASS and the approximate 1991 SASS CVs based on the 1988 CVS adjusted for 91 allocation. Graphs 1 - 3 show the results when estimating total number of teachers, while graphs 4 - 6 show the results for estimating total number of schools.

The line along the diagonal of the graphs represents where the 1988 CVs equal the 1991 CVs. Above this line, 1991 CVs have decreased compared to the 1988 CVs. Below the line, 1991 CVs have increased compared with 1988 CVs. The graphs show how smaller states' estimates have improved at the expense of the larger states. Selected states are highlighted with shaded boxes, to demonstrate the CV differences between large and small states. The small states have solid shades, while the large state have non-solid shades. The point to note is that the CVs for large states increased in 1991, while the small state CVs decreased.

In these graphs, some of the CVs for the smaller states have increased in 1991. In the state graphs, this occurred because of the way the 1991 CVs were approximated. The CVs were approximated using 1988 data, based on the QED frame. The definition of "school" differs on this frame from the definition used on the 1991 frame. Smaller states' CVs increased because in 1988, these states had large numbers of combined schools, which in turn had large CV estimates. In 1991, these states had few combined schools and therefore, fewer selected schools, causing the CV estimates to be even larger than in 1988, since the 1991 CV estimates were based on the 1988 CVs. The increase in the combined strata CVs was large enough to make the 1991 state CVs larger than the 1988 CVs, even though 1991 elementary and secondary CVs have either improved or remained the same. When real CVs are computed, the combined stratum CVs will be much smaller than the 1991 approximations provided here. This in turn will improve the state CVs.

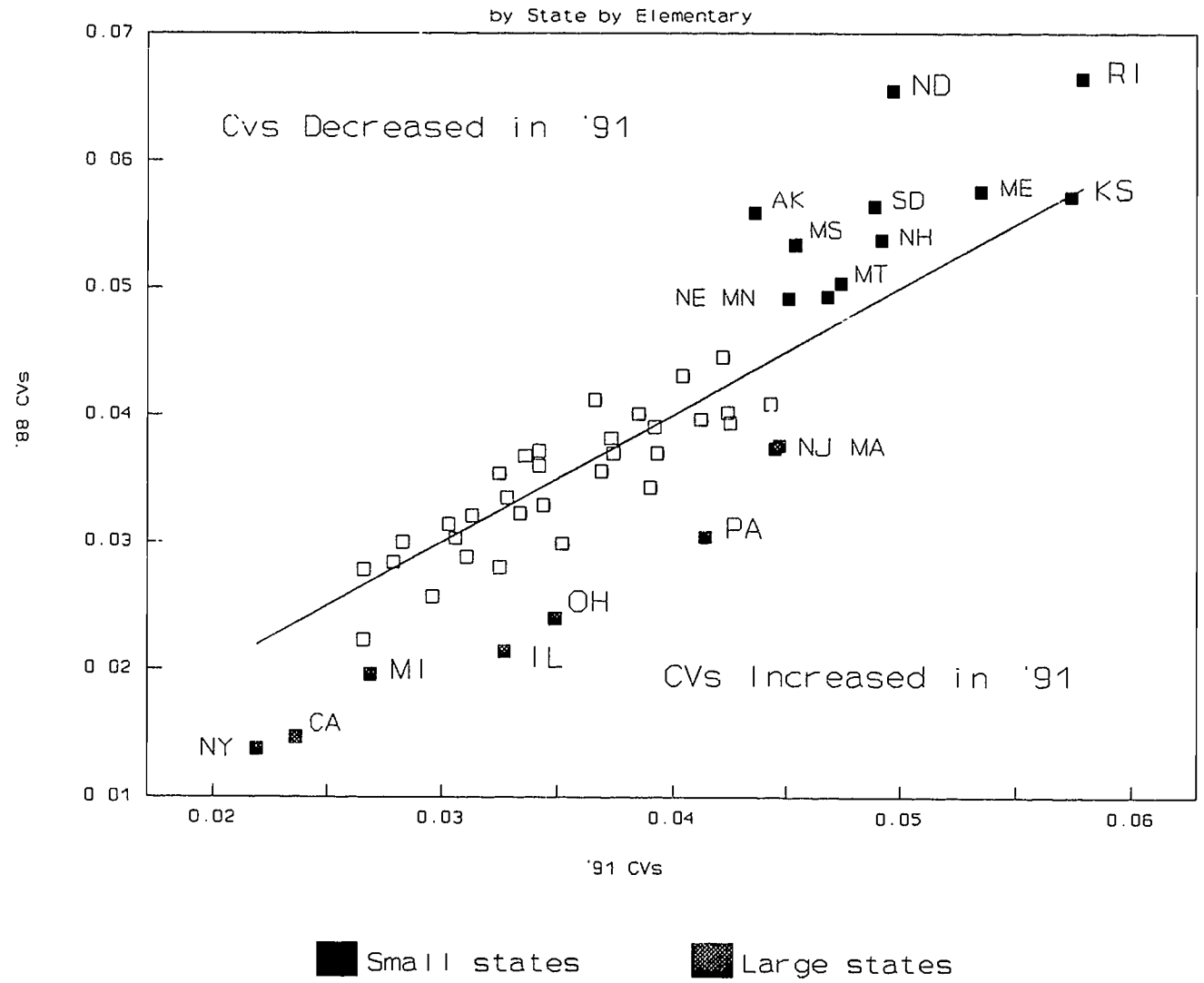
Graph 1

Estimated Number of Teachers CVs



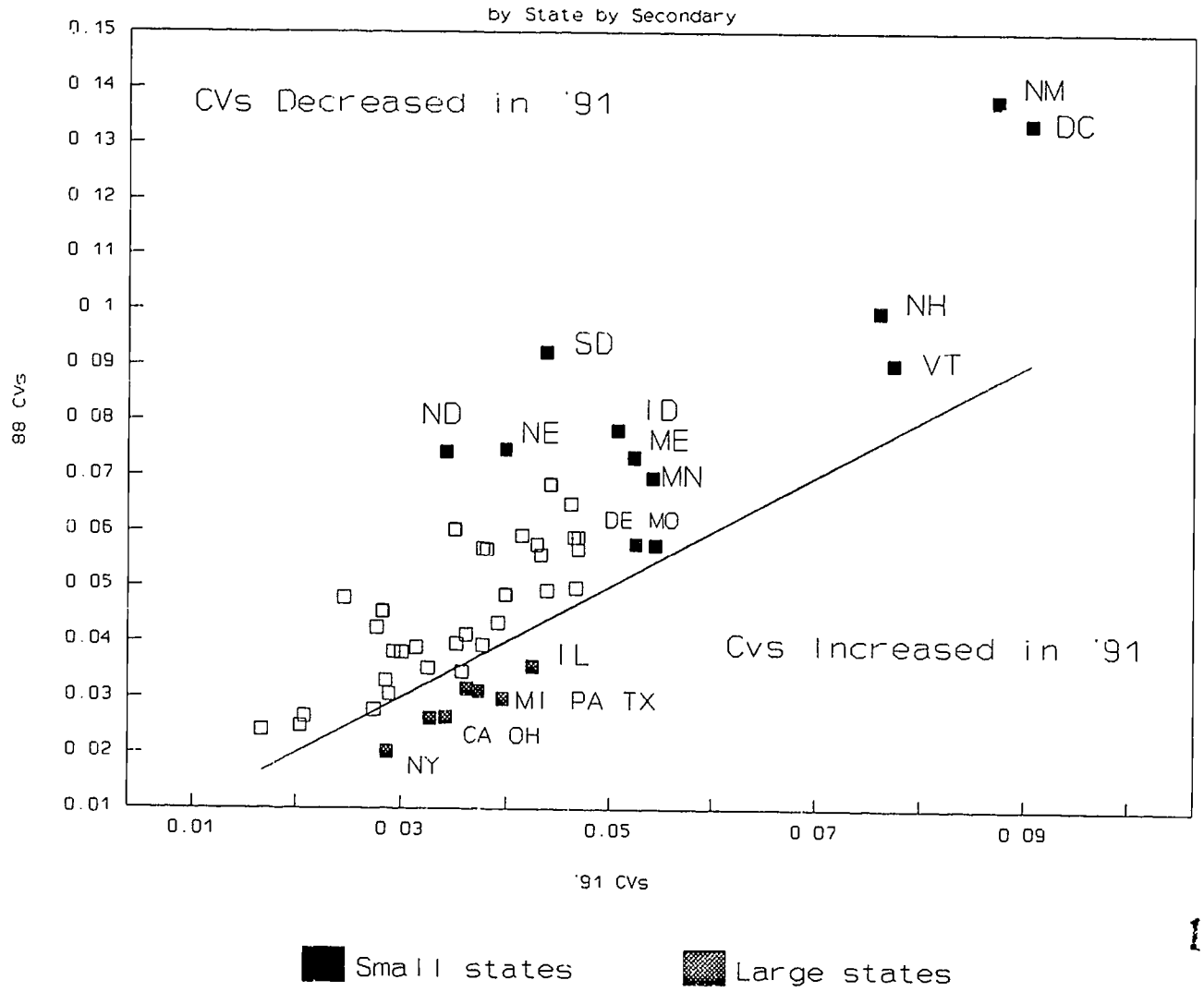
Graph 2

Estimated Number of Teacher CVs



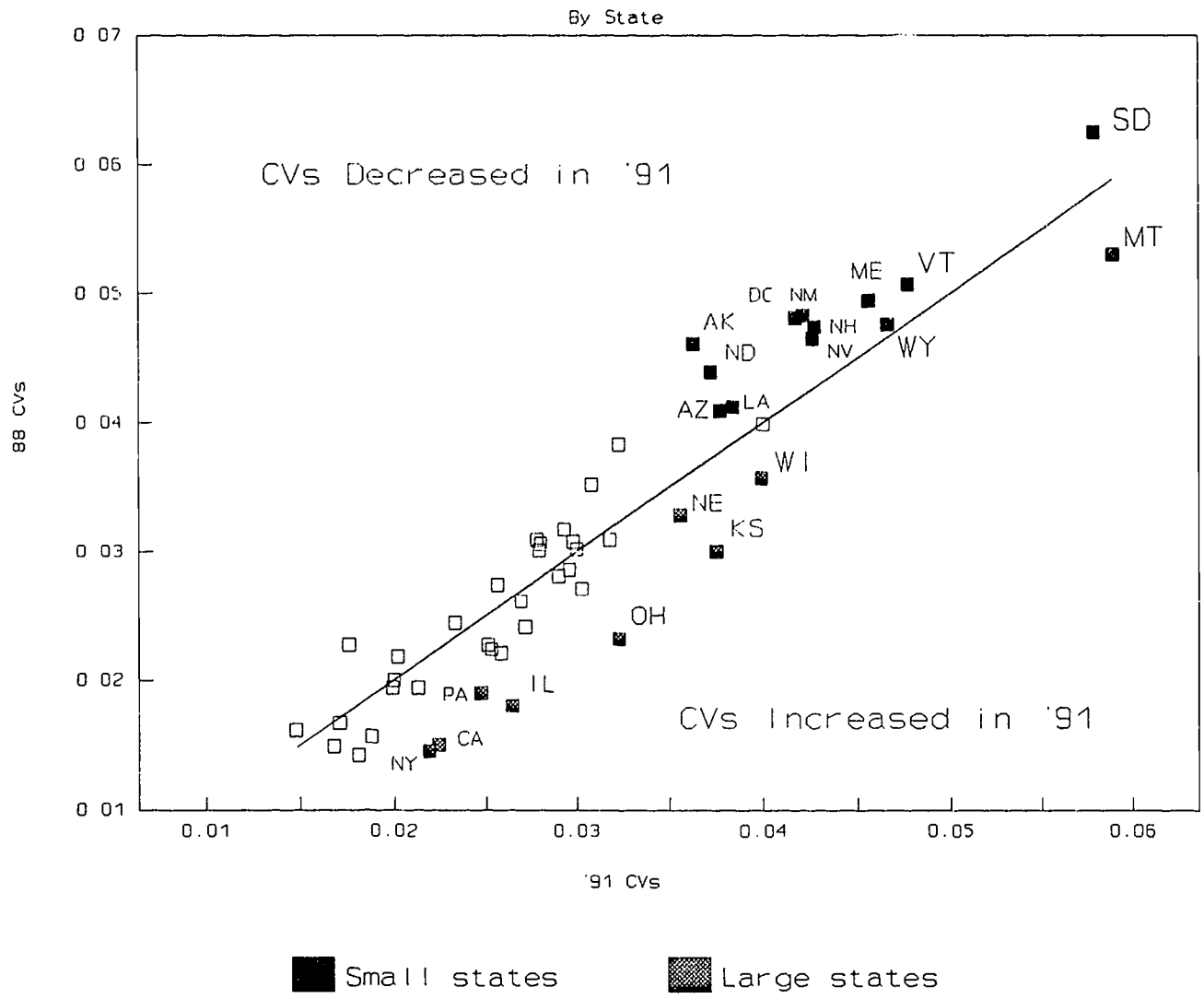
Graph 3

Estimated Number of Teachers CVs



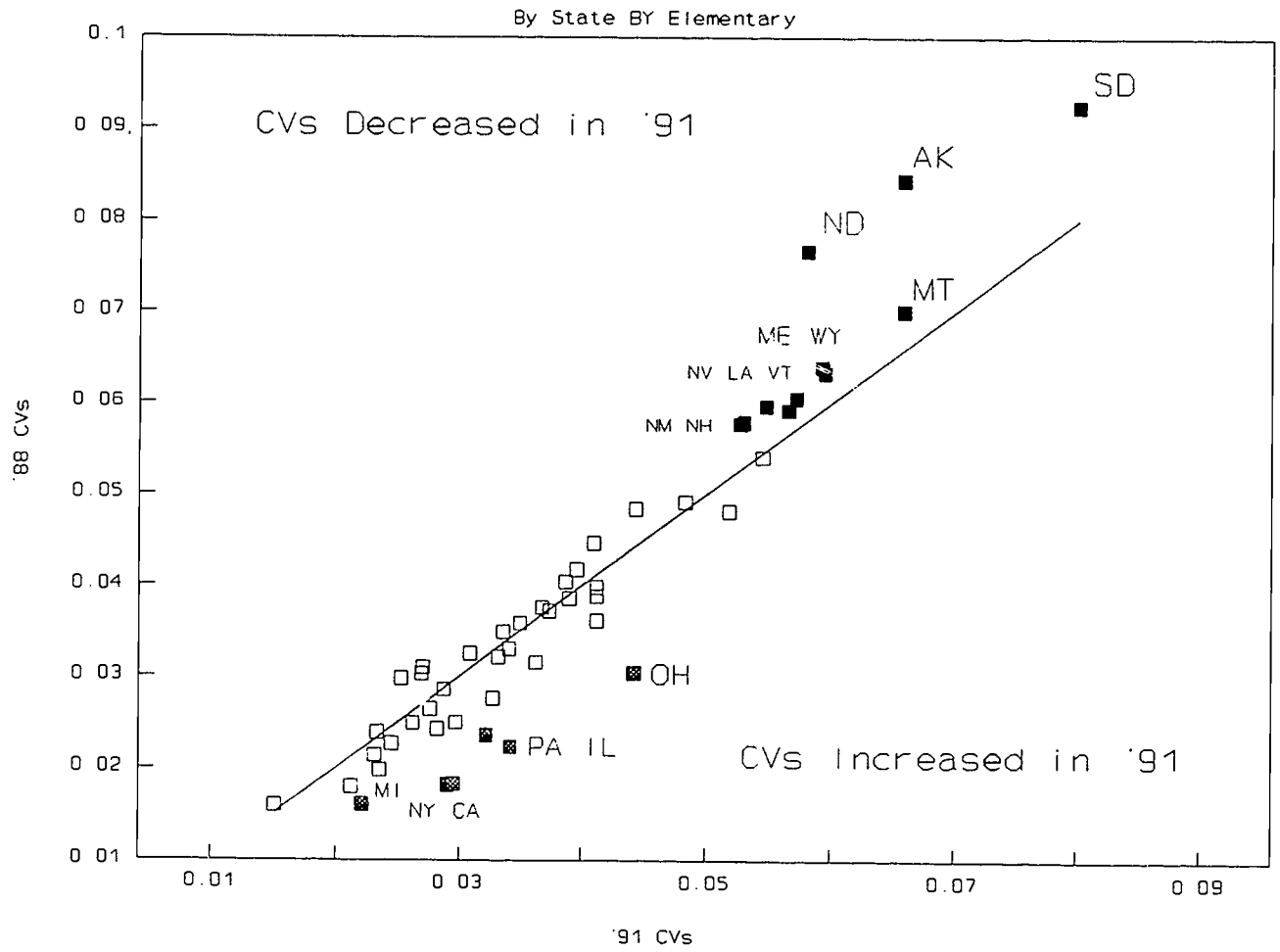
Graph 4

Estimated Number of Schools CVs



Graph 5

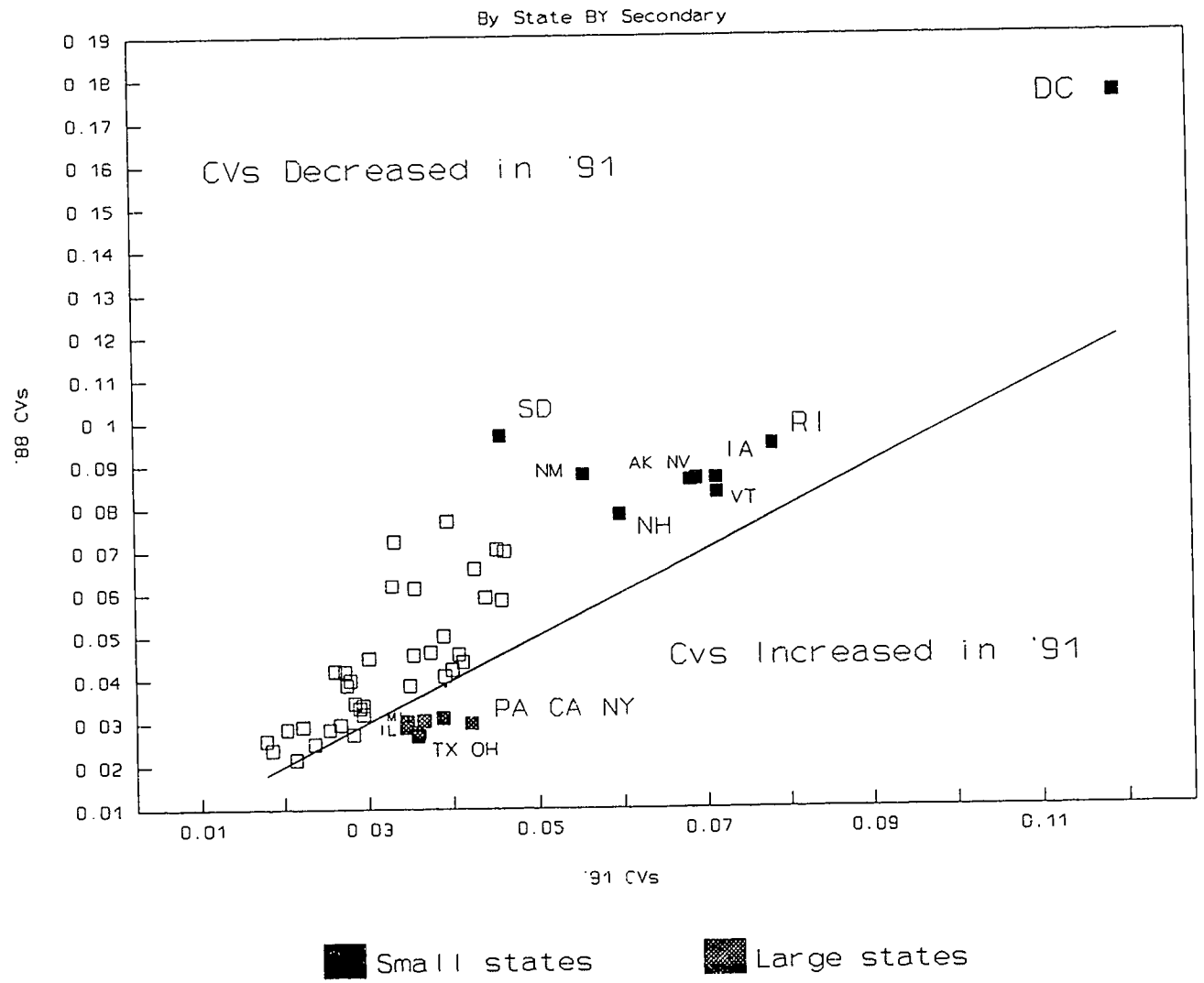
Estimated Number of Schools CVs



Small states Large states

Graph 6

Estimated Number of Schools CVs



Appendix 3

Controlling the School Overlap with the 1988 SASS

This appendix describes how the original 1991 SASS selection probabilities were adjusted so that the expected number of overlap schools between the 1991 and 1988 SASS can be set at a specific level without changing a school's overall selection probability for the 1991 SASS. To do this required knowledge of the 1991 and 1988 SASS selection probabilities for all schools in the frame. The 1991 SASS school sample selection will be dependent upon the 1988 SASS sample.

Since the overall probability of selection was the original 1991 SASS selection probability, the basic weights are the reciprocal of the original 1991 SASS school selection probability.

The details of this process are described below. First, required terminology and sets of schools are defined. Next, the definition of conditional selection probabilities are defined. Selecting the 1991 SASS sample with these conditional probabilities maintains the original 1991 SASS school selection probabilities, while controlling the expected overlap.

Terminology

S_1: 1988 SASS sample

S_2: 1991 SASS sample

i: school

$P_{h'i}(S_1)$: probability of selecting school i from stratum h' in the 1988 SASS.

$P_{hi}(S_2)$: probability of selecting school i from stratum h in the 1991 SASS.

$P_{hi}(S_2 | S_1)$: probability of selecting school i from stratum h in 1991 SASS given that this school was selected for 1988 SASS.

$P_{h'i}(NS_1)$: probability of not selecting school i from stratum h' in 1988 SASS.

$P_{hi}(S_2 | NS_1)$: probability of selecting school i from stratum h in the 1991 SASS given that this school was not selected for the 1988 SASS.

Conditional Selection Probabilities

The 1991 SASS sample was selected using the following conditional selection probabilities:

$$P_{hi}(S_2|S_1) = C_h \quad \text{if} \quad P_{hi}(S_2) \geq P_{\acute{h}i}(S_1) \quad \text{and} \quad P_{\acute{h}i}(S_1) + P_{hi}(S_2) \leq 1$$

$$P_{hi}(S_2|S_1) = C_h \frac{P_{hi}(S_2)}{P_{\acute{h}i}(S_1)} \quad \text{if} \quad P_{hi}(S_2) < P_{\acute{h}i}(S_1) \quad \text{and} \quad P_{\acute{h}i}(S_1) + P_{hi}(S_2) \leq 1$$

$$P_{hi}(S_2|S_1) = \frac{P_{\acute{h}i}(S_1) + P_{hi}(S_2) - 1}{P_{\acute{h}i}(S_1)} \quad \text{if} \quad P_{\acute{h}i}(S_1) + P_{hi}(S_2) > 1$$

$$P_{hi}(S_2|NS_1) = \frac{P_{hi}(S_2) - P_{\acute{h}i}(S_1)C_h}{1 - P_{\acute{h}i}(S_1)} \quad \text{if} \quad P_{hi}(S_2) \geq P_{\acute{h}i}(S_1) \text{ and } P_{\acute{h}i}(S_1) + P_{hi}(S_2) \leq 1$$

$$P_{hi}(S_2|NS_1) = \frac{P_{hi}(S_2)(1 - C_h)}{1 - P_{\acute{h}i}(S_1)} \quad \text{if} \quad P_{hi}(S_2) < P_{\acute{h}i}(S_1) \text{ and } P_{\acute{h}i}(S_1) + P_{hi}(S_2) \leq 1$$

$$P_{hi}(S_2|NS_1) = 1 \quad \text{if} \quad P_{\acute{h}i}(S_1) + P_{hi}(S_2) > 1$$

where $$C_h = \frac{M_h - M_{sh}}{M_{rh}}$$

$$M_{sh} = \sum_{i \in P_{\acute{h}i}(S_1) + P_{hi}(S_2) > 1} P_{hi}(S_2/S_1)$$

$$M_{rh} = \sum_{i \in P_{hi}(S_1) + P_{\acute{h}i}(S_2) \leq 1} P_{hi}(S_2/S_1)$$

M_h is the expected overlap sample size for stratum h.

It can be verified that these conditional selection probabilities will preserve the original 1991 SASS selection probabilities, $P_{hi}(S_2)$, while the expected overlap between 1991 SASS schools and 1988 SASS schools is equal to M_h. M_h's were chosen based on the following percentage of expected overlap in table 14 below:

Table 14.--Expected and actual school overlap from 1988 and 1991 by Association

Public Schools:	30%	
Private Schools:		
Association	**Expected Overlap**	**Actual Overlap**
01 Military Schools	100%	100%
02 Catholic	30%	31%
03 Friends	100%	100%
04 Episcopal	10%	12%
05 National Hebrew Day	6%	11%
06 Solomon Schechter	100%	100%
07 Other Jewish	2%	3%
08 Lutheran - Missouri Synod	30%	32%
09 Lutheran - Wisconsin Synod	30%	30%
10 Evangelical Lutheran Church	100%	100%
11 Other Lutheran	30%	27%
12 Seventh-Day Adventist	30%	30%
13 Christian Schools International	26%	33%
14 American Association of Christian Schools	0%	1%
15 National Association of Private Schools for Exceptional Children	22%	26%
16 Montessori	4%	6%
17 National Association of Independent Schools	1.5%	.5%
18 All Else	.6%	1%

* See Table 10 for the expected and actual overlap sample sizes.

Appendix 4

Effect of a Finite Population Correction on SASS Variance Estimates based on 1988 SASS

Replicate variance estimates assume sampling is done with replacement. For SASS, this was not the case. SASS variance estimates should only be a slight overestimate, unless the sampling rate is very high. None of the public school sampling rates were high enough to provide a overestimate of the variance. However, for some of the small private school affiliation strata, the sampling rates were high enough to possibly produce a large overestimate of the variance. For this reason, consideration was given to adjusting the private school replicate weights to improve variance estimates. The adjustment considered was the one appropriate for simple random sampling (i.e., 1-n/N, where n and N are the sample and universe sizes, respectively).

Some simple simulations were performed to measure the effect of such an adjustment when the sample was selected with probability proportionate to size. First, a school frame was assumed. Then, all possible school samples were generated and an estimate of total schools was produced for each sample. In each sample, 66 percent of the schools were selected. From this, the true standard error was computed.

Next, all possible replicate standard error estimates were computed for each sample. These standard error estimates were averaged to produce an average replicated standard error estimate.

Finally, the average replicated estimate was adjusted appropriately to reflect a simple random sample finite population correction factor (adjusted replicated estimate). With this information, it was possible to measure the error in both the unadjusted and adjusted replicated standard error estimates.

The three tables below show the results:

Table 15.--Effect of finite population correction (fpc) where the distribution of probabilities is unequal and skewed

School Frame			
Elementary		Secondary	
School	Probability	School	Probability
1	0.9	4	0.9
2	0.9	5	0.9
3	0.2	6	0.2

Two out of three schools were selected within each stratum using a probability proportionate to size sampling scheme. The probability column above provides the measures of size.

From the set of all possible samples, the following numbers can be computed.

(A)	(B)	
True Standard error: 2.199887	Average replicate Standard error: 2.459549	Ratio of (B) to (A) 1.118033

Unadjusted, the replicated standard error overestimates the true standard error by 12 percent. If the replicated standard error is adjusted by the square root of (1-n/N), where n and N are the sample and universe sizes, respectively, then the following numbers are obtained.

(C)	
FPC adjusted replicate standard error: 1.420021	Ratio of (C) to (A) 0.645497

Now, the adjusted replicated standard error is underestimating the standard error by 35 percent.

Table 16.--Effect of finite population correction (fpc) where the distribution of probabilities is unequal and not skewed

School Frame			
Elementary		Secondary	
School	Probability	School	Probability
1	0.5	4	0.5
2	0.7	5	0.7
3	0.8	6	0.8

Two out of three schools were selected within each stratum using a probability proportionate to size sampling scheme. The probability column above provides the measures of size.

From the set of all possible samples, the following numbers can be computed.

(A)	(B)	
True Standard error: 0.462910	Average replicate Standard error: 0.707106	Ratio of (B) to (A) 1.527525

Unadjusted, the replicated standard error overestimates the true standard error by 53 percent. If the replicated standard error is adjusted by the square root of (1-n/N), where n and N are the sample and universe sizes, respectively, then the following numbers can be computed.

(C)	
FPC adjusted replicate standard error: 0.408248	Ratio of (C) to (A) 0.881917

Now, the adjusted replicated standard error is underestimating the true standard error by 12 percent.

Table 17.--Effect of finite population correction (fpc) where the distribution of probabilities is almost equal and not skewed

School Frame			
Elementary		Secondary	
School	Probability	School	Probability
1	0.64	4	0.64
2	0.67	5	0.67
3	0.69	6	0.69

Two out of three schools were selected within each stratum using a probability proportionate to size sampling scheme. The probability column above provides the measures of size.

For the set of all possible samples, the following numbers can be computed.

(A)	(B)	
True Standard error: 0.066387	Average replicate Standard error: 0.113328	Ratio of (B) to (A) 1.707076

Unadjusted, the replicated standard error overestimates the true standard error by 71 percent. If the replicated standard error is adjusted by the square root of (1-n/N), where n and N are the sample and universe sizes, respectively, then the following numbers can be computed.

(C)	
FPC adjusted replicate standard error: 0.065430	Ratio of (C) to (A) 0.985581

Now, the adjusted replicated standard error is underestimating the standard error by 1 percent.

The tables show that, when the sampling rates are high, the adjusted standard errors underestimate the standard error, by as much as 35.5 percent (1-.645), if the selection probabilities are unequal and skewed. Since the underestimate can be so large, it is probably unwise applying an fpc adjustment.

Another argument against applying an fpc adjustment is the effect the area frame sampling rate has on the overall sampling rate. The 551 area sample schools represent 24 percent of the total number of private schools. These schools come from a sample of 75 out of 2,497 PSUs, a sampling rate of 3 percent. On average, one might expect that 24 percent of an affiliation's estimate would come from the area sample, with a 3 percent sampling rate. This should significantly reduce the overall sampling rate of affiliations with high list frame sampling rates. A similar situation arises if schools from other list frame strata with low sampling rates, report belonging to an affiliation with a high list frame sampling rate.

Table 18 below shows the impact of the area sample on the affiliation estimates with the highest list frame sampling rates.

This table also shows that except for Friends, military schools and Christian International schools, a significant part of the standard error comes from the area sample, which has a low overall sampling rate. Therefore, adjusting for a high list frame sampling rate should have minimal impact on the total variance. The high variance for Friends, military schools and Christian International schools is caused by schools from other list frame strata reporting they belong to Friends, military school or Christian International schools. The variance contribution from the Friends and Military schools list frame stratum is zero since they were all selected with certainty. The variance is solely coming from schools from other list strata that make up part of the estimate.

Because the overall affiliation sampling rate is greatly reduced by the low sampling rate of the area frame and some other list strata, and the fact that the proposed adjustment would produce a possibly large underestimate of the variance, the variance replicates were not adjusted for the high sampling rates. It is unlikely that any private school estimate would be greatly reduced by an appropriate finite population correction.

Table 18.--Estimate of number of schools by list and area stratum from 1988 SASS

School	List frame		Area Frame	
	Estimate	Standard Error	Estimate	Standard Error
Friends	73	11	2	2
Military Schools	51	19	0	0
Exceptional Children	199	44	109	52
Christian International	308	49	3	3
Episcopal	319	31	27	5
Montessori	467	54	216	106
Jewish	514	23	90	52

United States
Department of Education
Washington, D.C. 20208–5651

Official Business
Penalty for Private Use, $300

Postage and Fees Paid
U.S. Department of Education
Permit No. G–17

FOURTH CLASS BOOK RATE