NATIONAL CENTER FOR EDUCATION STATISTICS

Design Effects and Generalized Variance Functions for the 1990–91 Schools and Staffing Survey (SASS)

U.S. Department of Education
Office of Educational Research and Improvement NCES 95-342-I

NATIONAL CENTER FOR EDUCATION STATISTICS

Design Effects and Generalized Variance Functions for the 1990-91 Schools and Staffing Survey (SASS)

Sameena Salvucci
Stanley Weng
Synectics for Management Decision, Inc.

Steven Kaufman, Project Officer
National Center for Education Statistics

U.S. Department of Education
Office of Educational Research and Improvement **NCES 95-342-1**

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

Office of Educational Research and Improvement
Sharon P. Robinson
Assistant Secretary

National Center for Education Statistics
Jeanne E. Griffith
Acting Commissioner

National Center for Education Statistics

The purpose of the Center is to collect and report "statistics and information showing the condition and progress of education in the United States and other nations in order to promote and accelerate the improvement of American education."—Section 402(b) of the National Education Statistics Act of 1994 (20 U.S.C. 9001).

February 1995

For a complete copy of the report (#NCES 95-342-1) contact:
National Data Resource Center
(703) 845-3151

For technical questions contact:
Steven Kaufman
E-mail: steve_kaufman @ed.gov

CONTENTS

ACKNOWLEDGMENTS

This series of reports was prepared by Synectics for Management Decisions, Inc., a contractor to the National Center for Education Statistics, under Contract No. RN-91-0600.01.

The authors wish to thank all of those who contributed to the production of this report. Among Synectics staff, special mention goes to Fan Zhang, Mehrdad Saba, Michael Chang, and Nagarthun Movva who provided very capable programming support; Mayra Walker and Steven Fink who prepared the tables; and Elizabeth Walter, who carefully edited the report.

Steven Kaufman and Dan Kasprzyk, of the Elementary and Secondary Education Statistics Division, reviewed the manuscript through several drafts. We would also like to acknowledge the helpful comments of the following technical reviewers: Marilyn McMillen, Steve Broughman, Michael P. Cohen, Mary Rollefson, Randall J. Parmer, and Betty Jackson.

PREFACE

This user's manual summarizes the results and use of design effects and generalized variance functions to approximate standard errors for the 1990-91 Schools and Staffing Survey (SASS). It is Volume I of a two-volume publication that is part of the Technical Report Series published by the National Center for Education Statistics (NCES). Volume II is intended as a technical report describing the concept, methodology, and calculation/modeling of design effects and generalized variance functions (Salvucci et al. 1995). Users who are interested in knowing more about the background and methodological issues are referred to Volume II, the technical report. The methodological descriptions in Volume II, though not necessary for using this manual, would be very helpful for users to reach a better understanding of the methods and hence their use as illustrated by this manual.

EXECUTIVE SUMMARY

The Schools and Staffing Survey (SASS) is a periodic, integrated system of sample surveys conducted by the National Center for Education Statistics (NCES) of the U.S. Department of Education. The complex sample design of SASS produces sampling variances different from those produced by *simple random sampling* (srs) with fixed sample size. This is so for a number of reasons. There are gains in precision from stratification by geography, type of school, size of school, and so on. These gains, however, are counterbalanced by the effects of clustering of students and teachers within sampled schools. Weighting can be conducted to determine the contribution of sample units to the population estimates. However, the weights themselves are subject to sampling variability which may make nonlinear the statistics which are linear with simple random sampling. The calculation of variance estimates for SASS statistics are, therefore, more complex than the simple random sample variance estimation algorithms and computationally more expensive. Using the simple random sample methods for SASS complex samples almost always underestimates the true sampling variances and makes differences in the estimates appear to be significant when they are not. Unfortunately, general use statistical packages such as SAS, SPSS, etc., only calculate sampling variances based on simple random sample and are thus not appropriate for estimating variances for SASS.

This manual introduces two general techniques: the *design effect* and the *generalized variance function* (GVF), for estimating sampling variances for complex surveys such as SASS. These techniques differ from the direct estimation methods which either use point variance estimators or conduct replication procedures to obtain variance estimates individually for survey statistics. These general techniques use generalized analytical approaches, applied to groups of survey estimates, to produce complex sample variance estimates, for a variety of survey statistics, from srs variance estimates or from survey estimates themselves. The Introduction section of Volume II of this publication describes the rationale for developing and employing such general techniques.

The average design effect and GVF tables provided with this manual (appendix II and appendix III) are products of an empirical study as reported in Volume II of this publication. They can be used as alternatives to direct variance estimation for SASS, in particular, when appropriate statistical software is not available to conduct the balanced half-sample replication method (see section 1.3, Volume II) using the replicates provided on each SASS public use file (Kaufman and Huang 1993, Gruber et al. 1994). Generalized variance functions have been shown in some data settings to perform as well or better than direct variance estimators in terms of bias, precision, and confidence interval construction (Valliant 1987). The performance of the GVFs generally depends on the critical issue of selection of a set of survey variables for GVF modeling, the type of GVF model chosen including the method of estimating the parameters of the GVF model. A cautionary note is that there are likely to be survey variables (e.g., estimate of rare characteristics) whose GVF model differs considerably from that of most variables and for which GVFs will give poor results. Section 3.4 provides a list of specific types of variables in SASS for which GVFs may be inappropriate.

NCES has recently issued guidelines on recommended technical approaches for performing analysis on NCES survey data (Ahmed 1993b). The guidelines describe two categories of procedures and their order of preference. First, the preferred procedure is to use a program designed specifically for analyzing data from complex surveys, such as WESVAR/WESREG (Westat 1993), SUDAAN (Shah et al. 1992), and VPLX/CPLX (Fay 1995) to compute standard errors. Second, an alternative but acceptable procedure is to use a standard statistical package such as SAS or SPSS and a design effect correction to the standard error. The method of using generalized variance functions can be considered in the same category of alternative procedures as the design effect correction. When using the alternative procedures, choosing between design effect and GVF depends on the circumstances of the particular data analysis. Therefore, no general recommendation on using one or the other may be made here. These points will be made clearer in section 3.1 after discussion of the examples.

1. Overview

The purpose of this volume is to illustrate clearly the application of the two techniques, using the tables provided in this manual, to approximate variance estimates or standard errors for SASS. Following this overview, we first give a brief description of the SASS data (sections 1.1 and 1.2); then a conceptual introduction of the estimation and use of standard errors with complex survey data (section 1.3 through 1.5); and finally a description of the grouping of statistics regarding the structure of the tables provided with this manual (section 1.6). Sections 2 and 3 provide a brief review and a how-to guide on the use of the design effect tables and generalized variance function tables, respectively. For a more detailed methodological discussion of these techniques, users are referred to Volume II, section 3, Design Effect Methodology, and section 4, GVF Methodology, of this publication.

1.1 Source of Data

The data were collected in the second cycle of the Schools and Staffing Survey (SASS) conducted by the National Center for Education Statistics (NCES) in 1990-91. SASS provides data on public and private schools, public school districts, teachers, and administrators, and is used by educators, researchers, and policy makers. The survey includes several types of respondents: school district personnel, public school principals, private school principals, public school teachers, and private school teachers, among others. The 1990-91 SASS is a set of four interrelated national surveys.

The following elements make up the 1990-91 SASS:

a. The Teacher Demand and Shortage (TDS) Survey targeted public school district personnel who provided information about their district's student enrollment, number of teachers, position vacancies, new hires, teacher salaries and incentives, and hiring and retirement policies.

b. The School Administrator Survey collected background information from principals on their education, experience, and compensation and also asked about their perceptions of the school environment and the importance they placed on various educational goals.

c. The School Survey included information on student characteristics, staffing patterns, student-teacher ratios, types of programs and services offered, length of school day and school year, graduation and college application rates, and teacher turnover rates. The 1990-91 private school questionnaire incorporated questions on aggregate demand for both new and continuing teachers.

d. The Teacher Survey collected information on public and private school teachers' demographic characteristics, education, qualifications, income sources, working conditions, plans for the future, and perceptions of the school environment and the teaching profession.

1.2 Sample Design

The target populations for the 1990-91 SASS surveys included U.S. elementary and secondary public and private schools with students in any of grades 1-12, principals and classroom teachers in those schools, and local education agencies (LEAs) that employed elementary and/or secondary level teachers. In the private sector, since there is no counterpart to the LEAs, information on teacher demand and shortages was collected directly from individual schools. The sample was designed to produce 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, and 4) detailed association estimates and grade level estimates for private schools.

These are the three primary steps in the sample selection process followed during the 1990-91 SASS:

(1) A sample of schools was selected. The same sample was used for the School Administrator Survey. For the sample of private schools, the questions for the

Teacher Demand and Shortage Survey were included in the questionnaire for the School Survey.

(2) Each LEA that administered one or more of the sample schools in the public sector became part of the sample for the Teacher Demand and Shortage Survey.

(3) For each sample school, a list of teachers was obtained from which a sample was selected for inclusion in the Teacher Survey.

Details pertaining to the frame, stratification, sorting, and sample selection for each of the four surveys of SASS are described in the sections below (Kaufman and Huang 1993).

1.2 School Survey

The School Survey had two components: private schools and public schools. The primary frame for the public school sample was the 1988-89 Common Core of Data (CCD) file. The CCD survey includes an annual census of public schools, obtained from the states, with information on school characteristics and size. A supplemental frame was obtained from the Bureau of Indian Affairs, containing a list of tribal schools and schools operated by that agency. The school sample was stratified, with the allocation of sample schools among the strata designed to provide estimates for several analytical domains. Within each stratum, the schools in the frame were further sorted on several geographic and other characteristics. A specified number of schools were selected from each stratum with probability proportionate to the square root of the number of teachers as reported on the CCD file. The target sample size of public schools was 9,687.

A dual frame approach was used to select the samples of private schools. A list frame was the primary private school frame, and an area frame was used to find schools missing from the list frame, thereby compensating for the coverage problems of the list frame. To supplement the list frame, an area sample consisting of 123 primary sampling units (PSUs) was selected. The target sample size of private schools was 3,270, with 2,670 allocated to the list sample and 600 to the area sample. The list sample was allocated to 216 strata defined by association group, school level (elementary, secondary, combined), and census region (northeast, midwest, south, west). There were 18 association groups; for example, Catholic, National Society of Hebrew Day Schools, and National Association of Independent Schools. Within each stratum, schools were sorted by state and other variables within state. The area sample was allocated to strata defined by 123 PSUs and school level (elementary, secondary, combined). Within each stratum, schools were sorted by affiliation (Catholic, other religious, and nonsectarian), 1989 PSS enrollment, and school name. For both the list sample and the area sample, schools were systematically selected from each stratum with probability proportionate to the square root of the number of teachers as reported in the 1989-90 PSS. 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.

School Administrator Survey.

For the School Administrator Survey the target population consisted of the administrators of all public and private schools eligible for inclusion in the School Survey. Once the sample of schools was selected, no additional sampling was needed to select the sample of school administrators. Thus, the target sample size was the same as for the School Survey (n=12,957). Some of these schools did not have administrators, in which case the school was asked to return the questionnaire, but, with few exceptions, there was a one-to-one correspondence between the SASS samples of schools and school administrators.

Teacher Demand and Shortage Survey

The Teacher Demand and Shortage (TDS) Survey had two components: public schools and private schools.

For the public school sector, the target population consisted of all U.S. public school districts. These public school districts, often called local education agencies (LEAs), are government agencies administratively responsible for providing public elementary and/or secondary education. LEAs associated with the selected schools in the school sample received a TDS questionnaire. An additional sample of districts not associated with schools was selected and also received the TDS questionnaire. The target sample size was 5,424.

For the private school sector, the target population consisted of all U.S. private schools. Thus, the target sample size was the same as the private school sample of 3,270. The school questionnaire for the selected private schools included TDS questions for the school.

Teacher Survey

The target population for the Teacher Survey consisted of full-time and part-time teachers whose primary assignment was teaching in kindergarten through grade 12 (K-12). Data were collected from a sample of classroom teachers in each of the public and private schools that was included in the sample for the School Survey: the selected schools were asked to provide teacher lists for their schools and then those lists were used to select 56,051 public and 9,166 private school teachers. The survey designs for the public and private sectors were very similar. Within each selected school, teachers were stratified into one of five types in hierarchical order, as 1) Asian or Pacific Islander, 2) American Indian, Aleut, or Eskimo, 3) Bilingual/ESL (English as a Second Language), 4) New (less than three years teaching experience), or 5) Experienced (three or more years of teaching experience). Within each stratum, teachers were selected systematically with equal probability.

1.3 Accuracy of Estimates

SASS estimates are based on a sample; they may differ somewhat from the figures that would have been obtained if a complete census had been taken using the same questionnaire, instructions, and data collection procedure. There are two types of *errors* possible with an estimate based on a survey sample: *nonsampling errors* and *sampling errors*. We can provide estimates of the magnitude of SASS sampling errors, but not for nonsampling errors. The following of this section describes sources of nonsampling and sampling errors. The next sections describe sources of SASS nonsampling errors, followed by a discussion of sampling errors, their estimation, and their use in data analysis.

Nonsampling variability

Nonsampling errors can be attributed to many sources; e.g., inability to obtain information about all cases in the sample, definitional difficulties, differences in the interpretation of questions, inability or unwillingness on the part of the respondents to provide correct information, inability to recall information, errors made in collection such as in recording or coding the data, errors made in processing the data, errors made in estimating values for missing data, biases resulting from the differing recall periods caused by the interviewing pattern used, and failure of all units in the universe to have some probability of being selected for the sample (undercoverage). Quality control and edit procedures were used to reduce errors made by respondents, coders, and interviewers. For a further discussion, see *SASS Quality Profile* (Jabine 1994).

Undercoverage in SASS results from missed schools and from missed principals and teachers within sample schools. NCES used complex techniques to adjust the weights for nonresponse; the success of these techniques in avoiding bias has been examined (Synectics 1995).

Sampling Variability

Sampling errors are attributed to sampling variation; i.e., the variation that occurs by chance, because a sample, rather than a population, is surveyed. The sampling errors also partially measure the effect of some nonsampling errors in response and enumeration, but do not measure any systematic biases in the data. The reliability of an estimate is usually described in terms of a standard error (the square root of the estimated variance) that is primarily a measure of sampling variation; i.e., the variation that occurs by chance, because a sample, rather than a population, is surveyed. The chances are 68 out of 100 that an estimate from the sample would differ from a complete census figure by less than the standard error.

1.4 Uses of Standard Errors

Estimation/Confidence Intervals

A sample estimate and its associated standard error enable one to construct confidence intervals--ranges that include the average result of all possible samples with specified probabilities. For example, if all possible samples were selected with each being surveyed under essentially the same conditions and using the same sampling design, and if an estimate and associated standard error were calculated from each sample, then:

(1) Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average estimate from all possible samples.

(2) Approximately 90 percent of the intervals from 1.6 standard errors below the estimate to 1.6 standard errors above the estimate would include the average estimate from all possible samples.

(3) Approximately 95 percent of the intervals from two standard errors below the estimate to two standard errors above the estimate would include the average estimate from all possible samples.

The average estimate derived from all possible samples may or may not be contained in any particular computed confidence interval. However, for a particular sample, one can say with a specified confidence that the average estimate derived from all possible samples would be included in the confidence interval.

Hypothesis Testing

Standard errors may also be used for hypothesis testing, a statistical technique for distinguishing between population characteristics using sample estimates. The most common type of hypothesis testing is to test that the population characteristics among a set of groups are same against that they are different. Tests may be performed at various levels of significance, where a level of significance is the chance of concluding that the characteristics are different while, in fact, they are identical.

To perform the most common hypothesis test to compare a population characteristic between two groups, compute the difference X_A - X_B, where X_A and X_B are sample estimates of the population characteristic of interest for the two groups. Let se_{DIF} be the standard error of the difference X_A - X_B. If the value of $(X_A - X_B)/se_{DIF}$ is between -1.96 and 1.96, no conclusion about the difference of the characteristics between the two groups would be justified at the 5 percent significance level. If, however, $(X_A - X_B)/se_{DIF}$ is smaller than -1.96 or greater than 1.96, the observed difference would be justified significant at the 5 percent significance level. In this case, it is commonly accepted practice to say the characteristics are different between the two groups. Of course, sometimes this conclusion might be wrong. When the characteristics are, in fact, the same, there is a 5 percent chance of concluding that they are different. The test conducted here is called the *z-test*, where *z* is obtained from the standard normal distribution tables and 1.96 is called the critical value of the test at the 5 percent significance level. This test is applicable when the sample sizes from the two groups are sufficiently large so that the central limit theorem holds. If, however, the sample sizes are not sufficiently large, one has to assume that the two populations from which the samples are drawn are approximately normally distributed and the appropriate test is the *t*-test. The t-test has a somewhat similar formulation to the z-test described above and uses ‘t’ tables for critical values instead of the standard normal tables (Ott 1977). All statistical software can perform t-tests and include as output a statistic called a p-value indicating the observed significance level: if the p-value is less than 0.05, that is, the observed significance level is below the specified 5 percent significance level, the difference is justified significant; otherwise, it is not significant.

Note that as more hypothesis testings are performed, more erroneous significant differences may occur. For example, if 100 independent testings were performed at the 5 percent significance level in which there are no real differences, it is likely that about 5 erroneous conclusions would occur. Therefore, if a large number of testings are performed, the significance of any single test should be interpreted cautiously or a Bonferroni significance level adjustment (Mendenhall et al. 1981) should be made for each of the tests. This adjustment procedure will ensure that *all* of the confidence intervals will enclose their respective parameters with *at least* a certain probability.

1.4 Reliability of an Estimated Proportion

This section refers to the proportions of a group of individuals possessing particular attributes such as the proportion of teachers in public schools who are Hispanic. The reliability of an estimated proportion, computed by using sample data for both numerator and denominator, depends upon both the size of the proportion and the magnitude of the totals upon which the proportion is based. Estimated proportions are relatively more reliable than the corresponding estimates of the numerators of the proportions, particularly if the proportions are 0.5 or more (Short and Littman 1989).

1.5 Computation of Complex Survey Standard Errors

Complex sample designs--those that use stratification, clustering, unequal selection probabilities, and multi-stage sampling, such as SASS--require procedures for estimating sampling variation that are markedly different from the ones that apply when the data are from a simple random sample. In general, such complex designs yield statistics with larger standard errors than those from a simple random sample (Wolter 1985).

A class of techniques, called replication methods, provides a general approach to estimating standard errors for the types of sample designs and weighting procedures usually encountered in complex sample surveys such as SASS. In particular, the *balanced half-sample replication* (also called *balanced repeated replication*, abbreviated as BRR) method, as a direct estimation method, has been used to estimate the standard errors associated with the estimates for all of the 1990-91 SASS surveys. NCES has prepared public use data files for the 1990-91 SASS which include a set of 48 weighted replicates designed to produce balanced half-sample replication variance estimates (Kaufman and Huang 1993, Gruber et al. 1994). For a more detailed description of the balanced half-sample replication method, users are referred to section 1.3, Volume II of this publication.

The set of 48 BRR weighted replicate provided in the 1990-91 SASS public use data files can be utilized only by users who have software available to perform the balanced half-sample replication estimation. One instance of such software is a SAS (Statistical Analysis System) user-written procedure called PROC WESVAR developed by Westat, Inc. (Westat, 1993), which computes basic survey estimates and their associated sampling errors for user-specified characteristics. PROC WESVAR supports a BRR option which should be used along with the replicate weights which are prepared externally and supplied in the data file for estimation of sampling errors. In this manual, without indication, all standard errors, referred to as directly estimated, were produced through the BRR procedure using WESVAR.

With a variance estimation procedure such as BRR described above, it is possible to compute and show a standard error for each survey estimate in the results tables of SASS reports. However, the SASS data set contains approximately 1,500 variables. In addition, statistics such as totals, averages, proportions, and differences with respect to various subpopulations can also be estimated. Even if each published sample estimate was accompanied by its standard error, one could not predict the combinations of results (ratios, differences, etc.) that might be of interest to the user. Users will therefore not always find individual standard errors for each estimate published in SASS reports or other additional estimates of interest. The statistical software WESVAR, and another, SUDAAN (Shah et al. 1992), a main software for complex survey variance estimation, are not widely available for users to compute standard errors. These are the practical reasons that more general analytical techniques are desirable.

Standard errors, when estimated from sample data, are themselves subject to sampling error. The standard error for a survey statistic of interest generally has a larger relative (with respect to the magnitude of the standard error) sampling error than that for the estimated statistic. Thus the estimates of standard errors may vary considerably from one time of estimation to another or among related characteristics (that might be expected to have nearly the same magnitude of relative sampling error). Therefore, some techniques of stabilizing the standard error or variance estimates, for example, by generalizing or by averaging, are desired to improve their usefulness.

Empirical studies (Synectics 1992 and Volume II of this publication) have shown that appropriately formed groups of SASS statistics tend to have similar *design effects* (see section 2) and similar behavior, in some sense, of the *relative variance* (see section 3). Based on these studies, two general methods have been made available to calculate the standard errors for the 1990-91 SASS: the design effect method (section 2) and the generalized variance function (GVF) method (section 3), using the tables provided with this manual (appendix II and appendix III). Section 1.6 below describes, first, all the groups of statistics for which average design effects and GVFs are available from the tables. We will show how to use these tables in the following sections.

1.6 Groups of Statistics

NCES publishes SASS statistics for many characteristics (e.g., number of K-12 students in the U.S.) and some standard subpopulations (e.g., public and private schools). Based on these publications, and in anticipation of various combinations of results (e.g., totals, averages, and proportions) being of interest to users, table 1.1 below lists the groups of statistics for use in computing standard errors.

The first level of grouping was one of the four surveys: School, School Administrator, Teacher Demand and Shortage (TDS), or Teacher. There are a very large number of certainty and high probability districts in the public TDS sample. These districts also contain a very large proportion of the total number of teachers and students. For the complex SASS design, these districts contribute very little to the variance estimates of totals and averages. However, for a simple random sample design, these same districts do contribute a very large part of the variance estimates of totals and averages. Due to these differences in variance contribution, and depending on the subpopulation, the design effects can vary greatly. Often these design effects can be extremely small (design effects less than 0.2 are not uncommon). Hence, an average design effect would be inappropriate. District proportions have the same problem, but to a lesser extent. For this reason, we do not present average design effects or GVF tables for the public TDS.

The second level of grouping was within each survey--either totals, averages, or proportions were grouped together. For example, if a user needs to estimate the standard error of "the number of students in K-12 who are Hispanic," the user would first locate the correct design effect or GVF table based on one of these groups. In this example, the variable of interest (students in K-12 who are Hispanic) is found in the School Survey and the estimate of interest is a **total**; i.e., the total number of students. Therefore, the correct table to use would be found in the group labeled "School Survey - Student Totals."

Table 1.1 -- Groups of statistics in 1990-91 SASS

Survey	Group of Statistics
School	Student Totals (e.g., number of students enrolled in 1st grade) Teacher Totals (e.g., number of full-time K-12 teachers) School Proportions (e.g., proportion of schools offering kindergarten)
School Administrator	Administrator Totals (e.g., number of administrators with master's degrees) Administrator Proportions (e.g., proportion of male administrators)
Teacher Demand and Shortage (Private)	TDS Totals (e.g., number of full-time equivalent teachers with state certification) TDS Proportions (e.g., proportion of districts with retraining offered teachers: special education)
Teacher	Teacher Totals (e.g., number of male teachers) Teacher Averages (e.g., average number of years as a part-time teacher) Teacher Proportions (e.g., proportion of married teachers)

SOURCE: U.S. Department of Education, National Center for Education Statistics, Schools and Staffing Survey: 1990-91.

Table 1.2 describes the subpopulations available for each group of statistics in the four SASS surveys, and table 1.3 provides definitions of each subpopulation. For example, a user may need to estimate the standard error of the number of students in grades K-12 who are Hispanic in **private schools**. The subpopulation of interest in this example is "private schools," and the standard error is calculated by using the parameters available in the row labeled "Private" (under the subpopulation heading "Sector") in either the design effect or GVF table labeled "School Survey - Student Totals."

Table 1.2 -- Relevant subpopulations for groups of statistics in 1990-91 SASS

Survey	Subpopulation for each group of statistics
School	Sector Region Region within Sector School Level within Sector School Level within State (elementary and secondary public schools) Typology (private schools only) Community Type within Sector State (public schools only) School Size within Community Type within Sector Minority Status (of Students) within Community Type within Sector
School Administrator	Sector Region State (public schools only) Region within Sector School Level within Sector School Level within State (elementary and secondary public schools) Typology (private schools only)
Teacher Demand and Shortage (Private Only)	Region Typology School Level Minority Status (of Students)
Teacher	Sector Region Region within Sector Minority Status (of Students) within Sector State (public schools only)

SOURCE: U.S. Department of Education, National Center for Education Statistics, Schools and Staffing Survey: 1990-91.

Table 1.3 -- Definition of subpopulations in 1990-91 SASS

Subpopulation	Definition	
Sector	Public or Private Schools	
Region		
Northeast	Maine, New Hampshire, Vermont, Massachusetts, Rhode Island, Connecticut, New York, New Jersey, Pennsylvania	
Midwest	Ohio, Indiana, Illinois, Michigan, Wisconsin, Minnesota, Iowa, Missouri, North Dakota, South Dakota, Nebraska, Kansas	
South	Delaware, Maryland, District of Columbia, Virginia, West Virginia, North Carolina, South Carolina, Georgia, Florida, Kentucky, Tennessee, Alabama, Mississippi, Arkansas, Louisiana, Oklahoma, Texas	
West	Montana, Idaho, Wyoming, Colorado, New Mexico, Arizona, Utah, Nevada, Washington, Oregon, California, Alaska, Hawaii	
School Level	Elementary (no grade higher than 8 and at least one of grades 1-6), Secondary (grades 7-12), and Combined (any other combination of grades; e.g., 4-9, or 5-12)	
Typology	The private school typology separates private schools into three major groups and within each group into three subgroups: Catholic (parochial, diocesan, and private order), other religious (Conservative Christian, affiliated, and unaffiliated), and nonsectarian (regular, special emphasis, special education) (McMillen and Benson 1991)	
School Size	Enrollment of fewer than 150 students Enrollment of 150 to 499 students	Enrollment of 500 to 749 students Enrollment of 750 or more students
Community Type	Central City includes large central cities (Central cities of Standard Metropolitan Statistical Areas (SMSAs), with populations greater than or equal to 400,000 or population densities greater than or equal to 6,000 per square mile) and mid-size central cities (central cities of SMSAs, but not designated as large central cities). Urban Fringe/Large Town includes the urban fringes of large or mid-size cities (places located within SMSAs of large or mid-size central cities and defined as urban by the U.S. Bureau of the Census) and large towns (places not located within an SMSA, but that have populations greater than or equal to 25,000 and that are defined as urban by the U.S. Bureau of the Census). Rural/Small Town includes rural areas (places that have populations of fewer than 2,500 and that are defined as rural by the U.S. Bureau of the Census) and small towns (places not located within SMSAs, that have populations of fewer than 25,000, but greater than or equal to 2,500, and that are defined as urban by the U.S. Bureau of the Census).	
Minority Status	Minority enrollment (sum of all racial/ethnic groups other than white) of less than 20 percent, or greater than or equal to 20 percent.	
Field of Teaching	elementary general elementary special education elementary other secondary math secondary science	secondary English secondary social studies secondary vocational education secondary special education secondary other

SOURCE: U.S. Department of Education, National Center for Education Statistics, Schools and Staffing Survey: 1990-91.

2. Average Design Effects and Approximate Standard Errors

Regardless of which method is used to calculate the standard errors for statistics derived from the SASS data, they will be different from the standard errors that are based on the assumption that the data are from a simple random sampling. The SASS complex design differs from the simple random sampling. The impact of the complex design on the accuracy of a sample estimate, in comparison to the alternative simple random sampling, is often measured by the *design effect* (Deff), defined as the following ratio:

$$Deff = \frac{var_{COMPLEX}}{var_{SRS}}$$

$$= \frac{sampling\ variance\ of\ complex\ sample}{sampling\ variance\ of\ simple\ random\ sample}$$

One may think of this ratio as a measure of the efficiency of the actual design.

In a large scale sample survey such as SASS, data are collected for a large number of variables. This necessitates that the design effects be computed for at least some key variables. The average of these design effects can be considered as a measure of the efficiency of the survey design compared to the alternative simple random sampling. For the 1990-91 SASS, accordingly, an average design effect was derived for each group of statistics (table 1.1) and, within each group, for each classification of each subpopulation (table 1.2).

2.1 Design Effects and Their Use

Standard errors of complex survey statistics of various groups for various subpopulations can then be calculated approximately from the corresponding standard errors based on the alternative simple random sample and the average design effects corresponding to the groups and subpopulations. The calculation formula for the standard error of an estimate is expressed as follows:

$$se_{COMPLEX} = \sqrt{Deff \times v_{SRS}}$$

$$= \sqrt{Deff} \times se_{SRS}$$

where v_{SRS} is the estimated variance of the estimate from a simple random sample, and se_{SRS} is the corresponding standard error. The calculation formulas for v_{SRS} from sample data for three basic types of estimates, totals, averages, and proportions, are provided below. Let x be the variable of interest with sample values x_i, i = 1,...,n.

2.1.1 Calculation of Simple Random Sample Variance for Totals

$$v_{SRSTOT} = (\sum_1^n w_i)^2 \frac{1}{n} \frac{\sum_1^n w_i (x_i - \bar{x}_w)^2}{\sum_1^n w_i - 1}$$

$$= (\sum_1^n w_i)^2 \frac{1}{n} s_w^2$$

where w_i are the weights, n is the number of respondents in the sample,

$$\bar{x}_w = \frac{\sum_1^n w_i x_i}{\sum_1^n w_i},$$

and

$$s_w^2 = \frac{\sum_1^n w_i (x_i - \bar{x}_w)^2}{\sum_1^n w_i - 1}.$$

The above formula for v_{SRSTOT} can be written in terms of the standard error, say,

$$se_{SRSTOT} = (\sum_1^n w_i) \frac{s_w}{\sqrt{n}}$$

$$= (\sum_1^n w_i)\, se_{SRSAVG}.$$

Remark The quantity $s_w/n^{1/2} = se_{SRSAVG}$ is the standard error of the (weighted) mean of x (see section 2.1.2). It can be computed from SAS or SPSS procedures. An illustration of the SAS codes, using PROC MEANS, for computing se_{SRSAVG} and the total weight is provided below (SAS Institute Inc. 1990):

```
PROC MEANS DATA=SAS-data
    VARDEF=WDF VAR STD STDERR SUMWGT;
    VAR x;
    WEIGHT weight;
RUN;
```

where x is the variable for which the standard error of the (weighted) mean is requested, and *weight* is the variable for weights. The statistics VAR (the variance) and STD (the standard deviation) are included here for illustration purpose. The option VARDEF=WDF specifies the sum of weights minus one being used as the divisor in the calculation of the weighted VAR (as the s_w^2 above). The statistic STDERR (the standard error of the mean) is the desired se_{SRSAVG}, which is calculated by the weighted STD (as the s_w above) divided by the square root of the number of observations (as the n above). The statistic SUMWGT gives the total weight.

Note SAS is designed only for analyzing samples from infinite populations. To make the statistic STDERR in the form based on infinite population sampling, starting in release 6.11, with the procedures MEANS, SUMMARY, TABULATE and UNIVARIATE, the statistic STDERR for weighted mean will be calculated as the weighted STD (with VARDEF=DF) divided by the square root of the sum of weights. To use SAS 6.11 to compute se_{SRSAVG}, the codes need be modified accordingly.

Example 1 Consider the total enrollment of public school students in rural communities in K-12 plus those who are ungraded. In the School Survey data file, the variable is named ENRK12UG (Total Rural School Enrollment K-12 Plus Ungraded) (Gruber et al. 1994, appendix D-2). There are n = 4,993 records belonging to the subpopulation of interest, Public/Rural (i.e., Public/Rural-Small Town) under Sector/Community Type. Using the above SAS procedures, we can get se_{SRSAVG} = 4.1119, and the total weight 40,352. Thus, the simple random sample standard error for a total is the product of the se_{SRSAVG} and the total weight:

$$se_{SRSTOT} = 40{,}352 \times 4.1119 = 165{,}923.39.$$

Referring to the School Survey Design Effects table in appendix II, page II-9, the design effect for student total for the subpopulation Public/Rural under Sector/Community Type is Deff = 1.8167. Using the first equation of section 2.1 to calculate the approximate standard error for the total enrollment of public school students in rural communities in K-12 plus ungraded, we can substitute the above obtained values for se_{SRSTOT} and Deff:

$$se_{TOT} = \sqrt{Deff} \times se_{SRSTOT}$$

$$= \sqrt{1.8167} \times 165{,}923.39 = 223{,}639.9.$$

A direct estimate for this standard error is, say, *se*=189,642.5 (Choy et al. 1993, table B1, p.171). The relative difference in percent of se_{SRSTOT}, compared with the direct estimate *se*, is 100 x $|se_{DEFF}$ - $se|$ /*se* = 100 x |223,639.9 - 189,642.5| / 189,642.5 = 17.9(%).

For users who are more familiar with SPSS than SAS, we provide below an illustration of the SPSS codes for computing se_{SRSAVG} and the total weight (SPSS Inc., 1993a):

```
GET FILE=SPSS-data.
COMPUTE wvar=1.
EXECUTE.
WEIGHT BY weight.
DESCRIPTIVES VARIABLES=wvar
    /STATISTICS=SUM.
DESCRIPTIVES VARIABLES=x
    /STATISTICS=SEMEAN.
```

where *x* is the variable for which the standard error of the (weighted) mean is requested, and *weight* is the variable for weights. The first DESCRIPTIVES computes the sum of weights. In the second DESCRIPTIVES, the statistic SEMEAN, defined also as the standard error of the mean, is calculated as the weighted standard deviation divided by the square root of the sum of weights (SPSS Inc. 1993b), differently from SAS. Thus an additional calculation is needed to get the desired se_{SRSAVG}:
se_{SRSAVG} = SEMEAN x sqrt{sum of weights / number of observations}.

2.1.2 Calculation of Simple Random Sample Variance for Averages

$$v_{SRSAVG} = \frac{1}{n}\frac{\sum_{1}^{n} w_i(x_i - \bar{x}_w)^2}{\sum_{1}^{n} w_i - 1}$$

$$= \frac{1}{n}s_w^2 = \left(se_{SRSAVG}\right)^2$$

where w_i are the weights, and

$$\bar{x}_w = \frac{\sum_1^n w_i x_i}{\sum_1^n w_i}.$$

se_{SRSAVG}, as described in last section, can be obtained from SAS or SPSS.

Example 2 Consider the same variable and subpopulation as in Example 1, but for student average. The design effect for student average for the subpopulation Public/Rural (i.e., Public/Rural-Small Town) under Sector/Community Type, from the School Survey Design Effects table in appendix II, page II-9, is Deff = 1.6410. Then, with se_{SRSAVG} = 4.1119 from Example 1, the desired standard error is calculated as

$$se_{AVG} = \sqrt{Deff} \times se_{SRSAVG}$$

$$= \sqrt{1.6410} \times 4.1119 = 5.2674.$$

2.1.3 Calculation of Simple Random Sample Variance for Proportions

$$v_{SRSPROP} = \frac{p(1-p)}{n},$$

$$se_{SRSPROP} = \sqrt{\frac{p(1-p)}{n}}$$

where p denotes the estimate of proportion for a characteristic of interest, expressed as

$$p = \frac{\sum_1^n w_i I(i)}{\sum_1^n w_i}$$

where $I(i)$ = 1 if the characteristic is present for the sampled unit and 0 if it is absent.

Example 3 Consider the proportion of private school teachers who have bachelor's degrees as highest degree earned. There are n = 6,642 teacher records belonging to the subpopulation Private under Sector. An estimated (weighted) proportion is p = 0.619 (Choy et al. 1993, table 3.7, where the listed value is the percentage, 61.9). Thus, using the equation specified above, the standard error of p from the alternative simple random sample is

$$se_{SRSPROP} = \{0.619 \times (1 - 0.619) / 6642\}^{1/2} = 0.0060.$$

The design effect for teacher proportion for the subpopulation Private under Sector, from the Teacher Survey Design Effects table in appendix II, page II-39, is Deff = 1.9053. An approximate standard error for the proportion of interest is calculated as:

$$se_{PROP} = \sqrt{Deff} \times se_{SRSPROP}$$

$$= \sqrt{1.9053} \times 0.0060 = 0.0083.$$

An available direct estimate for this standard error is 0.009; see Choy et al. 1993, table B4, p.176, where the listed standard error, 0.90, being for percentage, is converted to the standard error, 0.009, for proportion. The relative (absolute) difference in percent of se_{PROP}, compared with the direct estimate, is $100 \times |0.0083 - 0.009|/0.009 = 7.8(\%)$.

2.2 Average Design Effect Tables

In appendix II, the tables give the average design effects for each survey and subpopulation. SASS users who do not have access to software for computing accurate standard errors can use the average design effects presented in these tables and the formulas in section 2.1 to approximate the standard errors of statistics based on the SASS data.

2.3 Outlier Variables in the Average Design Effect Groups

When examining the design effect tables, readers may notice some relatively high average design effects. These appear to be attributable to some highly skewed variables included in the surveys. Removal of those variables would produce homogeneous design effects. For surveys with a large number of variables, removal of a few highly skewed variables would not effect the calculation of average design effects. However, for some of the surveys in this study there were not many variables used in the average design effect calculation and therefore the highly skewed variables were kept in for calculating the average design effects. Table 2-1 below presents the highly skewed variables identified in each of the survey components.

Table 2.1 -- Variables with very high design effects

Survey	Type of estimate	Variable	Variable Label
School Survey	Student Totals	NUMBRPK	Number of students enrolled in pre-k
		NUMBR7	Number of students enrolled in grade 7
		NUMBR8	Number of students enrolled in grade 8
		BILINGNUM	Number of Bilingual Ed students
		AFTERNUM	Number of extended day students
School Administrator Survey	Totals	ASC017	Have a masters degree
	Average	ASC031	Number of years teaching experience before becoming a principal
	Average	ASC047	Number of years in other nonteaching, nonadministrator positions in elem/secondary education, e.g. a guidance counselor.
	Average	ASC048	Number of years in positions outside of elementary/secondary education.
Teacher Survey	Total	RACE=4	Race Ethnicity=White

SOURCE: U.S. Department of Education, National Center for Education Statistics, Schools and Staffing Survey: 1990-91.

3. Generalized Variance Functions and Approximate Standard Errors

Sampling variance or the *relative variance* of a survey estimator (defined as the sampling variance divided by the square of the mean of the estimator) can be related to the mean (expectation) of the estimator by simple mathematical relationships (Wolter 1985). A *generalized variance function* (GVF) is such a mathematical model which can be used to calculate the variance estimates (or standard errors) for survey items by evaluating the model at the corresponding survey estimates, avoiding computations of direct estimation. Thus, survey estimates with similar behavior of the relative variance (or its square root, the *coefficient of variation* (CV)[1]) were grouped together. Appropriate GVF (with two model parameters *A* and *B*) was developed for each group of survey estimates. The GVF for a group can be used to describe the behavior of the relative variance for all survey estimates in that group. The model parameters *A* and *B* vary by the group of statistics (totals, averages, proportions) and by the subpopulation (e.g., public schools) to which the estimate applies. The GVF tables in appendix III of this manual provide the parameters *A* and *B*, according to the groups of statistics and subpopulations as described in section 1.6, to be used for 1990-91 SASS estimates of interest.

It is noticed that, unlike the design effect approach, the GVF approach involves no need to calculate the simple random sample variance estimates. With the GVF tables provided, the calculation of a standard error takes only three simple steps:

(1) Read the parameters *A* and *B* from the GVF table corresponding to the survey estimate (*X*) of interest;

(2) Evaluate the GVF model at the survey estimate *X*, that is, calculate

$$CV(\%) = \sqrt{A + B/X}\,.$$

(3) Calculate the associated standard error of *X* as *se* = CV(%) x *X* /100.

Remark Because the CVs used to develop the GVF models were computed through WESVAR in the scale of percent (that is, 100 x (standard error/estimate)), the calculated CV from evaluating the GVFs will be also in the scale of **percent**. To get the CV to the normal scale, we need to divide by 100 the percent CV resulted from the GVF evaluation.

The *R*-squared column in the GVF table represents how well the model fits the 1990-91 SASS data. In practice, if a GVF has small *R*-squared value, say, less than 0.5, the GVF

[1] CV is estimated by the standard error of the estimate divided by the estimate.

would not be considered appropriate for use. For the GVFs for the 1990-91 SASS, there are only a few such cases.

Procedures for using the tables of the GVF parameters for the calculation of standard errors are illustrated through examples given in the following of this section.

3.1 Illustration of the Use of GVF Tables

GVFs were developed for the calculation of standard errors of totals, averages, and proportions of interest in the SASS surveys. GVF tables for totals, averages (see section 3.3), and proportions, by various subpopulations, are provided in appendix III of this manual. The following examples use the GVF tables to obtain the standard error for a total and a proportion estimates.

Example 1 Consider the total number of public school students in rural communities (see Example 1 of section 2.1). Table 3.1 below is an extract of the School Survey GVFs for student totals table for the subpopulations of Sector/Community Type (appendix III, page III-26). This table shows the GVF coefficients for the subpopulation Public/Rural, $A = 0.919$, and $B = 8{,}244{,}388.289$.

The estimated total number of the Public/Rural students is $X = 15{,}695{,}586$ (se = 189,642.5) (Choy et al. 1993, table 2.1, p.6, and table B1, p.171). The generalized CV (in percent) is calculated, by the formula in above step (2), as

$$\text{CV}(\%) = \{0.919 + (8{,}244{,}388.289/15{,}695{,}586)\}^{1/2} = 1.201777.$$

The GVF standard error (se_{GVF}) is then calculated as

$$\begin{aligned} se_{GVF} &= (\text{CV}/100) \times X \\ &= (1.201777/100) \times 15{,}695{,}586 = 188{,}625.942. \end{aligned}$$

This result can be compared with the published standard error for the total, from direct estimation, 189,642.5, as listed above with the estimate X. They appear quite close with a relative (absolute) difference in percent of $100 \times |se_{GVF} - se|/se$ = $100 \times |188{,}625.9 - 189{,}642.5|/189{,}642.5 = 0.536(\%)$.

The R-squared column in the GVF table represents how well the model fits the 1990-91 SASS data. For this case, the R-squared value is 0.8801.

The standard error calculated in this example was calculated as 223,639.9, by the design effect approach, in example 1 of section 2.1, with a relative difference 17(%), as compared to the direct estimate 189,642.5. For this example, the GVF approach appears having better performance than the design effect approach.

Table 3.1 -- GVFs for student totals (School Survey) (GVF model:*CV(%) = (A+B/X)*$^{1/2}$*)*

Sector / Community Type	Parameter		Measure of Fit
	A	*B*	*R*-squared
Public / Urban	4.260	11,127,626.44	0.6182
Public / Suburban	1.970	10,321,487.16	0.7684
Public / Rural	**0.919**	**8,244,388.289**	**0.8801**
Private / Urban	3.985	2,771,444.620	0.8751
Private / Suburban	5.076	3,600,659.902	0.7697
Private / Rural	16.455	4,420,924.491	0.7602

SOURCE: U.S. Department of Education, National Center for Education Statistics, Schools and Staffing Survey: 1990-91 (School Questionnaires).

Example 2 Consider the proportion of private school teachers with bachelor's degree as highest degree earned. Table 3.2 is an extract of the Teacher Survey GVFs for teacher proportions table for the subpopulations of Sector (appendix III, page III-101). This table shows the GVF coefficients for the subpopulation Private, A = -2.6522, and B = 2.6695.

The estimated proportion of the private school teachers with bachelor's degree is X = 0.619 (se = 0.0090) (Choy et al., 1993, table 3.7, p.45, and table B4, p.176. Listed in these tables are the estimated percentage, 61.9, and the associated standard error, 0.90. This percentage can be converted to proportion as 0.619, by a division by 100, and similarly the associated standard error converted to 0.0090). The generalized CV (in percent) is calculated, by the formula in above step (2), as

$$\text{CV}(\%) = \{-2.6522 + 2.6695/0.619\}^{1/2} = 1.2886.$$

The GVF standard error (se_{GVF}) of the estimated proportion is then calculated as

$$se_{GVF} = (\text{CV}/100) \times X$$
$$= (1.2886/100) \times 0.619 = 0.007976.$$

This result can be compared with the published standard error, from direct estimation, 0.0090, as listed above with the estimate *X*. The relative (absolute) difference in percent is $100 \times |se_{GVF} - se|/se = 100 \times |0.007976 - 0.0090|/0.0090 = 11.4(\%)$. The *R*-squared value for this GVF is quite high as 0.9807, listed in the *R*-square column of table 3.2.

The standard error calculated in this example was calculated as 0.0083, by the design effect approach, in example 3 of section 2.1, with a relative difference 7.8(%), also compared to the direct estimate 0.0090. For this example, the design effect approach appears having better performance than the GVF approach.

Table 3.2 -- GVFs for teacher proportions (Teacher Survey) (GVF model: ***CV(%) = (A+B/X)***$^{1/2}$**)**

Sector	Parameter		Measure of Fit
	A	*B*	*R*-squared
Public	-0.5385449013	0.5372155053	0.9725
Private	**-2.652233929**	**2.669488096**	**0.9807**

SOURCE: U.S. Department of Education, National Center for Education Statistics, Schools and Staffing Survey: 1990-91 (Teacher Questionnaires).

They might, of course, both perform poorly in some other cases. Generally, the two approaches lie on the same theoretical ground: an appropriately formed group of statistics for a subpopulation has similar behavior in the sampling variance. GVF and design effect represent two aspects of the similarity. Methodologically, regarding their applicability and accuracy delivered, they are considered in the same category. Therefore, there is no general criterion can be established for making decision of selecting between the two approaches.

3.2 Standard Error of a Ratio

To estimate the relative variance of an estimated proportion $\hat{R} = \hat{X} / \hat{Y}$,

where $\hat{Y}$ is an estimator of the total number of individuals in a certain subpopulation and $\hat{X}$ is an estimator of the number of those individuals with a certain attribute.

When $\hat{R}$ and the denominator $\hat{Y}$ are approximately uncorrelated, the relative variance $V^2{}_R$ of $\hat{R}$ can be approximately calculated from the relative variances $V^2{}_X$ of $\hat{X}$ and $V^2{}_Y$ of $\hat{Y}$ by

$$V^2{}_R = V^2{}_X - V^2{}_Y. \qquad (1)$$

Formula (1) has been shown to produce useful approximations. The estimate of $V^2{}_X$ and $V^2{}_Y$ can be read, approximately, from the appropriate GVF tables. $\hat{X}$ and $\hat{Y}$ are usually in the same group of statistics. With Model 1, more specifically, it follows

$$V_R^2 = B(\hat{X}^{-1} - \hat{Y}^{-1}).$$

This approach of approximating the relative variance of a proportion could be applied to ratios, under a similar assumption, that is, the correlation between the ratio $\hat{R}$ and the denominator $\hat{Y}$ is close to 0. The following is an illustrative example.

Example Consider the student-teacher ratio for national public schools. The teacher number in each school counted is for the full-time-equivalent (FTE) teachers, which is calculated as a combination of the numbers of full-time teachers and part-time teachers in the following way, according to the NCES guideline:

FTE teachers = full-time teachers + 0.54 part-time teachers.

(In SASS School Survey files, the variable for the number of full-time teachers in school is FULTEACH, and for the number of part-time teachers in school is PARTEACH. The variable for the number of students in school is ENRK12UG.)

The following table lists, for national public schools, the estimates of the student total, FTE teacher total, and their ratio, and the associated standard errors, as directly estimated via BRR. For convenience, a last column for CV (in percent) is added to the table.

Table 3.3 -- Student and teacher totals and their ratio for public schools

Variable	Total	Standard error	CV(%)
Students (X)	40103699	362552.64	0.9040
FTE teacher (Y)	2439057	20331.12	0.8336
Students/FTE teacher (R)	16.4423	0.05863	0.3566

SOURCE: U.S. Department of Education, National Center for Education Statistics, Schools and Staffing Survey: 1990-91.

Now use the formula $V^2_R = V^2_X - V^2_Y$, to calculate the CV for the ratio from the CVs for the numerator and the denominator of the ratio,

$$CV_R = \{(CV_X)^2 - (CV_Y)^2\}^{1/2}$$
$$= (0.9040^2 - 0.8336^2)^{½} = 0.3498.$$

This result of CV (in percent) is very close to the directly estimated CV (in percent) for the ratio, 0.3566, as listed in table 3.3. The relative (absolute) difference is 100 x |0.3498 - 0.3566|/0.3566 = 1.9(%).

We also use the GVF estimates of the relative variances for X and Y. From the School Survey GVFs for Student Totals table (appendix III, page III-19) under the subpopulation Public of Sector, the GVF parameters for X are $A_X = 0.590$ and $B_X =$ 9872132.241. The relative variance for X is then calculated as

$$V_X^2 = A_X + B_X/X$$
$$= 0.590 + 9872132.241/40103699$$
$$= 0.8362.$$

And from the School Survey GVFs for Teacher Totals table (appendix III, page III-35) under the subpopulation Public of Sector, the GVF parameters for Y are $A_Y = 0.6880$ and $B_Y =$ 119403.0681. The relative variance for Y is then calculated as

$$V_Y^2 = A_Y + B_Y/Y$$
$$= 0.6880 + 119403.0681/2439057$$
$$= 0.7370.$$

Thus the relative variance for R is calculated as

$$V_r^2 = V_x^2 - V_y^2$$
$$= 0.8362 - 0.7370 = 0.0992,$$

and the corresponding CV_R (in percent) is 0.3150. This result is fairly close to the directly estimated CV (in percent) for the ratio, 0.3566 (table 3.3). The relative (absolute) difference is $100 \times |0.3150 - 0.3566|/0.3566 = 11.7(\%)$.

Remark The assumption that $\hat{R}$ and $\hat{Y}$ are uncorrelated is critical for the formula (1) to give useful approximations. In practice, the relative variance estimate for the numerator may be smaller than that for the denominator, resulting in a negative relative variance for the ratio. This circumstance is an indication that the assumption is violated. In the case that the ratio is a proportion, and Model 1 GVF estimates are valid for the relative variances of $\hat{X}$ and $\hat{Y}$, the negative relative variance problem will not occur.

3.3 Standard Error of an Average

The standard error of an **average** can be derived approximately from the standard error of the corresponding **total** according to the following formula:

$$se_{COMPLEXAVG} = \frac{se_{COMPLEXTOT}}{\sum_{1}^{n} w_i}$$

where $se_{COMPLEXTOT}$ is the standard error associated with a total type estimate, either obtained using a GVF table or directly estimated, and w_i are the weights. The above formula is approximate because the domain over which the weights are summed (in the denominator) can vary randomly. The summing of weights is over the sample units within the group of interest. This total weight provides an estimated total number of individuals in the subpopulation defined by that group. For example, for the variable NUMBR4: “NMBR STUDENTS ENROLLED IN 4TH GRADE” in the School Survey (Gruber et al. 1994, appendix D-13), if our interest is in the Public/Region NE group, the total of weights would sum up the weights of the public schools in the sample which belong to the Northeast region; the total weight would be an estimated total number of public schools in the Northeast region.

Tables of total weights of the sample units over various subpopulations of interest are provided for each survey with this manual (appendix IV). However, it should be noticed that the total weights in these tables were calculated according to **all** sample units belonging to the subpopulation. That is, all sample units were considered as respondents. But that might not be the real case. For survey totals with a high item nonresponse rate, using the total weights corresponding to all sample units may cause unignorable error, resulting in an underestimate of the standard error for the average. There seems no convenient way to incorporate the individual item nonresponse rates into the tables of total weights which are produced for general use. In the case that, as mentioned above, the item nonresponse rate is high, caution must be taken and users are urged to calculate the total weights individually for that item by summing up weights over only the respondents for that item in the sample.

The following example illustrates the use of the formula.

Example Consider the variable HISPNSTU (NMBR K-12 STUDENTS ARE: HISPANIC) in School Survey (Gruber et al. 1994, appendix D-11) and the group Public/Urban of Sector/Community Type. A directly estimated standard error by BRR for the total is $se_{COMPLEXTOT}$ = 102,238.68. The total weight for the (responding) schools in that group is calculated from the data as 18,683.82. The derived standard error for the average is then

$$se_{COMPLEXAVG} = 102{,}238.68 \,/\, 18{,}683.82 = 5.472.$$

A directly estimated standard error by BRR for the average is, say, se_{AVG} = 5.3435. Compare the two results, and calculate the relative difference in percent:

$$100\ (se_{COMPLEXAVG} - se_{AVG}) \,/\, se_{AVG}$$

$$= 100 \times (5.472 - 5.3435) \,/\, 5.3435 = 2.4\ (\%).$$

Also, we can use the GVF approach to estimate the standard error for total from the estimated total. For this example, the estimated total is X = 2,318,226.59. From the GVF table, the School Survey GVFs for student totals, for the group Public/Urban of Sector/Community Type, it is found that the estimated coefficients are: A = 4.26 and B = 11,127,626.44. Thus, by the GVF model,

$$\text{CV (in percent)} = (A + B/X)^{1/2}$$

$$= (4.26 + 11{,}127{,}626.44/2{,}318{,}226.59)^{1/2} = 3.01,$$

and the GVF modeled standard error for the total is

$$se_{COMPLEXTOT} = X \times \mathrm{CV}$$
$$= 2{,}318{,}226.59 \times 3.01/100 = 69{,}778.62.$$

Using this estimate of the standard error for the total, the derived standard error for the average is

$$se_{COMPLEXAVG} = 69{,}778.62/18{,}683.82 = 3.7347,$$

where 18,683.82 is the total weight. A comparison between this estimate and the direct estimate is given by the relative difference:

$$100\ (se_{COMPLEXAVG} - se_{AVG}\,)\ /\ se_{AVG}$$
$$= 100 \times |3.7347 - 5.3435|\ /\ 5.3435 = 30\ (\%).$$

This time the result from using GVF seems not to give satisfactory accuracy. It is noticed that the R-squared value for the GVF used is 0.6182, so the model didn't fit very well.

3.4 Outlier Variables Found in the GVF Groups

Users are cautioned that during the GVF modeling process some variables were found to be outliers; i.e., they differed considerably from that of most of the variables in a group. GVF models used for these variables will give poor results. Table 3.4 provides a list of specific variables for which GVFs may be inappropriate.

Table 3.4 -- Outlier variables found in the GVF Groups

Survey/Estimate	Subgroup	Variable	Label
School: Student Totals	Illinois/ Secondary	NUMBRPK	Number of students enrolled in pre-Kindergarten
School: Teacher Totals	North Dakota	ASIANTCH	Number of K-12 teachers that are Asian/Pacific Islander
School: Teacher Totals	Private/Rural/750+	SPCLNEW	Number of new K-12 teachers, main assignment: special ed
Administrator: Totals	Catholic/Private	ASC072	Problem : Student apathy
Administrator: Proportions	Kansas	ASC124	Of Hispanic origin
Administrator: Proportions	New York	ASC123	Enrolled in recognized tribe
Administrator: Proportions	North Carolina	ASC123	Enrolled in recognized tribe
Administrator: Proportions	Idaho/Elementary	ASC042	Participated in training for aspiring school administrators
Administrator: Proportions	Idaho/Elementary	ASC043	Completed the Indian Education Administration Program
Administrator: Proportions	Illinois/Elementary	ASC124	Of Hispanic origin
Administrator: Proportions	Kansas/Secondary	ASC124	Of Hispanic origin
Administrator: Proportions	New York/Elementary	ASC123	Enrolled in recognized tribe
Administrator: Proportions	New York/Secondary	ASC124	Of Hispanic origin
Administrator: Proportions	North Carolina/ Elementary	ASC123	Enrolled in recognized tribe
Administrator: Proportions	North Carolina/ Secondary	ASC123	Enrolled in recognized tribe

SOURCE: U.S. Department of Education, National Center for Education Statistics, Schools and Staffing Survey: 1990-91.

References

Ahmed, S. W. (1993a), "Issues Arising in the Application of Bonferroni Procedures in Federal Surveys," *American Statistical Association 1993 Proceedings of the Section on Survey Research Methods.* Alexandria, VA: American Statistical Association.

Ahmed, S. W. (1993b), "Technical Approaches to Performing Regression and Other Multivariate Techniques on NCES Survey Data - Where We Stand." A note from the Chief Statistician. Washington, DC: National Center for Education Statistics.

Choy, S. P., Henke, R. R., Alt, M. N., and Medrich, E. A. (1993), *Schools and Staffing in the United States: A Statistical Profile, 1990-91*. NCES 93-146. Washington, DC: National Center for Education Statistics.

Cochran, W.G. (1977), *Sampling Techniques*, third edition. New York: John Wiley.

Copas, J. B. (1983), "Regression, Prediction and Shrinkage," *Journal of the Royal Statistical Society B*, 45, 311-354.

Edelman, M. W. (1967), "Curve Fitting of Keyfitz Variances." Unpublished memorandum. Washington, DC: U.S. Bureau of the Census.

Fay, R.E., (1995), *VPLX*. Washington DC: U.S. Bureau of the Census.

Gerald, E., McMillen, M., and Kaufman, S. (1992), *Private School Universe Survey, 1989-90*. E.D. Tabs, NCES 93-122. Washington, DC: National Center for Education Statistics.

Gruber, K. J., Rohr, C. L., and Fondelier, S. E. (1994), *1990-91 Schools and Staffing Survey: Data File User's Manual*, Vol. I. NCES 93-144-I. Washington, DC: National Center for Education Statistics.

Hanson, R. H. (1978), *The Current Population Survey: Design and Methodology*. Technical Paper 40. Washington, DC: U.S. Bureau of the Census.

Jabine, T. (1994), *Quality Profile for SASS*. Technical Report, NCES 94-340. Washington, DC: National Center for Education Statistics.

Johnson, E. G. and King, B. F. (1987), "Generalized Variance Functions for a Complex Sample Survey," *Journal of Official Statistics*, 3, 235-250.

Kaufman, S. and Huang, H. (1993), *1991 Schools and Staffing Survey: Sample Design and Estimation*. Technical Report, NCES 93-449. Washington, DC: National Center for Education Statistics.

Kish, L. (1965), *Survey Sampling*. New York: John Wiley.

Lee, K. H. (1972), “Partially balanced designs for half sample replication method of variance estimation,” *Journal of the American Statistical Association* 67, 324-334.

McCarthy, P. J. (1966), “Replication: An Approach to the Analysis of Data from Complex Surveys,” *Vital and Health Statistics*, Series 2, No. 14. U.S. Department of Health, Education and Welfare. Washington, DC: U.S. Government Printing Office.

McMillen, M., and Benson, P. (1991), *Diversity of Private Schools*. Technical Report. Washington, DC: National Center for Education Statistics.

Mendenhall, W., Scheaffer, R. L., and Wackerly, D. D. (1981), *Mathematical Statistics with Applications*. Boston, MA: Duxbury Press.

Ott, L. (1977), *An Introduction to Statistical Methods and Data Analysis*. North Scituate, MA: Duxbury Press.

Rothhaas, R. (1993), “CPS New and Improved Standard Error Parameters for Labor Force Characteristics.” Unpublished memorandum. Washington, DC: U.S. Bureau of the Census.

Rust, K. (1986), “Efficient Replicated Variance Estimation,” *Proceedings of the American Statistical Association Section on Survey Research Methods*, 81-87.

Salvucci, S., Holt, A. and Moonesinghe, R. (1995), *Generalized Variance Estimates for 1990-91 Schools and Staffing Survey (SASS) - Volume II*. NCES 95-340. Washington, DC: National Center for Education Statistics.

SAS Institute Inc. (1990), *SAS Procedures Guide, Version 6 Third Edition.*

Shah, B.V., Barnwell, B.G., Hunt, P., and LaVange, S.C. (1992), *SUDAAN User’s Guide* (software version 6.00, 1992). Research Triangle Park, NC: Research Triangle Institute.

Short, K. S. and Littman, M. S. (1989), *Transitions in Income and Poverty Status: 1984-85*. Current Population Reports, Series P-70, No. 15-RD-1. U.S. Government Printing Office. Washington, DC: U.S. Bureau of the Census.

SPSS Inc. (1993a), *SPSS Base System Syntax Reference Guide, Release 6.0*.

SPSS Inc. (1993b), *SPSS Statistical Algorithms*, 2nd Edition. New Jersey: Prentice-Hall.

Synectics for Management Decisions, Inc. (1992), *Generalized Variance Estimates for 1987-88 Schools and Staffing Survey (SASS)*. NCES Working Paper Series 94-02. Washington, DC: National Center for Education Statistics.

---. (1995), *Nonresponse Modeling: 1990-91 SASS.* Technical Report (upcoming). Washington, DC: National Center for Education Statistics.

Tomlin, P. (1974), "Justification of the Functional Form of the GATT Curve and Uniqueness of Parameters for the Numerator and Denominator of Proportions." Unpublished memorandum. Washington, DC: U.S. Bureau of the Census.

Valliant, R. (1987), "Generalized Variance Functions in Stratified Two-Stage Sampling," *Journal of the American Statistical Association*, 82, 499-508.

Westat, Inc. (1993), *The WESVAR SAS Procedure Version 1.2*.

Wolter, K. M. (1985), *Introduction to Variance Estimation*. New York: Springer-Verlag.

APPENDICES/Volumn I

Appendix Page