NATIONAL CENTER FOR EDUCATION STATISTICS

Working Paper Series

The Working Paper Series was initiated to promote the sharing of the valuable work experience and knowledge reflected in these preliminary reports. These reports are viewed as works in progress, and have not undergone a rigorous review for consistency with NCES Statistical Standards prior to inclusion in the Working Paper Series.

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

NATIONAL CENTER FOR EDUCATION STATISTICS

Working Paper Series

**Strengths and Limitations of Using
SUDAAN, Stata, and WesVarPC
for Computing Variances from
NCES Data Sets**

Working Paper No. 2000-03   February 2000

Contact: Ralph Lee
             Statistical Standards Program
             E-mail: ralph_lee@ed.gov

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

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

Office of Educational Research and Improvement
C. Kent McGuire
Assistant Secretary

National Center for Education Statistics
Gary W. Phillips
Acting Commissioner

The National Center for Education Statistics (NCES) is the primary federal entity for collecting, analyzing, and reporting data related to education in the United States and other nations. It fulfills a congressional mandate to collect, collate, analyze, and report full and complete statistics on the condition of education in the United States; conduct and publish reports and specialized analyses of the meaning and significance of such statistics; assist state and local education agencies in improving their statistical systems; and review and report on education activities in foreign countries.

NCES activities are designed to address high priority education data needs; provide consistent, reliable, complete, and accurate indicators of education status and trends; and report timely, useful, and high quality data to the U.S. Department of Education, the Congress, the states, other education policymakers, practitioners, data users, and the general public.

We strive to make our products available in a variety of formats and in language that is appropriate to a variety of audiences. You, as our customer, are the best judge of our success in communicating information effectively. If you have any comments or suggestions about this or any other NCES product or report, we would like to hear from you. Please direct your comments to:

> National Center for Education Statistics
> Office of Educational Research and Improvement
> U.S. Department of Education
> 555 New Jersey Avenue, NW
> Washington, DC 20208

The NCES World Wide Web Home Page is
http://nces.ed.gov

February 2000

Foreword

In addition to official NCES publications, NCES staff and individuals commissioned by NCES produce preliminary research reports that include analyses of survey results, and presentations of technical, methodological, and statistical evaluation issues.

The ***Working Paper Series*** was initiated to promote the sharing of the valuable work experience and knowledge reflected in these preliminary reports. These reports are viewed as works in progress, and have not undergone a rigorous review for consistency with NCES Statistical Standards prior to inclusion in the Working Paper Series.

Copies of Working Papers can be downloaded as pdf files from the NCES Electronic Catalog (http://nces.ed.gov/pubsearch/), or contact Sheilah Jupiter at (202) 219-1761, e-mail: sheilah_jupiter@ed.gov, or mail: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, 555 New Jersey Ave. NW, Room 400, Washington, D.C. 20208-5654.

Marilyn M. McMillen
Chief Mathematical Statistician
Statistical Standards Program

Ralph Lee
Mathematical Statistician
Statistical Standards Program

This page intentionally left blank.

Strengths and Limitations of Using SUDAAN, Stata, and WesVarPC for Computing Variances from NCES Data Sets

Prepared by:

Pam Broene
Keith Rust
Westat

Prepared for:

U.S. Department of Education
Office of Educational Research and Improvement
National Center for Education Statistics

February 2000

This page intentionally left blank.

Table of Contents

This page intentionally left blank.

1.   Introduction

Most NCES surveys are based on complex designs involving multi-stage sampling, stratification, clustering, and unequal probabilities of selection. When computing variance estimates from NCES survey data, it is important to incorporate these design features in order to obtain correct standard error estimates. Standard statistical packages such as SPSS and SAS are inappropriate for complex survey data, because they are based on the assumption of independent, identically distributed observations, or simple random sampling with replacement. This has led to the development of software packages that have the capability of estimating variances from complex survey data using specialized methods such as Taylor series linearization and replication.

Three such packages are evaluated in this paper for applicability to NCES surveys. These are Stata Intercooled release 5.0, SUDAAN version 7.11, and WesVarPC version 2.12. Stata and SUDAAN use a linearization approach,[1] WesVarPC employs two replication methods: jackknife and balanced repeated replication (BRR), including Bob Fay's BRR variation. This review covers only the PC versions, and is further restricted in the case of Stata to the survey analysis portion of the software. Resources did not permit all the procedures in each package to be evaluated, nor are cost and availability for different platforms addressed.

1.1   Other Recent Complex Survey Analysis Software Review Papers

Cohen (1996) reviews PC packages Stata release 5.0 (Intercooled), SUDAAN version 7.0, and WesVarPC version 2.02 using the household component of the 1987 National Medical Expenditure Survey (NMES). Means, totals, ratios, and their standard errors are computed, and the three software packages are compared with respect to the programming effort required, documentation, and computational efficiency and accuracy. The article also provides information on software capabilities, available platforms, and cost. The author concludes that none of the packages can be considered "best" or "worst" for all applications and situations, but notes differences in execution times, analytical capabilities, and ease of programming. SUDAAN was found to run the fastest of the three and Stata to be significantly slower, with WesVarPC in between. The author believes the replication approach to be a factor in the slower WesVarPC run times obtained. In his evaluation, 100 BRR replicate weights were

---

[1] RTI's newest release, SUDAAN 7.5, can calculate standard errors using jackknife and BRR replicate weights as well. The user may supply BRR replicate weights but the software does not accept jackknife replicate weights. However, it will create jackknife replicate weights internally so that the jackknife method can be used. This capability was not evaluated here.

created in WesVarPC. It should be noted that the newer 32-bit processor version of WesVarPC version 2.12 will run noticeably faster on a Pentium with 32MB of RAM.

Cohen finds all three packages to produce identical point estimates and identical standard error estimates for SUDAAN and Stata. The WesVarPC standard errors were found "generally convergent" with the linearization estimates. It would be of interest to know if any of the statistics produced for his evaluation are based on NMES variables with missing data, since no inconsistencies are mentioned between the SUDAAN and Stata standard errors (see section 6.1 on this point). A table comparing software capabilities is given, which shows all three packages can calculate descriptive statistics with standard errors and design effects and can perform linear and logistic regression analysis. WesVarPC does estimate standard errors for quantiles, though this is not shown in the table. All three can estimate linear contrasts and perform significance tests. Both WesVarPC and SUDAAN can perform tests of independence for survey data.

Lepkowski and Bowles (1996) catalog eight packages for PC that will calculate standard errors for complex survey data using either replication or linearization: CENVAR, CLUSTERS, EpiInfo, PC CARP, Stata, SUDAAN, VPLX, and WesVarPC. The catalog contains useful information on available platforms, cost, installation, purchasing, input files, and general software capabilities. Although no estimates are presented in the review, the authors found that estimated proportions, standard errors, and coefficients of variation were "similar" across the CENVAR, CLUSTERS, EpiInfo, SUDAAN, and WesVarPC packages. No estimates were produced for Stata, PC CARP, or VPLX. They note that while all packages can compute standard errors using the first stage PSUs under a with-replacement design, only SUDAAN has the capability to handle a without-replacement design where sampling is with unequal probabilities at the first stage (if one has the joint probabilities of selection of the PSUs), and to incorporate a finite population correction factor at subsequent stages of selection when sampling is with equal probabilities. They point out that linearization packages can only produce standard errors for statistics for which a Taylor Series approximation formula has been included in the software, whereas replication methods can be used for virtually any statistic. They consider the WesVarPC documentation the easiest to read and best laid out but warn that users unfamiliar with replication may find the specification of replicate weights confusing. They note that CENVAR, CLUSTERS, EpiInfo, VPLX, and WesVarPC are free, and, with the exception of CLUSTERS, can be downloaded off the Internet.

Weng, Zhang, and Cohen (1995) compare six packages that can handle variance estimation for complex survey data: VPLX, WESVAR/WESREG (apparently mainframe version 1.2), REPTAB, SUDAAN, PC CARP, and STRATTAB. They compute standard errors for proportions, means, ratios,

and regression coefficients using Teacher Survey data from the 1990-91 Schools and Staffing Survey. The packages are compared with respect to computational accuracy, ability to handle NCES survey designs, and treatment of missing data. The versions of each package and the platforms on which the jobs were run are not mentioned in their paper.

A set of BRR replicates created by NCES and a set of jackknife replicate weights generated by VPLX were used in their evaluation. A with-replacement design was assumed for all packages. The authors found all packages to give identical point estimates. Of the three linearization packages (SUDAAN, PC CARP, and STRATTAB), they found SUDAAN and PC CARP to give identical standard error estimates for descriptive statistics but not for regression estimates. The STRATTAB package gave extremely different standard errors that were much smaller than the other two. All three replication packages (VPLX, WESVAR, and REPTAB) produced identical standard errors for a set of BRR replicate weights. The replication and linearization standard errors were found to be similar for the descriptive statistics, but large differences were found for one of the regression models. A few comments are made about the treatment of missing data in the version of WESVAR and WESREG used for their evaluation. This older version of WESVAR and WESREG is no longer supported and the remarks made are not pertinent to the recent PC versions of WesVarPC. In WesVarPC the treatment of missing data is very different from what is described in the authors' paper. Another misunderstanding in their paper seems to be the two-PSU per stratum requirement for creating JK2 (stratified jackknife) and BRR replicate weights within WesVarPC. This does not mean that replicate weights created outside WesVarPC cannot be used to obtain correct standard errors using the software, even when there are not exactly two PSUs per stratum. Although the current version of WesVarPC requires two PSUs per stratum in order to create replicate weights within the software, the next release, due in mid-1998, will allow more than two PSUs per stratum.

The current evaluation attempts to provide a more indepth and comprehensive evaluation of each software's estimation procedures, including the treatment of missing data, for a broader range of statistics and sample designs than each of the above three papers. Test statistics and data processing issues are mentioned briefly. A comparison of standard errors calculated using bootstrap replicate weights as well as jackknife and BRR replicate weights is made with linearization standard errors. Recommendations are made specifically for NCES surveys.

1.2   Description of NCES Surveys

The 1988 National Education Longitudinal Study (NELS:88) baseline survey and the 1993-94 Schools and Staffing Survey (SASS) were chosen for the evaluation because both contain many of the complex sample design and estimation features common to NCES surveys. These features include multistage sampling, stratification, sorting, clustering, unequal probabilities of selection, high sampling fractions, nonresponse adjustments, and poststratification, all of which affect the estimation of variance.

1.2.1   NELS:88 Base year

The NELS:88 is a survey of the school-related experiences and accomplishments of a nationally representative sample of eighth graders. The target population consisted of eighth graders in all public and private schools containing eighth grades in the fifty states and the District of Columbia (Spencer et al., 1990). A sample of schools was drawn, and students were sampled within each school. School data were collected from the administrator at each sampled school. The sampled students were administered achievement tests in different academic subjects and were asked to respond to a questionnaire about their school and home life. Data on the students and on teacher practices and curriculum were also collected from the parents and teachers of the sampled students for use in analyzing student data.

Students were selected using a two-stage stratified, clustered design. At the first stage, schools were sampled from a list frame using probability proportional to size systematic sampling within 28 strata. Each school's probability was proportional to its eighth grade enrollment. Superstrata were formed by school type (public, Catholic, NAIS,[2] other private) and region. Substrata within primary strata were formed by urban status and by percent minority class (public schools only). Prior to sampling, schools were sorted within stratum by eighth grade enrollment.

The number of public schools sampled in each primary stratum was determined by proportional allocation using the total estimated eighth grade enrollment of all schools in the stratum. Private schools were oversampled relative to public schools to increase their sample size for analysis. Of the 1,734 schools sampled, 1,035 responded and were eligible. Students were stratified by Asian, Hispanic, and all other within each school, and were sampled with equal probabilities within each stratum. A maximum of 24 "all other" students per school were selected. These were supplemented with

---

[2] National Association of Independent Schools

an oversampling of Asian and Hispanic students. The total sample size of students was approximately 26,400, of which 24,599 students have data for analysis.

Both the school and student baseline sample files were used in this software evaluation, but only student estimates are discussed. The important features of the sample design to reflect in the variance estimates are the stratification and sorting of schools, the clustering of students within schools, and the unequal school and student weights. The school sampling fractions within each stratum were small so no finite population correction factor is needed and with-replacement sampling can be assumed for variance estimation.

1.2.2   1993-94 SASS

The 1993-94 SASS is a survey of public and private elementary and secondary schools in the United States. It has public school district (local education agency—LEA), teacher, administrator, library, and student components as well, all conducted during the same school year. Additionally, the Teacher Follow-up Survey is conducted on a subsample of teachers 1 year later. Data collected include school and teacher characteristics, school operations, program and policies, teacher demand and supply, and the opinions of teachers and administrators about policies and working conditions. These files can all be linked for analysis (Abramson et al., 1996). For the software evaluation, only the school and teacher survey files were used.

Schools were selected from public and private school list frames, with the addition of a small area sample of private schools to improve coverage. With the exception of the area sample, schools were selected as a one-stage stratified design using probability proportional to size systematic sampling. Each school's probability of selection was proportional to the square root of the number of its teachers.

The variables used to stratify public schools were state, school level (elementary, secondary, and combined), whether BIA-run (Bureau of Indian Affairs) or not, and the percentage of Native American students. The sample was allocated to strata using optimal allocation for estimating the total number of teachers. Within each non-BIA stratum, schools were sorted by LEA metro status, ZIP Code (recoded to keep schools in an LEA together within the stratum), LEA ID, highest grade, percentage of minority students, and enrollment prior to sampling. All 176 BIA schools were taken with certainty. In addition, schools in the high percentage Native American enrollment stratum were oversampled. A total of 9,956 public schools were sampled, of which 8,767 responded and were eligible.

Private schools on the list frame were stratified by school association membership, grade level, and Census region. The sample was allocated to strata proportionally but in such a way as to balance the goals of producing private school association and national level estimates, and public versus private sector comparisons. Within each stratum, schools were sorted by state, highest grade, urban status, ZIP Code (first two digits), and enrollment. Of the 19 associations, 4 were small enough that all schools were taken into the sample with certainty. In another four associations, the school sampling fraction was about 0.5 or greater.

For the area sample, PSUs consisting of a single county, independent city, or cluster of contiguous counties or independent cities were first sampled. Prior to sampling, the PSUs were stratified by Census region, MSA status, and percentage of minority students into 16 strata. A minimum of two PSUs were allocated to each stratum for variance estimation purposes, with 26 additional PSUs allocated to more nearly approximate a uniform sampling fraction of PSUs from each stratum.

The 1993-94 SASS area frame was designed to produce approximately 50 percent overlap with the 1991 SASS; thus 58 or half of the PSUs selected in 1991 SASS were retained for the 1993-94 survey. Eight PSUs that had been selected with certainty in 1991 SASS were again taken with certainty in the 1993-94 survey. An additional 58 PSUs were selected independently for the 1993-94 survey using probability proportional to size systematic sampling, with PSU probabilities proportional to the square root of their 1988 population projections.

Within each of the sampled PSUs, a list was compiled of all private schools and matched with the private school list frame already in use. From the nonmatching schools, a sample of schools was drawn. A total of 3,162 private schools were sampled from the list frame and 153 (after unduplication) from the area frame. The total number of responding, eligible private schools available for the software evaluation was 2,585.

Overall, the teacher sample is a two-stage stratified sample; the exception is the teacher sample from schools in area sample, which is a three-stage sample. Within the sampled schools teachers were stratified into one of five teacher types in the following hierarchical order: 1) American Indian, Aleut, or Eskimo; 2) Asian or Pacific Islander; 3) Bilingual/ESL; 4) new; and 5) experienced. Then they were sampled systematically with equal probabilities within strata. Teachers in the first three strata were oversampled in both public and private schools; new teachers in private schools were also oversampled. The maximum number of sampled teachers per school was set at 20. The total number of teachers sampled was 68,284, of which 55,477 were eligible respondents to the teacher questionnaire. Of these, 51,604 teachers also were in responding schools.

The sample of LEAs consisted of the LEAs associated with the sampled public schools, plus a few others that were sampled directly because they hire teachers but are not associated with any schools. Consequently, each LEA's probability of selection is a function of the probabilities of selection of its constituent schools. The school administrator sample consisted of the administrator at each sampled public and private school. The library sample consisted of the libraries associated with a subsample of the public school sample. The student sample was obtained by subsampling three teachers from each school in a subsample of schools, then subsampling two students from each teacher. A student records questionnaire was sent to the school administrator. LEAs, administrators, and librarians were administered separate questionnaires.

The aspects of the 1993-94 SASS sample design that should be taken into account in estimating variances are the stratification of schools and PSUs, the sorting of schools within explicit strata at the first stage of selection, the clustering of area sample schools within PSUs, the clustering of teachers within schools, unequal school and teacher weights, high sampling fractions in some of the states and private school strata, and the selection of BIA schools and schools in some private school associations with certainty.

2.   How This Software Evaluation Was Conducted

The criteria used to evaluate the three software packages are as follows:

- How well each package can handle the sample design features which are relevant for variance estimation;

- Analysis capabilities;

- How missing data are handled;

- Accuracy of computational algorithms and test statistics;

- Ease of use (including programming, inputting and outputting data);

- Execution time; and

- Documentation.

A range of statistics was computed using each software package, including means, medians, totals, differences in means, proportions, odds ratios, and logistic regression coefficients. The resulting estimates with their standard errors, design effects, and test statistics were compared across packages. A

sample of these computations is presented in tables 1a – 1d and figure 1. The treatment of missing data by each package is summarized in table 2. Table 3 contains execution times for several runs for each package. Table 4 summarizes the main survey analysis features available with each one. The criteria and tables are discussed in the remainder of the report.

3.   File Preparation

Regardless of the software package chosen, the preparation of input files for variance estimation can be a very time-consuming task, depending on the availability of sample design variables and replicate weights on the analysis files. To create replicate weights for use with WesVarPC, or sample design variables for SUDAAN or Stata, the user must have a complete understanding of the sample design, the details of which may not be available in publicly released documents. At NCES, usually a public use analysis file is released with a particular variance estimation method in mind. For example, the 1993-94 SASS files contain bootstrap and BRR replicate weights that are designed to be used with software that can accommodate the BRR replication method; thus the user would presumably have no need for stratification variables. However, a SUDAAN or Stata user would require these. The NELS:88 files contain the stratum codes that SUDAAN or Stata need (masked for confidentiality reasons), but no replicate weights.

The creation of replicate weights can present a major roadblock to analysts who are unfamiliar with replication methods or the sample design of the survey or to those who do not have access to the required sample design variables. Replicate weights may be created in WesVarPC by specifying the variables that contain the variance strata (VARSTRAT) and PSU codes (VARUNIT) in the PREP step. For this process to work properly, in the current version of WesVarPC there must be exactly two PSUs or half-samples per variance stratum; i.e., VARUNIT must equal 1 or 2 for each level of VARSTRAT.[3] Outside programming is necessary to create the variance strata and PSU variables when there are more than two PSUs per stratum, when there are strata with only one PSU in them, or when there are certainty PSUs. This may require grouping PSUs to form half-samples which function as the VARUNIT in WesVarPC, or it may require collapsing strata. The variance strata should be created in such a way that there are sufficient degrees of freedom for stable variance estimation for the domains of interest. This may involve splitting and combining the original sampling strata to form variance strata.

---

[3] A new version of WesVarPC, due for release in mid-1998, will create jackknife replicate weights when there are more than two PSUs per stratum.

All three packages will accept ASCII files, but the layout of each record must be specified either in an external file or keyed in when the file is read in. Both SUDAAN and WesVarPC will accept SAS version 6.02 through 6.04 files (PC SAS files for DOS), though they will not accept later SAS file versions. SAS for Windows version 6.08 and later files can be easily converted to version 6.04 by specifying the V604 engine in the LIBNAME statement when outputting the file. WesVarPC will also accept SPSS for Windows, SAS Transport, and DBF files. Stata comes with an easy-to-use Stat/Transfer program that converts SAS Transport files into Stata files (.dta files). The Stat/Transfer program will also convert the following types of files into Stata files that can be read by the software: Lotus, Excel, Clipper, Alpha Four, CRUNCH, dBase II – IV, FoxBase, Gauss, Paradox, Quattro Pro, S-Plus, SPSS Export, and Systat. Stata can also read comma- or tab-delimited ASCII files created by a spreadsheet or database program with its insheet command.

In this evaluation, the first step was to create SAS version 6.04 files from the restricted use ASCII files. Additional sample design variables were obtained from NORC for the NELS:88 school file and the Census Bureau for the 1993-94 SASS private school sample. For the NELS:88 school and students sample files, a set of 100 JK2 and 100 BRR replicate weights were created for use with WesVarPC. These replicate weights were created outside of WesVarPC using Westat's internal SAS WESWGT programs, because the NELS:88 sample design did not consist of two PSUs per stratum and because some variance strata (after splitting the original sampling strata) consisted of three PSUs. The replicate weights reflect the sort order of selection of schools within superstrata and substrata. The analysis file containing replicate weights was then imported in the PREP step of WesVarPC to create the .VAR file that WesVarPC uses.

For SUDAAN and Stata, variance strata (VARSTRAT) and variance PSU (VARPSU) variables were created to communicate the sample design to the software for the 1993-94 SASS school and teacher and NELS:88 school and student files. In general, for both 1993-94 SASS and NELS:88, the variance strata were set equal to the school strata and the variance PSUs were set equal to the school ID. Variance strata containing only one school were collapsed with the adjacent stratum. For the area sample schools in 1993-94 SASS, the variance strata were set equal to the strata used to select the geographic PSUs, and the variance PSUs were set equal to the geographic PSU code. For SUDAAN, these variables must be numeric and the input files must be sorted by variance strata and variance PSU in the order they appear on the NEST statement. For Stata, SAS transport files were created from the SAS version 6.04 files and converted to Stata files using Stat/Transfer. In the Stat/Transfer run, the lengthy 1993-94 SASS school and teacher control numbers were set to "double precision" numeric data types instead of the "float" default. This extra precision was needed so that Stata would read in the entire field and count the number of PSUs correctly.

All three packages have flexible recoding and subsetting capabilities, though Stata's replace, generate, and egen (extensions to generate) commands are the most powerful of the three for recoding existing variables and creating new variables quickly. All three support the creation of labels for variables and levels of categorical variables.

4. Handling of Complex Survey Design Features

4.1 WesVarPC

In replication methods, the sample design is built into the replicate weights. Thus, if the replicate weights are contained on the input file, it is only necessary to identify them by variable name to WesVarPC and to select the method used to create them in the PREP step. All the components of variance corresponding to the different stages of sampling will be included in the variance estimates. The effects of stratification, clustering, certainty selections (if any), and unequal weighting will be accounted for if the replicates were created properly. It is possible to incorporate an fpc into jackknife and BRR replicate weights in WesVarPC if sampling at the first stage is with equal probabilities, but this condition was not present with NELS:88 or SASS since schools were selected with probability proportional to size. However, the bootstrap replicate weights account for the high sampling fractions in SASS.

4.1.1 NELS:88

WesVarPC was able to handle the most important aspects of the sample design for both of the NELS:88 baseline surveys examined, primarily because these were built into the replicate weights, which were created outside of WesVarPC. To create replicate weights in WesVarPC, additional programming would be required before importing the school or sample file in the PREP step. A variance stratum variable (VARSTRAT) must be created with exactly two half-samples or PSUs per VARSTRAT. A VARUNIT variable must also be created that identifies the half-sample or PSU.

In WesVarPC, it is possible to reflect raking and poststratification adjustments to the weights in the variance estimates if the raking dimension or poststratification cell variables and control totals are available on the file, and also the replicate weights prior to the raking/poststratification. However, the raking dimension variables and control totals were not available on the NELS:88 files, so this was not done in creating replicate weights for this evaluation.

4.1.2   1993-94 SASS

The 1993-94 SASS files already have 48 bootstrap replicate weights, which NCES created for their public and restricted use files. These replicate weights incorporate the aspects of the sample design and weighting that are relevant for variance estimation, so the user does not have to worry about this in setting up WesVarPC runs. Bootstrap replicate weights were created for the school, administrator, teacher, and LEA samples because it was considered important to reflect the high sampling fractions in some of the school sampling strata (and in the LEA sample) as well as the systematic sampling in the variance estimates. BRR and Jackknife replicate weights cannot be responsive to this requirement when the first stage of selection is with unequal probabilities. The creation of these replicate weights is described in Abramson et al. (1996) and in Kaufman (1993, 1994). WesVarPC will compute the correct variance using bootstrap replicate weights if the BRR method is selected after identifying the replicate weight variables in the PREP step because the formula for calculating the variance from the replicate estimates is the same for both methods. In the creation of bootstrap replicate weights by NCES, all nonresponse and poststratification weighting adjustments were replicated so that WesVarPC will correctly include the impact of these adjustments in its variance estimates. Although poststratification reduces the variance, extra sampling variability arises from weighting adjustments for poststratification and nonresponse, because different samples would give rise to different weighting adjustment factors.

4.2   SUDAAN

SUDAAN accommodates six types of sample designs:

- **WR:** multi-stage design, with replacement sampling at the first stage, PSUs sampled with equal or unequal probabilities of selection.

- **STRWR:** single-stage design, with replacement sampling with equal or unequal probabilities of selection.

- **WOR:** multi-stage design, without replacement sampling assumed at the first stage, equal probabilities of selection at the first stage and subsequent stages where sampling is without replacement.

- **STRWOR:** single-stage design, without replacement sampling with equal probabilities of selection.

- **UNEQWOR:** single or multi-stage design, without replacement sampling at the first stage, PSUs sampled with unequal probabilities of selection, sampling with or without replacement with equal probabilities at subsequent stages.

- **SRS:** simple random sampling with replacement.

The STRWR and STRWOR options are for one-stage stratified samples where each record is a PSU. WR is used when there are two or more stages of sampling; this design option treats the PSUs as "ultimate clusters." The three without replacement designs (WOR, STRWOR, UNEQWOR) calculate a finite population correction factor (fpc) for the first stage of sampling. These three designs require that the total number of PSUs in the population for each stratum be contained in a variable placed on every record, or else be available on a user-created PSU-level PSUDATA file that SUDAAN accesses during execution. In addition, UNEQWOR requires the joint probabilities of selection for each pair of PSUs.

Many survey designs can be adequately described by one of SUDAAN's six design options. In particular, many designs are covered approximately by the WR option. However, there may be cases where it is inappropriate to use the linearization method of variance estimation. For example, the multiplicity sampling used to obtain the sample of LEAs in the 1993-94 SASS Teacher Demand and Shortage Survey does not fit a linearization scheme. Two-phase or double sampling is another example which neither linearization nor replication methods can fully account for.

4.2.1   NELS:88

SUDAAN was able to handle the NELS:88 sample design easily since the school and student samples fit well into the STRWR and WR design options. It was only necessary to collapse beforehand one superstratum (SSTRATID) that contained a single responding school with another superstratum. In SUDAAN it is possible to account for the sorting of schools within superstrata prior to sampling by listing the primary sort variables after SSTRATID in the NEST statement and using the PSULEV option to indicate which variable corresponds to the PSU. This was done for one set of runs, but it made very little difference in the particular set of variance estimates examined. If a statistic is highly correlated with the sort variables however, including these variables on the NEST statement should have the effect of reducing the variance.

4.2.2   1993-94 SASS

SUDAAN was less successful in taking account of the Schools and Staffing Survey design, primarily because of the high school sampling fractions and because of the way the LEA sample was drawn. However, most aspects of the sample design relevant for variance estimation could be taken in account for schools and teachers (and by extension, administrators, librarians, and students). In the case of the LEA sample, linearization variance estimates are not possible because LEAs were not sampled directly but were obtained through the sampled schools. For the School and Teacher Surveys, to reflect the high sampling fractions in some school strata, the DESIGN=UNEQWOR option would need to be specified because schools were sampled with unequal probabilities. This design option requires giving SUDAAN the joint probabilities of selection for each pair of PSUs on the JOINTPROB statement. However, NCES does not have the joint probabilities of selection, nor would they be easy to compute. Such computations, in fact, would also require the variance stratum population counts to be included on the TOTCNT statement. These population counts are not currently on the restricted use file, though presumably they could be obtained from NCES. Because of these difficulties, the DESIGN=WR option was used in this evaluation for teacher-level estimates from the teacher file, and DESIGN=STRWR was used for school-level estimates from the school file.

Another issue for the Schools and Staffing Survey is the treatment of certainty PSUs, which would include the BIA schools and schools in the four private school associations from which all schools were taken. (The eight certainty PSUs in the private school area sample are not a concern here, because schools from these PSUs were added to the list frame prior to sampling.) To reflect the certainty schools, each certainty school must be recoded as a separate variance stratum and teachers in these schools must be recoded as PSUs. In other words, one would set VARSTRAT=school ID and VARPSU=teacher ID for each certainty school. There must be at least two teachers for each certainty school, or SUDAAN will stop executing and generate an error message.[4] It is probably not a good idea to try to incorporate the teacher stratification in creating VARSTRAT for the certainty schools, since many of the teacher strata are small and would need to be collapsed within a school to meet the two PSU per stratum requirement. Note that the certainty schools cannot be recoded this way on the school file because there is only one record per school, and SUDAAN will not execute when there is only one PSU per stratum for a with-replacement design option. However, certainty schools represent only 1.8 percent and 7.1 percent of the public and private schools samples, respectively. For this reason, the extra recoding may not be considered worth the effort. The effect of not taking into account the certainty schools is that the

---

[4] This error message was obtained even when the MISSUNIT option was specified in the NEST statement, contrary to the manual's direction.

responding certainty schools are treated as a random sample of the school universe in each stratum, meaning they will contribute a small amount to the variance.

SUDAAN can reflect the effect of poststratification in its variance estimates if the poststratification cell variables and control totals are available on the analysis file.

4.3   Stata

Stata handles complex survey data through its svy commands (svyset, svydes, svymean, svytotal, svyprop, svyratio, svylogit, svymlog, svyolog, svyprobt, svyoprob, svylc, svytest).[5] Although it accommodates fewer types of sample designs than SUDAAN, it can handle the equivalent of SUDAAN's DESIGN=WR, STRWR, and STRWOR design options.   It employs the "between PSU" variance estimator or "ultimate cluster" sampling model. Unlike SUDAAN, it does not handle second or third stage sampling within PSUs. It does not have the capability to incorporate an fpc after the first stage of sampling or with unequal probabilities of selection at the first stage, though it will correctly calculate an fpc for the first stage of sampling if sampling was with equal probabilities. To do this, the analyst must specify either the stratum total number of PSUs or the stratum sampling fraction on the svyset fpc command.

Note that although Stata can calculate standard errors based on bootstrap replication with its BSTRAP command, its bootstrapping procedure assumes the original sample was selected using simple random sampling. Therefore the BSTRAP command is not appropriate for complex survey data.

4.3.1   NELS:88

Stata was able to estimate correct standard errors for the NELS:88 school and student files using the same VARSTRAT and VARPSU variables that were created for the SUDAAN input file with the DESIGN=WR option. One difference between the two packages is that Stata permits just one variable to be specified as the stratification variable on the svyset strata statement. This means one could not as easily include the effects of the school order of selection within superstrata as one can with SUDAAN, since a new stratification variable would have to be created which incorporated the sort

---

[5] The svymlog, svyolog, and svyoprob commands are not included with the Stata software diskettes shipped to new users nor do they appear in the version 5.0 manuals. They must be downloaded from the Stata web site.

variables. However, it does allow the collapsing of strata within the program; this does not have to be done in SAS beforehand as with SUDAAN.

4.3.2   1993-94 SASS

Stata can calculate approximately correct standard errors for the 1993-94 SASS School, Teacher, Administrator, Library, and Student Surveys using the same method as SUDAAN's DESIGN=WR option. The standard errors are identical to those calculated by SUDAAN with the exception of estimated totals when there is missing data. To account for the certainty schools on the teacher file, the certainty schools would have to be recoded to separate variance strata and their teachers to variance PSUs as described above for SUDAAN. As with SUDAAN, the certainty schools cannot be accounted for on the school file by making them their own variance stratum, since a variance cannot be estimated for strata containing only one PSU. To ensure that they contribute zero to the variance for school level estimates, the certainty schools could all be placed in a single stratum, and a sampling fraction of one specified in the svyset fpc command for that stratum. A sampling fraction of zero would have to be specified for the other strata so that Stata would compute an fpc equal to one for them. This effectively prevents Stata from computing an incorrect fpc in the noncertainty strata where schools are sampled with unequal probabilities.

5.   Ease of Use

5.1   Documentation

The manuals for all three packages provide clear instructions and examples on how to use the software. The Stata manuals are especially complete and easy to read. Helpful cross-references under each topic made learning the software much easier. E-mail requests for technical support to Stata's Web site were answered in a timely fashion. The WesVarPC manual is also easy to read, but one should not expect very much guidance from it on how to create replicate weights. Both Stata and WesVarPC manuals provide the formulas used by the software, but the SUDAAN manual is uneven in this regard. Although the test statistics are documented in the manual, most of the formulas for variances and design effects (with the exception of the MULTILOG procedure) can only be found in a separate document (Shah et al., 1996).

5.2   **Programming**

Of the three packages, WesVarPC is probably the easiest to learn and use once replicate weights are on the input file. A user does not need a background in sample design to perform data analysis, nor does he/she need to learn much in the way of syntax. It is menu driven, and programs are built by clicking on options, variable lists, and mathematical symbols on screens. As an alternative to the point-and-click method of creating table requests, it does have a utility for creating large numbers of table requests in a single file. This file can be edited as a text file and submitted in batch mode. The main advantage of this feature is that table settings, variables, and functions can be reused for many tables without having to re-enter them each time using the point-and-click graphical interface. Alternatively, each table in a single request file can be customized; i.e., the statistics calculated and the output settings for each table in a single request can be different. The software also includes an Excel macro Table Generator to produce user-defined tables using output from WesVarPC's table requests. This is helpful since WesVarPC's table request output is in list format, but the macro may not be useful for all table formats.

SUDAAN is the most difficult to learn and set up initially, but once this hurdle is overcome it is a very powerful and economical package in terms of the number of statements required. Its commands are reminiscent of SAS. The tabular output is very readable and many output options exist to control the appearance of the table and the statistics that are output. A PROC PRINTTAB procedure exists for further processing summary statistics from other SUDAAN procedure output. SUDAAN can also produce ASCII output files which can often be directly imported into spreadsheets such as Excel with a minimum of SAS programming intervention.[6] Many tables can be requested in a single statement, but the drawback of this is that the table settings, statistics calculated, and output must be the same for every table. In addition, it is necessary to list all categorical variables and the number of levels for each in a SUBGROUP and LEVEL statement, which can be an error-prone task when many variables are involved. SUDAAN also requires the levels of categorical variables to be recoded to consecutive integers beginning with 1. A careful reading of the manual is required to understand and correctly use SUDAAN's syntax and output options, especially the use of its special variables such as _ONE_.

Stata is easy to learn, but can require many more programming statements than WesVarPC or SUDAAN to accomplish the same thing. This is particularly true in generating tables. Standard errors for proportions can be calculated in Stata using either the svyprop or the svymean commands, but design effects and confidence intervals are not available with svyprop. Proportions may be calculated as means

---

[6] Note that the PC stand-alone version of SUDAAN does not output SAS files, but a SAS-callable version available for VAX VMS and IBM MVS will output SAS version 6 files.

of 0-1 indicator variables using the svymean command, which does produce design effects and confidence intervals, but the user is responsible for creating these or recoding existing variables. For a multilevel categorical variable, a separate indicator variable must be created for each level of the variable. In Stata, multiple table requests cannot be submitted with a single statement as in SUDAAN, but require separate svyprop or svymean statements for each crosstab. Another inconvenience is that the marginals and weighted totals are not a byproduct of the table request but must be produced in separate commands. In WesVarPC or SUDAAN, getting weighted totals, proportions, unweighted cell counts, confidence intervals, and design effects for categorical variables in one request is simply a matter of specifying these statistics in the output. Output from Stata svy commands can be saved for access by other software, but this is not as easy as in WesVarPC or SUDAAN.

All three packages compute standard errors for proportions and for continuous statistics such as means, totals, ratios, and differences in these quantities. In addition, SUDAAN and WesVarPC can calculate standard errors for quantiles and odds ratios. WesVarPC's COMPUTE and FUNCTION screens allow the user great flexibility in calculating standard errors for statistics that may not be available in SUDAAN or STATA.

All three packages will do regression and logistic regression. SUDAAN also does polychotomous logistic regression, survival analysis, and loglinear models, but Stata is the only package that will estimate probit models. However, in fitting regression models Stata requires the user to create dummy variables for the different levels of categorical regressors, whereas these are generated automatically in WesVarPC and SUDAAN. The estimated model coefficients and their standard errors can be output for use by other software in all three packages, but this is not at all straightforward in Stata. Stata does have some very useful commands (predict, summarize, and plot) that enable the user to interactively compute the predicted values and residuals, then summarize and display them in plots.

6.   Computational Accuracy

6.1   Point Estimates and Standard Errors

Table 1a contains selected estimates from the NELS:88 baseline student sample and tables 1b – d and figure 1 contain selected estimates from the 1993-94 SASS teacher and school samples, so that computational results from WesVarPC, SUDAAN, and Stata may be compared. As can be seen in Tables 1a – d, all three packages produced identical point estimates, except that Stata's medians differed slightly in the first decimal place due to the different algorithm it uses for obtaining quantiles.

WesVarPC and SUDAAN both estimate the median and confidence interval of the median using the Woodruff method (Woodruff, 1952) on either grouped or ungrouped data. However, WesVarPC uses replication and SUDAAN uses linearization to estimate the standard error of the cumulative distribution function. These methods are documented in the WesVarPC manual and in Shah et al. (1996). The difference between the three packages is that WesVarPC and SUDAAN interpolate between two order statistics to get the estimated quantile, whereas Stata simply takes the arithmetic average. Also, Stata does not provide the confidence interval or standard error for quantiles of weighted survey data.

Tables 1a – d show that standard error estimates were identical for Stata and SUDAAN except for certain estimated totals of variables containing missing data. SUDAAN appears to be setting missing values to zero and including these observations in its calculation of the standard error. This was confirmed by setting missing values to zero for several variables and computing the standard errors in Stata. (Normally, Stata excludes observations with missing values from a calculation.) Thereafter, the standard errors matched SUDAAN's exactly. Examples are the estimated student total for "Father's highest level of education" in "Other Private" schools in table 1a, and estimated public school student totals for ESL, Chapter 1, and Kindergarten in table 1c. SUDAAN users should be warned about this aspect of its treatment of missing data, since it is not mentioned anywhere in the SUDAAN documentation.

A bug was discovered in Stata when computing standard errors for estimated public school student totals for ESL, Chapter 1, and Kindergarten. The problem occurs when Stata checks for strata with fewer than two PSUs when subpopulation estimates are requested. The svy commands do not automatically drop strata with no members of the subpopulation before checking whether there is only one PSU per stratum. As a result, the user receives an error message about "Stratum with one PSU detected" and the program aborts. The public schools all had at least two schools per stratum on the input file with reported data (a small amount of collapsing of strata was done beforehand to ensure that this condition was met for just these three variables), although there were numerous private school strata with only one school in them. To force the Stata program to run, it was necessary to subset the file on just the public schools when requesting public school estimates. This produced the standard errors for these three variables in table 1c, which do not appear to be correct. Stata Corporation acknowledges this problem and plans to fix it in the near future.

Table 1a. **Estimates, standard errors and design effects for WesVarPC, SUDAAN, and Stata: NELS:88 student file**

Statistic	Variable	Estimate	WesVarPC: JK2		WesVarPC: BRR		SUDAAN		Stata			
			S.E.	DEFF	S.E.	DEFF	S.E.	DEFF	S.E.	DEFF1[1]	DEFF2	MEFF
Means	Reading score											
n=18,403	Public......................	9.92	0.083	3.38	0.084	3.4	0.10	4.98	0.10	4.53	4.98	4.46
n=2,525	Catholic	12.26	0.20	3.29	0.21	3.59	0.21	2.57	0.21	3.57	2.57	3.56
n=2,488	Other private............	13.69	0.28	6.34	0.26	5.48	0.26	2.3	0.26	5.36	2.3	5.77
	Math score											
n=18,403	Public......................	15.55	0.17	4.17	0.18	4.59	0.20	6.27	0.20	5.63	6.27	5.57
n=2,525	Catholic	18.15	0.48	5.31	0.49	5.49	0.50	4.12	0.50	5.73	4.12	5.64
n=2,488	Other private............	22.44	0.58	7.73	0.56	7.25	0.55	2.9	0.55	6.77	2.9	7.09
Medians	Reading score											
n=18,403	Public......................	9.75	0.26		0.26		0.18		--[2]	--	--	--
n=2,525	Catholic	13.08	0.29		0.30		0.21		--	--	--	--
n=2,488	Other private............	14.56	0.35		0.34		0.42		--	--	--	--
	Math score											
n=18,403	Public......................	14.40	0.21		0.23		0.36		--	--	--	--
n=2,525	Catholic	17.57	0.68		0.70		0.74		--	--	--	--
n=2,488	Other private............	24.00	0.72		0.69		0.83		--	--	--	--
Totals	Total students with father highest level of education BA+											
n=15,922	Public......................	586,650	17,306	--[3]	18,185	--[3]	19,437	6.86	19,437	6.38	6.86	6.41
n=2,157	Catholic	74,565	6,531	--	6,843	--	6,511	4.78	6,511	10.45	4.78	10.49
n=2,371	Other private............	74,511	7,408	--	8,138	--	8,005	7.21	7,916	44.82	7.07	57.04
	Total students with mother highest level of education BA+											
n=16,875	Public......................	485,657	13,249	--	13,179	--	14,806	4.50	14,806	4.14	4.50	4.17
n=2,248	Catholic	60,227	4,928	--	5,032	--	5,440	4.09	5,440	7.91	4.09	7.95
n=2,381	Other private............	60,736	5,890	--	6,368	--	6,429	5.67	6,429	27.23	5.67	29.84
Proportions	Percent with mother/ female guardian still living											
n=19,058	Public......................	99.33	0.07	1.44	0.07	1.34	0.07	1.48	0.07	1.33	1.48	1.29
n=2,545	Catholic	99.31	0.15	0.83	0.15	0.78	0.17	0.73	0.17	1.02	0.73	1.00
n=2,524	Other private............	99.78	0.07	0.58	0.07	0.59	0.07	0.26	0.07	0.62	0.26	0.31
	Percent with father highest level of ed. BA+	(%)										
n=15,922	Public......................	26.73	0.66	3.50	0.66	3.56	0.73	4.87	0.73	4.32	4.87	4.34
n=2,157	Catholic	38.87	2.65	6.37	2.75	6.86	2.32	3.54	2.32	4.86	3.54	4.88
n=2,371	Other private............	62.62	3.33	11.25	3.28	10.87	2.93	3.57	2.93	8.69	3.57	11.07
	Percent with mother highest level of ed. BA+											
n=16,875	Public......................	20.90	0.50	2.50	0.47	2.23	0.54	3.35	0.54	2.97	3.35	3.00
n=2,248	Catholic	30.15	2.14	4.87	2.18	5.05	2.05	3.25	2.05	4.46	3.25	4.49
n=2,381	Other private............	50.52	2.65	6.67	2.62	6.56	2.45	2.36	2.45	5.71	2.36	6.26

[1]With SRSSUBPOP option.

[2]Stata does not calculate standard errors for quantiles of survey data.

[3]WesVarPC does not calculate design effects for totals.

Table 1b. **Estimates, standard errors and design effects for WesVarPC, SUDAAN, and Stata: 1993-94 SASS teacher file**

Statistic	Variable	Estimate	WesVarPC		SUDAAN		Stata			
			S.E.	DEFF	S.E.	DEFF	S.E.	DEFF1[1]	DEFF2	MEFF
Proportions	Percent with bachelor's degree									
n=44,041	Public	98.98	0.06	1.74	0.07	2.23	0.07	2.18	2.23	1.75
n=7,563	Private................................	93.09	0.46	2.44	0.50	2.55	0.50	3.00	2.55	3.65
	Percent with master's degree									
n=44,041	Public	46.33	0.36	2.25	0.41	3.12	0.41	3.04	3.12	3.07
n=7,563	Private................................	31.99	0.67	1.56	0.72	1.55	0.72	1.82	1.55	1.80
	Percent ever threatened by a student									
n=44,041	Public................................	25.26	0.39	3.46	0.39	3.56	0.39	3.47	3.56	3.41
n=7,563	Private................................	6.82	0.39	1.79	0.49	2.47	0.49	2.90	2.47	2.74
Totals	Total number teachers with master's degree									
n=44,041	Public................................	1,103,352	12,554	--[3]	12,022	4.18	12,022	4.51	4.18	4.56
n=7,563	Private................................	108,649	3,277	--	3,037	1.68	3,037	2.78	1.68	2.75
	Total number teachers with bachelor's degree									
n=44,041	Public................................	2,357,371	20,770	--	14,293	12.29	14,293	157.00	12.29	122.00
n=7,563	Private................................	316,193	5,299	--	4,920	1.64	4,920	25.00	1.64	30.00
Logistic Regression Coefficients	Dependent variable: "Would become a Teacher again " 1=Certainly, Probably 0=Neutral,Probably not, Certainly not Independent variables:									
	Intercept.............................	1.40	0.094	--[3]	0.12	2.41	0.12	--[2]	2.41	2.42
	Age	-0.018	0.002	--	0.002	2.43	0.002	--	2.43	2.36
	Race/ethnicity									
	American Indian/ Alaska Native	-0.11	0.15	--	0.17	1.95	0.17	--	1.95	3.84
	Asian/Pacific Islander........	-0.091	0.10	--	0.12	1.20	0.12	--	1.20	2.14
	Black, not Hispanic	0.061	0.080	--	0.10	2.69	0.10	--	2.69	2.33
	White, not Hispanic	-0.12	0.069	--	0.084	2.71	0.084	--	2.71	2.51
	Hispanic.............................	0								
	Verbal abuse of teachers a problem									
	Serious	-0.74	0.066	--	0.068	2.40	0.068	--	2.40	2.51
	Moderate............................	-0.50	0.051	--	0.053	2.55	0.053	--	2.55	2.60
	Minor	-0.21	0.043	--	0.044	2.56	0.044	--	2.56	2.52
	Not a problem....................	0								
	Poverty is a problem at this school									
	Serious	-0.001	0.063	--	0.069	2.95	0.069	--	2.95	2.86
	Moderate............................	0.059	0.052	--	0.059	2.84	0.059	--	2.84	2.74
	Minor	-0.0077	0.049	--	0.055	2.77	0.055	--	2.77	2.65
	Not a Problem....................	0								
	Satisfied with teacher salary									
	Strongly Agree...................	1.24	0.058	--	0.061	2.75	0.061	--	2.75	2.71
	Somewhat Agree................	0.93	0.038	--	0.042	2.80	0.042	--	2.80	2.85
	Somewhat Disagree	0.46	0.040	--	0.043	2.78	0.043	--	2.78	2.88
	Strongly Disagree	0								

[1]With SRSSUBPOP option.

[2]SRSSUBPOP option not available for estimates pertaining to the entire sample.

[3]WesVarPC does not calculate design effects for totals or logistic regression coefficients.

Table 1b. Estimates, standard errors and design effects for WesVarPC, SUDAAN, and Stata: 1993-94 SASS teacher file (continued)

Statistic	Variable	Estimate	WesVarPC S.E.	WesVarPC DEFF	SUDAAN S.E.	SUDAAN DEFF	Stata S.E.	Stata DEFF1[1]	Stata DEFF2	Stata MEFF
	Teachers participate in making most important decisions at this school									
	Strongly agree....................	0.78	0.051	--	0.063	2.75	0.063	--	2.75	2.78
	Somewhat agree................	0.53	0.043	--	0.049	2.70	0.049	--	2.70	2.72
	Somewhat disagree	0.19	0.043	--	0.051	2.63	0.051	--	2.63	2.66
	Strongly disagree	0								
	Sector									
	Public...............................	-0.50	0.038	--	0.051	1.97	0.051	--	1.97	2.14
	Private..............................	0								
	Physical conflicts among students a problem									
	Serious	-0.044	0.072	--	0.087	2.92	0.087	--	2.92	2.91
	Moderate...........................	0.0065	0.061	--	0.060	2.69	0.060	--	2.69	2.81
	Minor	0.012	0.039	--	0.049	2.62	0.049	--	2.62	2.74
	Not a Problem...................	0								
	Sex									
	Male..................................	-0.31	0.029	--	0.035	2.55	0.035	-	2.55	2.92
	Female	0								

[1]With SRSSUBPOP option.

[2]SRSSUBPOP option not available for estimates based on the entire sample.

[3]WesVarPC does not calculate design effects for totals or logistic regression coefficients.

Table 1c. Estimated totals, means, and standard errors for WesVarPC, SUDAAN, and Stata: 1993-94 SASS school file

| Variable	Number of sample schools	Estimate	WesVarPC S.E.	SUDAAN S.E.	Stata S.E.
Public Schools					
Total # full-time teachers	n=8,767	2,338,092	20,532	12,032	12,032
Total # American Indian teachers...........	n=8,767	10,616	556	781	781
Total # Asian/PI teachers	n=8,767	27,293	1,047	1,141	1,141
Total # Hispanic teachers.......................	n=8,767	94,404	4,712	4,836	4,836
Total # black teachers	n=8,767	220,111	4,537	6,152	6,152
Total # white teachers	n=8,767	2,208,871	18,321	12,466	12,466
Total # students	n=8,767	41,621,660	393,746	217,541	217,541
Total # students in ESL	n=3,542	1,650,736	98,697	92,185	88,643
Total # students in Chap. 1.....................	n=4,787	5,847,768	176,629	171,326	140,117
Total # kindergartners	n=3,579	3,175,690	50,722	48,989	31,761
Private Schools					
Mean yearly salary for teachers with bachelor's degree	n=1,746	$16,329	137	181	181
Mean yearly salary for teachers with master's degree.......................................	n=1,746	$25,189	277	339	338

Table 1d. Estimated totals and standard errors for WesVarPC, SUDAAN, and Stata: 1993-94 school file

	# sample schools	Estimated total # full-time teachers	WesVarPC S.E.	SUDAAN S.E.	Stata S.E.
United States	n=8,767	2,338,092	20,532	12,032	12,032
Alabama	n=224	42,826	1,543	989	989
Alaska	n=169	7,286	359	238	238
Arkansas...	n=156	28,316	1,009	699	699
Arizona...	n=189	34,602	1,116	905	905
California	n=353	196,127	7,304	4,392	4,392
Colorado...	n=164	32,324	1,729	856	856
Connecticut	n=148	32,176	1,059	985	985
District of Columbia..............................	n=55	4,931	191	199	199
Delaware	n=63	6,397	208	296	296
Florida..	n=228	104,024	3,211	3,125	3,125
Georgia...	n=168	71,701	1,399	1,531	1,531
Hawaii	n=85	9,911	435	314	314
Iowa	n=158	29,586	1,339	1,138	1,138
Idaho	n=158	10,803	507	314	314
Illinois	n=238	101,152	2,755	3,204	3,204
Indiana..	n=166	52,000	1,627	1,302	1,302
Kansas	n=149	25,894	1,137	1,049	1,049
Kentucky	n=149	38,864	1,664	989	989
Louisiana..	n=207	46,197	1,116	1,081	1,081
Massachusetts	n=208	50,257	1,480	1,503	1,503
Maryland	n=135	40,170	702	1,000	1,000
Maine	n=145	13,507	638	414	414
Michigan	n=202	73,053	3,225	1,787	1,787
Minnesota...	n=160	37,819	1,848	1,289	1,289
Missouri	n=168	55,969	2,322	2,104	2,104
Mississippi	n=195	28,824	1,134	692	692
Montana	n=179	10,271	511	412	412
North Carolina.......................................	n=151	66,419	2,775	1,964	1,964
North Dakota...	n=165	6,707	362	291	291
Nebraska	n=139	17,276	1,034	838	838
New Hampshire.....................................	n=117	11,209	705	413	413
New Jersey	n=167	75,003	3,381	1,844	1,844
New Mexico..	n=160	18,181	747	529	529
Nevada	n=109	11,805	347	511	511
New York..	n=269	166,105	5,996	4,654	4,654
Ohio	n=176	99,291	3,416	2,764	2,764
Oklahoma..	n=306	36,926	1,366	800	800
Oregon..	n=159	22,297	930	684	684
Pennsylvania	n=169	103,152	6,003	2,880	2,880
Rhode Island...	n=88	7,943	519	328	328
South Carolina.......................................	n=141	36,909	1,832	1,223	1,223
South Dakota...	n=165	8,746	457	375	375
Tennessee	n=179	44,196	1,527	1,120	1,120
Texas..	n=380	214,725	8,659	4,375	4,375
Utah..	n=174	17,955	644	422	422
Virginia	n=158	59,795	3,138	1,910	1,910
Vermont	n=97	6,308	331	215	215
Washington	n=200	43,172	1,472	1,146	1,146
Wisconsin..	n=164	53,523	1,898	2,017	2,017
West Virginia..	n=154	19,134	1,070	745	745
Wyoming...	n=131	6,328	355	284	284

Figure 1. Percent Difference in Bootstrap and Linearization Standard Errors for 1993-94 SASS Estimated Public School Full-time Teacher Totals
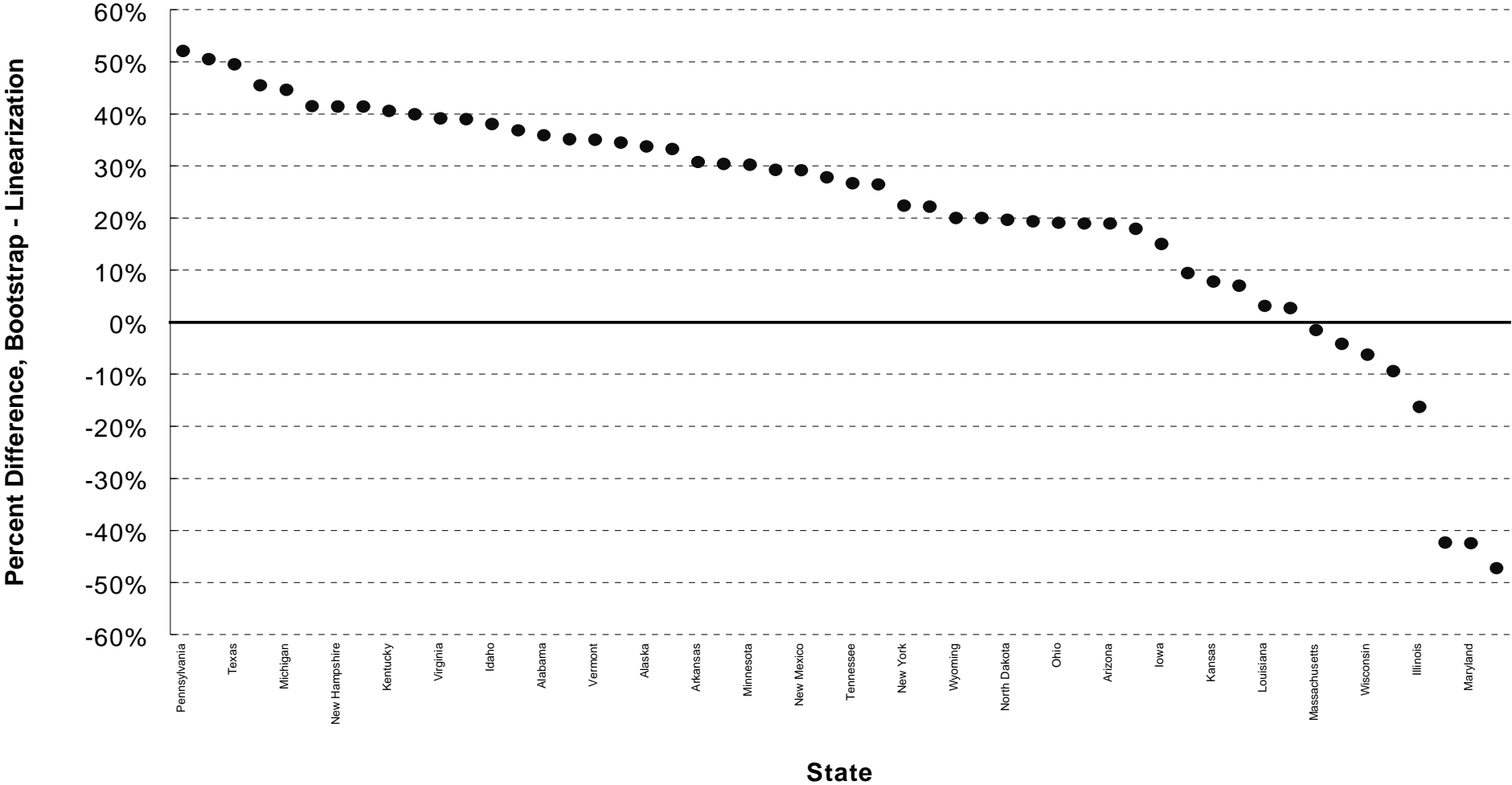
The linearization standard errors from SUDAAN and Stata were also compared with the replication standard errors from WesVarPC. These should not be expected to be identical, even if the variance strata used in constructing replicates were exactly the same as those used by SUDAAN and Stata, because replication and linearization are two different methods of variance estimation. There is also some randomness or "noise" in the standard error estimates which can account for some of the differences, especially if the number of replicates is small. However, both methods give consistent, asymptotically unbiased estimates of the standard error (Wolter, 1985, Theorems B.8, B.9). Therefore, if the sample design is appropriately reflected in the construction of the replicate weights and the SUDAAN and Stata input specifications, the linearization standard errors should be close to the replication estimates when the number of PSUs in each stratum is large. This was, in fact, the case for the NELS:88 student estimates, with the exception of the median reading and math scores.

The linearization and replication standard error estimates for 1993-94 SASS teacher estimates were similar for proportions and logistic regression coefficients (see table 1b), although the linearization estimates were usually somewhat larger (9 percent for proportions and 14 percent for logistic regression coefficients). This may be because the bootstrap replicate weights have accounted for the large sampling fractions in the school sampling, the impact on the standard error of the weighting adjustments, the order of selection of the schools, and the certainty schools, whereas the SUDAAN and Stata estimates have not. Thus the linearization estimates would be expected to be more conservative.

The school-based estimates of totals presented in tables 1c and 1d, on the other hand, tell a very different story. The linearization standard errors were usually significantly smaller, as can be seen in the estimates of total students in table 1c and especially in the public school state estimates in table 1d. In table 1d, the SUDAAN and Stata standard errors are on average 21 percent smaller than the WesVarPC bootstrap standard errors. Figure 1 graphically shows the difference between the bootstrap and linearization standard errors. Difference of this magnitude is contrary to what one would expect based on the literature (Kovar et al., 1988). This puzzling result cannot be attributed to the reasons cited above, because those factors would tend to make the SUDAAN and Stata standard error estimates larger than the bootstrap estimates, not smaller. Further discussion on the use of the bootstrap with the Schools and Staffing Survey is found in Kaufman (1993, 1994).

6.2   Design Effects

WesVarPC and SUDAAN calculate design effects differently, while Stata allows one the flexibility of calculating design effects using either the WesVarPC or the SUDAAN approach. The

difference in design effects can be substantial for subpopulations that have been oversampled. This can be seen in the design effects for "Other private" student estimates in table 1a. In WesVarPC, the design effect is defined as the ratio of the variance under the complex survey design to the variance under simple random sampling for a sample of the same size. WesVarPC conditions on the actual sample size achieved for the domain or subpopulation of interest in calculating the simple random sample variance. The formula WesVarPC uses to calculate the simple random sample variance of a proportion for the design effect is:

$$\frac{\hat{p}(1-\hat{p})}{n}$$

where $\hat{p}$ is estimated using the sample weights and n is the achieved sample size on which the estimate of p is based. The formula it uses for the variance of a mean under simple random sampling is:

$$\frac{\sum_{i=1}^{n} w_i (y_i - \bar{y})^2}{n \sum_{i=1}^{n} w_i}$$

where n is the number of observations on which the estimated mean is based, and w_i is the full-sample weight for the i-th observation. SUDAAN instead uses the expected sample size for the domain or subpopulation as if a simple random sample of the entire population had been drawn with no oversampling. For example, suppose the design effect is calculated on average income for black persons and that this domain was deliberately oversampled so that it has a sample size of 40 percent of the national sample, rather than the 13 percent that would be expected in a national simple random sample. The SRS variance is calculated in WesVarPC with $n=.4n_{total}$, where n_{total} is the national sample size. SUDAAN uses $n=.13\ n_{total}$ in its calculation.

To illustrate, in an r by c table if $\hat{p}_{rc}$ is the estimated proportion of the population in row r that is in cell (r,c), the simple random sample variance calculated by SUDAAN is given in Shah et al. (1996) as:

$$\frac{\hat{p}_{rc}(1-\hat{p}_{rc})}{(n-1)\dfrac{\hat{N}_{r.}}{\hat{N}}}$$

where $\hat{p}_{rc} = \hat{N}_{rc} / \hat{N}_{r.}$, n is the actual number of sampled observations with nonzero weights[7] on the file, $\hat{N}_{rc}$ is the sum of the weights of the sample observations in the (r,c) cell, $\hat{N}_{r.}$ is the sum of the weights of the sample observations in the r-th row, and $\hat{N}$ is the sum of the weights of all the sample observations

[7]Sampled observations with nonzero weights are presumably the eligible respondents.

(Shah et al., 1996, eqn. 6.10). The expression in the denominator is an estimate of the expected number of cases in the r-th row under SRS without any oversampling.

For an estimated total for the entire population, the simple random sample variance of the total is calculated by SUDAAN as:

$$\frac{\hat{N}^2 \sum\limits_{i=1}^{n} w_i (y_i - \bar{y})^2}{\hat{N}(n-1)}$$

and the simple random sample variance of a mean for the total population is calculated as:

$$\frac{\sum\limits_{i=1}^{n} w_i (y_i - \bar{y})^2}{\hat{N}(n-1)}$$

where n is the actual sample size in both numerator and denominator since the total or mean is estimated for the entire population, and $\hat{N} = \sum\limits_{i=1}^{n} w_i$ (Shah, 1996, eqn. 2.5). The design effect formula used by SUDAAN in calculating design effects for subpopulations is not given for totals or means in the manual. However, it would appear that if the estimate applies to a subpopulation, n in the denominator is the expected sample size for the subpopulation in a simple random sample from a population of size N. The summation in the numerator is still over the entire sample, since the simple random sample is assumed to be drawn from across the entire population. The n in the denominator is not necessarily the same n that WesVarPC uses. WesVarPC conditions on the actual sample size achieved for a subpopulation and does not use the expected sample size.

Stata can calculate the design effect using either method. The SRSSUBPOP DEFF option produces design effects which are conceptually the same as WesVarPC's. These are labeled DEFF1 in tables 1a and b. They are not exactly the same, because the standard errors produced from replication and linearization are not necessarily exactly the same for reasons explained earlier. Leaving out SRSSUBPOP will produce design effects identical to SUDAAN's. These are labeled DEFF2 in table 1b. If the user includes an "svyset fpc" command, an fpc will also be calculated as n/$\hat{N}$. In addition, Stata calculates a misspecification effect (MEFF). The misspecification effect is calculated as the ratio of the variance under the complex survey design to the variance under simple random sampling with replacement, ignoring all aspects of the sample design, including weighting. The formula Stata uses for the simple random sample variance in the MEFF is:

$$\frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n'(n'-1)}$$

where n' in the denominator is the expected sample size for a national simple random sample with no oversampling, unless the SRSSUBPOP option is specified, and n is the actual sample size. The summation in the numerator is still over the entire sample. If SRSSUBPOP is specified, n' in the denominator is the number of observations actually obtained in the sample for the subpopulation ($n = n'$). The misspecification effect is only useful for statistics that would not be too badly biased without using the weights, such as proportions, means, and regression coefficients. As can be seen in table 1b, the misspecification effects for estimated totals are quite large.

6.3   Hypothesis Testing

For categorical data analysis, WesVarPC and SUDAAN both produce adjusted chi-square statistics to test for independence in two-way tables. WesVarPC calculates a Rao-Scott first-order adjusted chi-square (RS2) and a Satterthwaite or Rao-Scott second-order adjusted chi-square statistic (RS3). In its PROC CROSSTAB procedure SUDAAN calculates two test statistics based on the Wald chi-square (CHISQ, LLCHISQ) and a survey data version of the Cochran-Mantel-Haenszel statistic (CMH). CHISQ and CMH are based on the observed minus the expected values; LLCHISQ is based on the log odd ratios. Stata does not calculate adjusted chi-square statistics for survey data.

The adjusted Wald F statistic (Skinner, Holt, and Smith, 1989) for the logistic regression model in table 1b was identical for Stata and SUDAAN since both are based on the same number of degrees of freedom for the variance estimator. WesVarPC's adjusted Wald F statistic is larger because the replication variance estimate is based on fewer degrees of freedom. Stata and SUDAAN can also calculate an unadjusted Wald F. SUDAAN will output a Satterthwaite Adjusted F, Satterthwaite adjusted chi- square, and Wald chi- square as well.

All three software packages have the capability to test hypotheses for linear combinations of means, ratios, totals, proportions, and regression coefficients. In WesVarPC this is done by creating functions of weighted totals; e.g., differences in means, proportions, ratios, or odds ratios in the COMPUTE and FUNCTION screens, or in the Create Hypothesis Testing screen when defining a regression model. In SUDAAN, hypotheses for linear contrasts in estimated means, total, proportions and ratios can be tested in PROC DESCRIPT and PROC RATIO through the DIFFVAR, CONTRAST, PAIRWISE, and POLYNOMIAL statements. Hypotheses for linear contrasts of regression coefficients

can be tested in the regression procedures (PROC REGRESS, PROC LOGISTIC) with CONTRAST statements. Both single degree and multiple degree of freedom hypothesis tests may be performed in PROC HYPTEST as well by inputting estimates obtained from other SUDAAN PROCs. Stata does hypothesis testing of linear combinations of survey estimates, including regression model coefficients, with its svylc and svytest commands.

7.   Missing Data

In WesVarPC, if an observation has a missing full-sample or replicate weight, it is excluded from the calculations. In modeling, any observation having a missing value for either the dependent variable or any of the independent variables is excluded from all estimates for that model. This exclusion is at the model level; i.e., the inclusion of variables in one model has no effect on the other models in the same program request (.wvr file). In tables, the treatment of missing values for analysis variables and variables involved in compute statistics is determined through the use of the MISSING and COMPLETE options. If the COMPLETE option is on, only observations with nonmissing data for all analysis variables and compute statistics in the request program will be included. The implication thus is that all statistics in the program will be based on the same number of observations. If the COMPLETE option is off, all available cases will be used for each statistic estimated. For example, in computing the pupil-teacher ratio, only schools with both the total number of pupils and teachers reported are used by WesVarPC. But in the same request program (.wvq file), the estimate of the total number of teachers in the population would be based on all schools reporting the number of teachers, and the estimate of the total number of pupils would be based on all schools reporting the number of pupils. Thus, the estimated pupil-teacher ratio could be based on a different number of observations than either of the estimated total pupils or teachers.

SUDAAN excludes observations with missing values for sample design variables on the NEST and WEIGHT statements. The default treatment of observations with missing data for analysis variables is the same as WesVarPC's for modeling procedures. In PROC DESCRIPT, CROSSTAB, and RATIO, the treatment of missing data is the same as when WesVarPC's COMPLETE option is set to off. It is possible to treat missing values as a separate level of a categorical variable by coding the INCLUDE option on the PROC statement. Observations with missing values will then be included in the marginal counts.

Stata has two ways of treating missing data in tables through the use of its COMPLETE and AVAILABLE options. These correspond to WesVarPC's COMPLETE=ON and COMPLETE=OFF options, respectively. Stata treats missing data the same way WesVarPC and SUDAAN do in estimating models and calculating ratios. However, it apparently is not possible to treat missing values as a separate level for categorical variables in tables. An idiosyncrasy that may be important in recoding or creating new variables with the generate, egen, or replace commands is that relationally Stata considers numeric missing values to be larger than any numeric value. This is the opposite of SAS, WesVarPC, and SUDAAN.

The treatment of missing data is summarized for the three packages in table 2.

Table 2. Available missing data options for tables in WesVarPC, SUDAAN, and Stata

Option	WesVarPC	SUDAAN	Stata
Treat missing values as a separate level	Yes: set MISSING option to ON in Tables screen	Yes: specify INCLUDE option in PROC statement; specify DISPLAY to see counts	Not available
Delete observations with missing data for any analysis variables in the table request (use only observations with complete data for all analysis variables used in the table)	Yes: set COMPLETE option to ON in Table request screen	Not available	Yes: specify COMPLETE in svy statement
Include all available observations with data for each analysis variable separately	Yes: set COMPLETE option to OFF in Table request screen (this is the default)	Yes: this is the default	Yes: this is the default (or specify AVAILABLE in svy statement)

8.   Execution Times

Execution times are compared in table 3 below for five jobs. The time needed for the PREP step in WesVarPC and file preparation outside of SUDAAN and Stata are not included. All four jobs were run under Windows 95 on a Compaq DeskPro 133MHz Pentium with 32 MB of ram and a 1 GB hard disk. The 32-bit processor version of WesVarPC was run. The input files for each software package are not the same size because the WesVarPC files contain replicate weights, and because Stata and WesVarPC store the input files in their own internal format. However, these are the only differences. The size of the NELS:88 student input file was 23 MB for WesVarPC, including the

replicate weights, 3 MB for SUDAAN, and 2.2 MB for Stata. The 1993-94 SASS teacher input file was 43 MB for WesVarPC (including replicate weights), 23 MB for SUDAAN, and 12 MB for Stata.

Due to the size of the input files, the default memory allocation in Stata had to be increased to 9,000K for the NELS:88 student file and to 19,000K file for the 1993-94 SASS teacher file. The setting of this parameter can greatly affect the execution time for Stata. If the memory setting exceeds the physical memory capacity of the machine on which it is executing, Stata will begin using hard disk space as virtual memory, which can slow the job down considerably. For example, at a memory allocation setting of 20,000K, Stata began "paging" through the hard disk, and the logistic regression program in table 3 took 2 hours and 35 minutes!

Table 3.   Execution times for WesVarPC, SUDAAN, and Stata

File	Statistics	WesVarPC	SUDAAN	Stata
NELS:88 Students n=23,416	4 Means	30 seconds	30 seconds	20 seconds
	4 Medians, ungrouped method	40 seconds	1 minute, 20 seconds	not available in Stata
n=24,599	12 Proportions	30 seconds	1 minute, 30 seconds	17 minute, 50 seconds (svyprop)
1993-94 SASS Teachers n=51,604	logistic regression model	37 minutes, 30 seconds	7 minutes, 30 seconds	3 minutes
	15 proportions	2 minutes	2 minutes, 12 seconds	4 minutes, 25 seconds (svymean)

For many of the runs in this evaluation, the execution times were not greatly different. The exceptions were in the case of logistic regression and proportions. SUDAAN and Stata were noticeably faster than WesVarPC in the logistic regression run, but Stata was slower in computing proportions, especially when the svyprop command was used. The inclusion of replace and generate statements to perform recodes in a Stata program will also slow its execution time. Note that recoding is done in WesVarPC after the PREP step, but before submitting the program request, and is not included in the execution times. The PREP step for the NELS:88 student file took 18 minutes; the RECODE step for all the variables involved in the evaluation took 20 minutes.

9.   Summary and Recommendations

For many people, ease of use, cost, and whether the software package does what they need are more important considerations than computer execution time. While SUDAAN is probably the most

powerful package of the three, it is also the most difficult to learn. WesVarPC is both easy to learn and powerful, but it requires replicate weights, and it does not have some of the model fitting capabilities that SUDAAN does. Stata is more limited in its survey data analysis capabilities and can be slower to run, but it does enable one to easily plot and examine predicted values and residuals when model-fitting. This may make it a good choice for regression or probit analysis, since neither WesVarPC nor SUDAAN have any graphical capabilities. For categorical data analysis and large-scale table production, we recommend SUDAAN or WesVarPC. The survey data analysis capabilities are summarized in table 4 for each of the three packages.

The choice of software may also be dictated by the availability of sample design variables and replicate weights on the public or restricted use analysis files. To allow for more choices in software, we recommend that the variables needed for variance estimation with linearization packages and a set of replicate weights be created internally by sampling statisticians knowledgeable with the sample design and data collection procedures, and that these be placed on the NCES analysis files. This will also decrease the likelihood of errors in variance calculations, since all analysts would be using the same set of variance strata, PSUs, and replicate weights.

All three software packages will give good approximations to the standard error for the base year NELS:88 files (but note the caveat below about estimating standard errors for totals in SUDAAN); therefore, it is largely a matter of convenience which one is chosen. With the next release of WesVarPC it will be easier to create replicate weights within the software package, which should remove some of the difficulty in using WesVarPC with these files.

Table 4. **Analysis Capabilities for WesVarPC, SUDAAN, and Stata**

	WesVarPC	SUDAAN	Stata
Standard errors and design effects for means, totals, proportions, ratios..	X[8]	X	X
Quantiles..	X	X	
Finite Population Correction Factor:			
1st stage only, equal probabilities of selection	X	X	X
1st stage with unequal probabilities of selection		X	
Linear Regression	X	X	X
Logistic Regression			
Dichotomous..	X	X	X
Polychotomous		X	X[9]
Probit Models			X
Loglinear Models..		X	
Tests of Independence in Tables......................................	X	X	
Linear Contrasts, Differences..	X	X	X
Survival Analysis..		X	
Graphics..			X
Batch Processing Available	X	X	X
Output Useful for Importing into Spreadsheets	X	X	

For the Schools and Staffing Survey LEA files, the use of linearization packages is not advised. Conservative variance estimates will generally be obtained with SUDAAN and Stata for the teacher files because of the high school sampling fractions. However, we caution that for estimates of totals this conclusion does not always hold; in fact, the bootstrap replication standard errors were often significantly larger than the linearization estimates , especially for estimates of totals based on the school file. In addition, we have noticed inconsistencies between Stata and SUDAAN standard errors for totals in the presence of missing data. This is because Stata treats missing data basically the same way WesVarPC does, but SUDAAN sets missing values to zero when estimating the variance of a total. For all of the above reasons we recommend using WesVarPC with the bootstrap replicate weights created by NCES for Schools and Staffing Survey estimates.

---

[8] No design effects available for totals.

[9] Available to registered users as of November, 1997 by downloading the appropriate ado-files from the Stata web site.

The system requirements to run each of the three packages under Windows 95 are given below:

- WesVarPC: a 486 or higher PC with at least 8 MB of RAM and a 5.7 MB hard disk
- SUDAAN: a 386 or higher PC with at least 4 MB of RAM and a 5 MB hard disk
- Stata: a 386 or higher PC with at least 8 MB of RAM and a math co-processor.

Because the software world is fast-changing, we recommend contacting each software provider directly to obtain current information on cost and new releases. The Web sites for each provider are given below.

SUDAAN:	http://www.rti.org
Stata:	http://www.stata.com
WesVarPC:	http://www.westat.com

REFERENCES

Abramson, R., Cole, C., Fondelier, S., Jackson, R., and Kaufman, S., (1996), *1993-94 Schools and Staffing Survey: Sample Design and Estimation,* U.S. Department of Education. Washington, DC.

Binder, D. (1991), *Use of Estimating Functions for Interval Estimation from Complex Surveys,* 1991 ASA Survey Research 34-42 (Disc. 43-45).

Brick, M., et al. (1997). *A User's Guide to WesVarPC.* Westat, Inc. Rockville, MD.

Cohen, S. (1996), *An Evaluation of Alternative PC Based Software Packages Developed for the Analysis of Complex Survey Data.* Agency for Health Care Policy and Research. Rockville, MD.

Fuller, W., and Francisco, C. (1991), *Quantile Estimation with a Complex Survey Design,* 1991 Annals of Statistics 19, 454-469.

Kaufman, S. (1993). A Bootstrap Variance Estimator for the Schools and Staffing Survey. *Proceedings of the Section on Survey Research Methods, American Statistical Association,* pp. 675-680. Alexandria, VA: American Statistical Association.

Kaufman, S. (1994). Properties of the Schools and Staffing Survey Bootstrap Variance Estimator. *Proceedings of the Section on Survey Research Methods, American Statistical Association,* pp. 1116-1121. Alexandria, VA: American Statistical Association.

Kovar, J.G., Rao, J.N.K., and Wu, C.F.J. (1988). Bootstrap and Other Methods to Measure Errors in Survey Estimates, *The Canadian Journal of Statistics,* Vol. 16, Supplement, 25-45.

Lepkowski, J. and Bowles, J. (1996). *Sampling Error Software for Personal Computers.* Survey Statistician, No. 35, Vol. 10-17.

Newton, H. J. (1996). *Stata Technical Bulletin,* College Station, TX.

Nigam, A.K. and Rao, J.N.K. (1996). On Balanced Bootstrap for Stratified Multistage Samples*, Statistical Sinica*, 6, 199-214.

Shah, B.V., Barnwell, B., and Bieler, G. (1995). *SUDAAN User's Manual: Software for Analysis of Correlated Data,* Release 6.40. Research Triangle Institute, Research Triangle Park, NC.

Shah, B.V., et al. (1996). *Technical Manual: Statistical Methods and Algorithms Used in SUDAAN Release 7.0.* Research Triangle Institute, Research Triangle Park, NC.

Sitter, R. R. (1990). *Comparing Three Bootstrap Methods for Survey Data.* Technical Report Series of the Laboratory for Research in Statistics and Probability, No. 152.

Skinner, C., Holt, D., and Smith, T. (1989). *Analysis of Complex Surveys.* New York: Wiley & Sons.

Spencer, B., Frankel, M., Ingels, S., Rasinski, K., and Tourangeau, R., *National Education Longitudinal Study of 1988 Base Year Sample Design Report,* U.S. Department of Education. Washington, DC.

Stata Statistical Software User's Guide, Release 5 (1997), Stata Corporation, College Station, TX.

Weng, S.S., Zhang, F., and Cohen, M.P. (1995). *Variance Estimates Comparison by Statistical Software,* 1995 Proceedings of the Section on Survey Research Methods, American Statistical Association, 333-338.

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

Woodruff, R.S. (1952). *Confidence Intervals for Medians and Other Position Measures.* Journal of the American Statistical Association 47, 635-646.

This page intentionally left blank.

Listing of NCES Working Papers to Date

Working papers can be downloaded as pdf files from the NCES Electronic Catalog (http://nces.ed.gov/pubsearch/). You can also contact Sheilah Jupiter at (202) 219-1761 (sheilah_jupiter@ed.gov) if you are interested in any of the following papers

Listing of NCES Working Papers by Program Area

No.	Title	NCES contact
Baccalaureate and Beyond (B&B)		
98-15	Development of a Prototype System for Accessing Linked NCES Data	Steven Kaufman
Beginning Postsecondary Students (BPS) Longitudinal Study		
98-11	Beginning Postsecondary Students Longitudinal Study First Follow-up (BPS:96-98) Field Test Report	Aurora D'Amico
98-15	Development of a Prototype System for Accessing Linked NCES Data	Steven Kaufman
1999-15	Projected Postsecondary Outcomes of 1992 High School Graduates	Aurora D'Amico
Common Core of Data (CCD)		
95-12	Rural Education Data User's Guide	Samuel Peng
96-19	Assessment and Analysis of School-Level Expenditures	William J. Fowler, Jr.
97-15	Customer Service Survey: Common Core of Data Coordinators	Lee Hoffman
97-43	Measuring Inflation in Public School Costs	William J. Fowler, Jr.
98-15	Development of a Prototype System for Accessing Linked NCES Data	Steven Kaufman
1999-03	Evaluation of the 1996-97 Nonfiscal Common Core of Data Surveys Data Collection, Processing, and Editing Cycle	Beth Young
Decennial Census School District Project		
95-12	Rural Education Data User's Guide	Samuel Peng
96-04	Census Mapping Project/School District Data Book	Tai Phan
98-07	Decennial Census School District Project Planning Report	Tai Phan
Early Childhood Longitudinal Study (ECLS)		
96-08	How Accurate are Teacher Judgments of Students' Academic Performance?	Jerry West
96-18	Assessment of Social Competence, Adaptive Behaviors, and Approaches to Learning with Young Children	Jerry West
97-24	Formulating a Design for the ECLS: A Review of Longitudinal Studies	Jerry West
97-36	Measuring the Quality of Program Environments in Head Start and Other Early Childhood Programs: A Review and Recommendations for Future Research	Jerry West
1999-01	A Birth Cohort Study: Conceptual and Design Considerations and Rationale	Jerry West
Education Finance Statistics Center (EDFIN)		
94-05	Cost-of-Education Differentials Across the States	William J. Fowler, Jr.
96-19	Assessment and Analysis of School-Level Expenditures	William J. Fowler, Jr.
97-43	Measuring Inflation in Public School Costs	William J. Fowler, Jr.
98-04	Geographic Variations in Public Schools' Costs	William J. Fowler, Jr.
1999-16	Measuring Resources in Education: From Accounting to the Resource Cost Model Approach	William J. Fowler, Jr.
High School and Beyond (HS&B)		
95-12	Rural Education Data User's Guide	Samuel Peng
1999-05	Procedures Guide for Transcript Studies	Dawn Nelson
1999-06	1998 Revision of the Secondary School Taxonomy	Dawn Nelson
HS Transcript Studies		
1999-05	Procedures Guide for Transcript Studies	Dawn Nelson
1999-06	1998 Revision of the Secondary School Taxonomy	Dawn Nelson

No.	Title	NCES contact

International Adult Literacy Survey (IALS)
| 97-33 | Adult Literacy: An International Perspective | Marilyn Binkley |

Integrated Postsecondary Education Data System (IPEDS)
| 97-27 | Pilot Test of IPEDS Finance Survey | Peter Stowe |
| 98-15 | Development of a Prototype System for Accessing Linked NCES Data | Steven Kaufman |

National Assessment of Educational Progress (NAEP)
95-12	Rural Education Data User's Guide	Samuel Peng
97-29	Can State Assessment Data be Used to Reduce State NAEP Sample Sizes?	Steven Gorman
97-30	ACT's NAEP Redesign Project: Assessment Design is the Key to Useful and Stable Assessment Results	Steven Gorman
97-31	NAEP Reconfigured: An Integrated Redesign of the National Assessment of Educational Progress	Steven Gorman
97-32	Innovative Solutions to Intractable Large Scale Assessment (Problem 2: Background Questionnaires)	Steven Gorman
97-37	Optimal Rating Procedures and Methodology for NAEP Open-ended Items	Steven Gorman
97-44	Development of a SASS 1993-94 School-Level Student Achievement Subfile: Using State Assessments and State NAEP, Feasibility Study	Michael Ross
98-15	Development of a Prototype System for Accessing Linked NCES Data	Steven Kaufman
1999-05	Procedures Guide for Transcript Studies	Dawn Nelson
1999-06	1998 Revision of the Secondary School Taxonomy	Dawn Nelson

National Assessment of Adult Literacy Survey (NAALS)
98-17	Developing the National Assessment of Adult Literacy: Recommendations from Stakeholders	Sheida White
1999-09a	1992 National Adult Literacy Survey: An Overview	Alex Sedlacek
1999-09b	1992 National Adult Literacy Survey: Sample Design	Alex Sedlacek
1999-09c	1992 National Adult Literacy Survey: Weighting and Population Estimates	Alex Sedlacek
1999-09d	1992 National Adult Literacy Survey: Development of the Survey Instruments	Alex Sedlacek
1999-09e	1992 National Adult Literacy Survey: Scaling and Proficiency Estimates	Alex Sedlacek
1999-09f	1992 National Adult Literacy Survey: Interpreting the Adult Literacy Scales and Literacy Levels	Alex Sedlacek
1999-09g	1992 National Adult Literacy Survey: Literacy Levels and the Response Probability Convention	Alex Sedlacek

National Education Longitudinal Study of 1988 (NELS:88)
95-04	National Education Longitudinal Study of 1988: Second Follow-up Questionnaire Content Areas and Research Issues	Jeffrey Owings
95-05	National Education Longitudinal Study of 1988: Conducting Trend Analyses of NLS-72, HS&B, and NELS:88 Seniors	Jeffrey Owings
95-06	National Education Longitudinal Study of 1988: Conducting Cross-Cohort Comparisons Using HS&B, NAEP, and NELS:88 Academic Transcript Data	Jeffrey Owings
95-07	National Education Longitudinal Study of 1988: Conducting Trend Analyses HS&B and NELS:88 Sophomore Cohort Dropouts	Jeffrey Owings
95-12	Rural Education Data User's Guide	Samuel Peng
95-14	Empirical Evaluation of Social, Psychological, & Educational Construct Variables Used in NCES Surveys	Samuel Peng
96-03	National Education Longitudinal Study of 1988 (NELS:88) Research Framework and Issues	Jeffrey Owings
98-06	National Education Longitudinal Study of 1988 (NELS:88) Base Year through Second Follow-Up: Final Methodology Report	Ralph Lee
98-09	High School Curriculum Structure: Effects on Coursetaking and Achievement in Mathematics for High School Graduates—An Examination of Data from the National Education Longitudinal Study of 1988	Jeffrey Owings
98-15	Development of a Prototype System for Accessing Linked NCES Data	Steven Kaufman
1999-05	Procedures Guide for Transcript Studies	Dawn Nelson
1999-06	1998 Revision of the Secondary School Taxonomy	Dawn Nelson
1999-15	Projected Postsecondary Outcomes of 1992 High School Graduates	Aurora D'Amico

No.	Title	NCES contact

National Household Education Survey (NHES)

No.	Title	NCES contact
95-12	Rural Education Data User's Guide	Samuel Peng
96-13	Estimation of Response Bias in the NHES:95 Adult Education Survey	Steven Kaufman
96-14	The 1995 National Household Education Survey: Reinterview Results for the Adult Education Component	Steven Kaufman
96-20	1991 National Household Education Survey (NHES:91) Questionnaires: Screener, Early Childhood Education, and Adult Education	Kathryn Chandler
96-21	1993 National Household Education Survey (NHES:93) Questionnaires: Screener, School Readiness, and School Safety and Discipline	Kathryn Chandler
96-22	1995 National Household Education Survey (NHES:95) Questionnaires: Screener, Early Childhood Program Participation, and Adult Education	Kathryn Chandler
96-29	Undercoverage Bias in Estimates of Characteristics of Adults and 0- to 2-Year-Olds in the 1995 National Household Education Survey (NHES:95)	Kathryn Chandler
96-30	Comparison of Estimates from the 1995 National Household Education Survey (NHES:95)	Kathryn Chandler
97-02	Telephone Coverage Bias and Recorded Interviews in the 1993 National Household Education Survey (NHES:93)	Kathryn Chandler
97-03	1991 and 1995 National Household Education Survey Questionnaires: NHES:91 Screener, NHES:91 Adult Education, NHES:95 Basic Screener, and NHES:95 Adult Education	Kathryn Chandler
97-04	Design, Data Collection, Monitoring, Interview Administration Time, and Data Editing in the 1993 National Household Education Survey (NHES:93)	Kathryn Chandler
97-05	Unit and Item Response, Weighting, and Imputation Procedures in the 1993 National Household Education Survey (NHES:93)	Kathryn Chandler
97-06	Unit and Item Response, Weighting, and Imputation Procedures in the 1995 National Household Education Survey (NHES:95)	Kathryn Chandler
97-08	Design, Data Collection, Interview Timing, and Data Editing in the 1995 National Household Education Survey	Kathryn Chandler
97-19	National Household Education Survey of 1995: Adult Education Course Coding Manual	Peter Stowe
97-20	National Household Education Survey of 1995: Adult Education Course Code Merge Files User's Guide	Peter Stowe
97-25	1996 National Household Education Survey (NHES:96) Questionnaires: Screener/Household and Library, Parent and Family Involvement in Education and Civic Involvement, Youth Civic Involvement, and Adult Civic Involvement	Kathryn Chandler
97-28	Comparison of Estimates in the 1996 National Household Education Survey	Kathryn Chandler
97-34	Comparison of Estimates from the 1993 National Household Education Survey	Kathryn Chandler
97-35	Design, Data Collection, Interview Administration Time, and Data Editing in the 1996 National Household Education Survey	Kathryn Chandler
97-38	Reinterview Results for the Parent and Youth Components of the 1996 National Household Education Survey	Kathryn Chandler
97-39	Undercoverage Bias in Estimates of Characteristics of Households and Adults in the 1996 National Household Education Survey	Kathryn Chandler
97-40	Unit and Item Response Rates, Weighting, and Imputation Procedures in the 1996 National Household Education Survey	Kathryn Chandler
98-03	Adult Education in the 1990s: A Report on the 1991 National Household Education Survey	Peter Stowe
98-10	Adult Education Participation Decisions and Barriers: Review of Conceptual Frameworks and Empirical Studies	Peter Stowe

National Longitudinal Study of the High School Class of 1972 (NLS-72)

No.	Title	NCES contact
95-12	Rural Education Data User's Guide	Samuel Peng

National Postsecondary Student Aid Study (NPSAS)

No.	Title	NCES contact
96-17	National Postsecondary Student Aid Study: 1996 Field Test Methodology Report	Andrew G. Malizio

National Study of Postsecondary Faculty (NSOPF)

No.	Title	NCES contact
97-26	Strategies for Improving Accuracy of Postsecondary Faculty Lists	Linda Zimbler
98-15	Development of a Prototype System for Accessing Linked NCES Data	Steven Kaufman
2000-01	1999 National Study of Postsecondary Faculty (NSOPF:99) Field Test Report	Linda Zimbler

Private School Universe Survey (PSS)

95-16	Intersurvey Consistency in NCES Private School Surveys	Steven Kaufman
95-17	Estimates of Expenditures for Private K-12 Schools	Stephen Broughman
96-16	Strategies for Collecting Finance Data from Private Schools	Stephen Broughman
96-26	Improving the Coverage of Private Elementary-Secondary Schools	Steven Kaufman
96-27	Intersurvey Consistency in NCES Private School Surveys for 1993-94	Steven Kaufman
97-07	The Determinants of Per-Pupil Expenditures in Private Elementary and Secondary Schools: An Exploratory Analysis	Stephen Broughman
97-22	Collection of Private School Finance Data: Development of a Questionnaire	Stephen Broughman
98-15	Development of a Prototype System for Accessing Linked NCES Data	Steven Kaufman

Recent College Graduates (RCG)

98-15	Development of a Prototype System for Accessing Linked NCES Data	Steven Kaufman

Schools and Staffing Survey (SASS)

94-01	Schools and Staffing Survey (SASS) Papers Presented at Meetings of the American Statistical Association	Dan Kasprzyk
94-02	Generalized Variance Estimate for Schools and Staffing Survey (SASS)	Dan Kasprzyk
94-03	1991 Schools and Staffing Survey (SASS) Reinterview Response Variance Report	Dan Kasprzyk
94-04	The Accuracy of Teachers' Self-reports on their Postsecondary Education: Teacher Transcript Study, Schools and Staffing Survey	Dan Kasprzyk
94-06	Six Papers on Teachers from the 1990-91 Schools and Staffing Survey and Other Related Surveys	Dan Kasprzyk
95-01	Schools and Staffing Survey: 1994 Papers Presented at the 1994 Meeting of the American Statistical Association	Dan Kasprzyk
95-02	QED Estimates of the 1990-91 Schools and Staffing Survey: Deriving and Comparing QED School Estimates with CCD Estimates	Dan Kasprzyk
95-03	Schools and Staffing Survey: 1990-91 SASS Cross-Questionnaire Analysis	Dan Kasprzyk
95-08	CCD Adjustment to the 1990-91 SASS: A Comparison of Estimates	Dan Kasprzyk
95-09	The Results of the 1993 Teacher List Validation Study (TLVS)	Dan Kasprzyk
95-10	The Results of the 1991-92 Teacher Follow-up Survey (TFS) Reinterview and Extensive Reconciliation	Dan Kasprzyk
95-11	Measuring Instruction, Curriculum Content, and Instructional Resources: The Status of Recent Work	Sharon Bobbitt & John Ralph
95-12	Rural Education Data User's Guide	Samuel Peng
95-14	Empirical Evaluation of Social, Psychological, & Educational Construct Variables Used in NCES Surveys	Samuel Peng
95-15	Classroom Instructional Processes: A Review of Existing Measurement Approaches and Their Applicability for the Teacher Follow-up Survey	Sharon Bobbitt
95-16	Intersurvey Consistency in NCES Private School Surveys	Steven Kaufman
95-18	An Agenda for Research on Teachers and Schools: Revisiting NCES' Schools and Staffing Survey	Dan Kasprzyk
96-01	Methodological Issues in the Study of Teachers' Careers: Critical Features of a Truly Longitudinal Study	Dan Kasprzyk
96-02	Schools and Staffing Survey (SASS): 1995 Selected papers presented at the 1995 Meeting of the American Statistical Association	Dan Kasprzyk
96-05	Cognitive Research on the Teacher Listing Form for the Schools and Staffing Survey	Dan Kasprzyk
96-06	The Schools and Staffing Survey (SASS) for 1998-99: Design Recommendations to Inform Broad Education Policy	Dan Kasprzyk
96-07	Should SASS Measure Instructional Processes and Teacher Effectiveness?	Dan Kasprzyk
96-09	Making Data Relevant for Policy Discussions: Redesigning the School Administrator Questionnaire for the 1998-99 SASS	Dan Kasprzyk
96-10	1998-99 Schools and Staffing Survey: Issues Related to Survey Depth	Dan Kasprzyk
96-11	Towards an Organizational Database on America's Schools: A Proposal for the Future of SASS, with comments on School Reform, Governance, and Finance	Dan Kasprzyk
96-12	Predictors of Retention, Transfer, and Attrition of Special and General Education Teachers: Data from the 1989 Teacher Followup Survey	Dan Kasprzyk
96-15	Nested Structures: District-Level Data in the Schools and Staffing Survey	Dan Kasprzyk
96-23	Linking Student Data to SASS: Why, When, How	Dan Kasprzyk

No.	Title	NCES contact
96-24	National Assessments of Teacher Quality	Dan Kasprzyk
96-25	Measures of Inservice Professional Development: Suggested Items for the 1998-1999 Schools and Staffing Survey	Dan Kasprzyk
96-28	Student Learning, Teaching Quality, and Professional Development: Theoretical Linkages, Current Measurement, and Recommendations for Future Data Collection	Mary Rollefson
97-01	Selected Papers on Education Surveys: Papers Presented at the 1996 Meeting of the American Statistical Association	Dan Kasprzyk
97-07	The Determinants of Per-Pupil Expenditures in Private Elementary and Secondary Schools: An Exploratory Analysis	Stephen Broughman
97-09	Status of Data on Crime and Violence in Schools: Final Report	Lee Hoffman
97-10	Report of Cognitive Research on the Public and Private School Teacher Questionnaires for the Schools and Staffing Survey 1993-94 School Year	Dan Kasprzyk
97-11	International Comparisons of Inservice Professional Development	Dan Kasprzyk
97-12	Measuring School Reform: Recommendations for Future SASS Data Collection	Mary Rollefson
97-14	Optimal Choice of Periodicities for the Schools and Staffing Survey: Modeling and Analysis	Steven Kaufman
97-18	Improving the Mail Return Rates of SASS Surveys: A Review of the Literature	Steven Kaufman
97-22	Collection of Private School Finance Data: Development of a Questionnaire	Stephen Broughman
97-23	Further Cognitive Research on the Schools and Staffing Survey (SASS) Teacher Listing Form	Dan Kasprzyk
97-41	Selected Papers on the Schools and Staffing Survey: Papers Presented at the 1997 Meeting of the American Statistical Association	Steve Kaufman
97-42	Improving the Measurement of Staffing Resources at the School Level: The Development of Recommendations for NCES for the Schools and Staffing Survey (SASS)	Mary Rollefson
97-44	Development of a SASS 1993-94 School-Level Student Achievement Subfile: Using State Assessments and State NAEP, Feasibility Study	Michael Ross
98-01	Collection of Public School Expenditure Data: Development of a Questionnaire	Stephen Broughman
98-02	Response Variance in the 1993-94 Schools and Staffing Survey: A Reinterview Report	Steven Kaufman
98-04	Geographic Variations in Public Schools' Costs	William J. Fowler, Jr.
98-05	SASS Documentation: 1993-94 SASS Student Sampling Problems; Solutions for Determining the Numerators for the SASS Private School (3B) Second-Stage Factors	Steven Kaufman
98-08	The Redesign of the Schools and Staffing Survey for 1999-2000: A Position Paper	Dan Kasprzyk
98-12	A Bootstrap Variance Estimator for Systematic PPS Sampling	Steven Kaufman
98-13	Response Variance in the 1994-95 Teacher Follow-up Survey	Steven Kaufman
98-14	Variance Estimation of Imputed Survey Data	Steven Kaufman
98-15	Development of a Prototype System for Accessing Linked NCES Data	Steven Kaufman
98-16	A Feasibility Study of Longitudinal Design for Schools and Staffing Survey	Stephen Broughman
1999-02	Tracking Secondary Use of the Schools and Staffing Survey Data: Preliminary Results	Dan Kasprzyk
1999-04	Measuring Teacher Qualifications	Dan Kasprzyk
1999-07	Collection of Resource and Expenditure Data on the Schools and Staffing Survey	Stephen Broughman
1999-08	Measuring Classroom Instructional Processes: Using Survey and Case Study Fieldtest Results to Improve Item Construction	Dan Kasprzyk
1999-10	What Users Say About Schools and Staffing Survey Publications	Dan Kasprzyk
1999-12	1993-94 Schools and Staffing Survey: Data File User's Manual, Volume III: Public-Use Codebook	Kerry Gruber
1999-13	1993-94 Schools and Staffing Survey: Data File User's Manual, Volume IV: Bureau of Indian Affairs (BIA) Restricted-Use Codebook	Kerry Gruber
1999-14	1994-95 Teacher Followup Survey: Data File User's Manual, Restricted-Use Codebook	Kerry Gruber
1999-17	Secondary Use of the Schools and Staffing Survey Data	Susan Wiley

Listing of NCES Working Papers by Subject

No.	Title	NCES contact

Adult education

96-14	The 1995 National Household Education Survey: Reinterview Results for the Adult Education Component	Steven Kaufman
96-20	1991 National Household Education Survey (NHES:91) Questionnaires: Screener, Early Childhood Education, and Adult Education	Kathryn Chandler
96-22	1995 National Household Education Survey (NHES:95) Questionnaires: Screener, Early Childhood Program Participation, and Adult Education	Kathryn Chandler
98-03	Adult Education in the 1990s: A Report on the 1991 National Household Education Survey	Peter Stowe
98-10	Adult Education Participation Decisions and Barriers: Review of Conceptual Frameworks and Empirical Studies	Peter Stowe
1999-11	Data Sources on Lifelong Learning Available from the National Center for Education Statistics	Lisa Hudson

American Indian – education

1999-13	1993-94 Schools and Staffing Survey: Data File User's Manual, Volume IV: Bureau of Indian Affairs (BIA) Restricted-Use Codebook	Kerry Gruber

Assessment/achievement

95-12	Rural Education Data User's Guide	Samuel Peng
95-13	Assessing Students with Disabilities and Limited English Proficiency	James Houser
97-29	Can State Assessment Data be Used to Reduce State NAEP Sample Sizes?	Larry Ogle
97-30	ACT's NAEP Redesign Project: Assessment Design is the Key to Useful and Stable Assessment Results	Larry Ogle
97-31	NAEP Reconfigured: An Integrated Redesign of the National Assessment of Educational Progress	Larry Ogle
97-32	Innovative Solutions to Intractable Large Scale Assessment (Problem 2: Background Questions)	Larry Ogle
97-37	Optimal Rating Procedures and Methodology for NAEP Open-ended Items	Larry Ogle
97-44	Development of a SASS 1993-94 School-Level Student Achievement Subfile: Using State Assessments and State NAEP, Feasibility Study	Michael Ross
98-09	High School Curriculum Structure: Effects on Coursetaking and Achievement in Mathematics for High School Graduates—An Examination of Data from the National Education Longitudinal Study of 1988	Jeffrey Owings

Beginning students in postsecondary education

98-11	Beginning Postsecondary Students Longitudinal Study First Follow-up (BPS:96-98) Field Test Report	Aurora D'Amico

Civic participation

97-25	1996 National Household Education Survey (NHES:96) Questionnaires: Screener/Household and Library, Parent and Family Involvement in Education and Civic Involvement, Youth Civic Involvement, and Adult Civic Involvement	Kathryn Chandler

Climate of schools

95-14	Empirical Evaluation of Social, Psychological, & Educational Construct Variables Used in NCES Surveys	Samuel Peng

Cost of education indices

94-05	Cost-of-Education Differentials Across the States	William J. Fowler, Jr.

Course-taking

95-12	Rural Education Data User's Guide	Samuel Peng
98-09	High School Curriculum Structure: Effects on Coursetaking and Achievement in Mathematics for High School Graduates—An Examination of Data from the National Education Longitudinal Study of 1988	Jeffrey Owings

No.	Title	NCES contact
1999-05	Procedures Guide for Transcript Studies	Dawn Nelson
1999-06	1998 Revision of the Secondary School Taxonomy	Dawn Nelson

Crime

97-09	Status of Data on Crime and Violence in Schools: Final Report	Lee Hoffman

Curriculum

95-11	Measuring Instruction, Curriculum Content, and Instructional Resources: The Status of Recent Work	Sharon Bobbitt & John Ralph
98-09	High School Curriculum Structure: Effects on Coursetaking and Achievement in Mathematics for High School Graduates—An Examination of Data from the National Education Longitudinal Study of 1988	Jeffrey Owings

Customer service

1999-10	What Users Say About Schools and Staffing Survey Publications	Dan Kasprzyk
2000-02	Coordinating NCES Surveys: Options, Issues, Challenges, and Next Steps	Valena Plisko

Data quality

97-13	Improving Data Quality in NCES: Database-to-Report Process	Susan Ahmed

Design effects

2000-03	Strengths and Limitations of Using SUDAAN, Stata, and WesVarPC for Computing Variances from NCES Data Sets	Ralph Lee

Dropout rates, high school

95-07	National Education Longitudinal Study of 1988: Conducting Trend Analyses HS&B and NELS:88 Sophomore Cohort Dropouts	Jeffrey Owings

Early childhood education

96-20	1991 National Household Education Survey (NHES:91) Questionnaires: Screener, Early Childhood Education, and Adult Education	Kathryn Chandler
96-22	1995 National Household Education Survey (NHES:95) Questionnaires: Screener, Early Childhood Program Participation, and Adult Education	Kathryn Chandler
97-24	Formulating a Design for the ECLS: A Review of Longitudinal Studies	Jerry West
97-36	Measuring the Quality of Program Environments in Head Start and Other Early Childhood Programs: A Review and Recommendations for Future Research	Jerry West
1999-01	A Birth Cohort Study: Conceptual and Design Considerations and Rationale	Jerry West

Educational attainment

98-11	Beginning Postsecondary Students Longitudinal Study First Follow-up (BPS:96-98) Field Test Report	Aurora D'Amico

Educational research

2000-02	Coordinating NCES Surveys: Options, Issues, Challenges, and Next Steps	Valena Plisko

Employment

96-03	National Education Longitudinal Study of 1988 (NELS:88) Research Framework and Issues	Jeffrey Owings
98-11	Beginning Postsecondary Students Longitudinal Study First Follow-up (BPS:96-98) Field Test Report	Aurora D'Amico

Faculty – higher education

97-26	Strategies for Improving Accuracy of Postsecondary Faculty Lists	Linda Zimbler
2000-01	1999 National Study of Postsecondary Faculty (NSOPF:99) Field Test Report	Linda Zimbler

Finance – elementary and secondary schools

94-05	Cost-of-Education Differentials Across the States	William J. Fowler, Jr.
96-19	Assessment and Analysis of School-Level Expenditures	William J. Fowler, Jr.
98-01	Collection of Public School Expenditure Data: Development of a Questionnaire	Stephen Broughman

No.	Title	NCES contact
1999-07	Collection of Resource and Expenditure Data on the Schools and Staffing Survey	Stephen Broughman
1999-16	Measuring Resources in Education: From Accounting to the Resource Cost Model Approach	William J. Fowler, Jr.

Finance – postsecondary

97-27	Pilot Test of IPEDS Finance Survey	Peter Stowe

Finance – private schools

95-17	Estimates of Expenditures for Private K-12 Schools	Stephen Broughman
96-16	Strategies for Collecting Finance Data from Private Schools	Stephen Broughman
97-07	The Determinants of Per-Pupil Expenditures in Private Elementary and Secondary Schools: An Exploratory Analysis	Stephen Broughman
97-22	Collection of Private School Finance Data: Development of a Questionnaire	Stephen Broughman
1999-07	Collection of Resource and Expenditure Data on the Schools and Staffing Survey	Stephen Broughman

Geography

98-04	Geographic Variations in Public Schools' Costs	William J. Fowler, Jr.

Inflation

97-43	Measuring Inflation in Public School Costs	William J. Fowler, Jr.

Institution Data

2000-01	1999 National Study of Postsecondary Faculty (NSOPF:99) Field Test Report	Linda Zimbler

Instructional resources and practices

95-11	Measuring Instruction, Curriculum Content, and Instructional Resources: The Status of Recent Work	Sharon Bobbitt & John Ralph
1999-08	Measuring Classroom Instructional Processes: Using Survey and Case Study Field Test Results to Improve Item Construction	Dan Kasprzyk

International comparisons

97-11	International Comparisons of Inservice Professional Development	Dan Kasprzyk
97-16	International Education Expenditure Comparability Study: Final Report, Volume I	Shelley Burns
97-17	International Education Expenditure Comparability Study: Final Report, Volume II, Quantitative Analysis of Expenditure Comparability	Shelley Burns

Libraries

94-07	Data Comparability and Public Policy: New Interest in Public Library Data Papers Presented at Meetings of the American Statistical Association	Carrol Kindel
97-25	1996 National Household Education Survey (NHES:96) Questionnaires: Screener/Household and Library, Parent and Family Involvement in Education and Civic Involvement, Youth Civic Involvement, and Adult Civic Involvement	Kathryn Chandler

Limited English Proficiency

95-13	Assessing Students with Disabilities and Limited English Proficiency	James Houser

Literacy of adults

98-17	Developing the National Assessment of Adult Literacy: Recommendations from Stakeholders	Sheida White
1999-09a	1992 National Adult Literacy Survey: An Overview	Alex Sedlacek
1999-09b	1992 National Adult Literacy Survey: Sample Design	Alex Sedlacek
1999-09c	1992 National Adult Literacy Survey: Weighting and Population Estimates	Alex Sedlacek
1999-09d	1992 National Adult Literacy Survey: Development of the Survey Instruments	Alex Sedlacek
1999-09e	1992 National Adult Literacy Survey: Scaling and Proficiency Estimates	Alex Sedlacek
1999-09f	1992 National Adult Literacy Survey: Interpreting the Adult Literacy Scales and Literacy Levels	Alex Sedlacek
1999-09g	1992 National Adult Literacy Survey: Literacy Levels and the Response Probability Convention	Alex Sedlacek

No.	Title	NCES contact
1999-11	Data Sources on Lifelong Learning Available from the National Center for Education Statistics	Lisa Hudson

Literacy of adults – international
97-33	Adult Literacy: An International Perspective	Marilyn Binkley

Mathematics
98-09	High School Curriculum Structure: Effects on Coursetaking and Achievement in Mathematics for High School Graduates—An Examination of Data from the National Education Longitudinal Study of 1988	Jeffrey Owings
1999-08	Measuring Classroom Instructional Processes: Using Survey and Case Study Field Test Results to Improve Item Construction	Dan Kasprzyk

Parental involvement in education
96-03	National Education Longitudinal Study of 1988 (NELS:88) Research Framework and Issues	Jeffrey Owings
97-25	1996 National Household Education Survey (NHES:96) Questionnaires: Screener/Household and Library, Parent and Family Involvement in Education and Civic Involvement, Youth Civic Involvement, and Adult Civic Involvement	Kathryn Chandler
1999-01	A Birth Cohort Study: Conceptual and Design Considerations and Rationale	Jerry West

Participation rates
98-10	Adult Education Participation Decisions and Barriers: Review of Conceptual Frameworks and Empirical Studies	Peter Stowe

Postsecondary education
1999-11	Data Sources on Lifelong Learning Available from the National Center for Education Statistics	Lisa Hudson

Postsecondary education – persistence and attainment
98-11	Beginning Postsecondary Students Longitudinal Study First Follow-up (BPS:96-98) Field Test Report	Aurora D'Amico
1999-15	Projected Postsecondary Outcomes of 1992 High School Graduates	Aurora D'Amico

Postsecondary education – staff
97-26	Strategies for Improving Accuracy of Postsecondary Faculty Lists	Linda Zimbler
2000-01	1999 National Study of Postsecondary Faculty (NSOPF:99) Field Test Report	Linda Zimbler

Private schools
96-16	Strategies for Collecting Finance Data from Private Schools	Stephen Broughman
97-07	The Determinants of Per-Pupil Expenditures in Private Elementary and Secondary Schools: An Exploratory Analysis	Stephen Broughman
97-22	Collection of Private School Finance Data: Development of a Questionnaire	Stephen Broughman

Projections of education statistics
1999-15	Projected Postsecondary Outcomes of 1992 High School Graduates	Aurora D'Amico

Public school finance
1999-16	Measuring Resources in Education: From Accounting to the Resource Cost Model Approach	William J. Fowler, Jr.

Public schools
97-43	Measuring Inflation in Public School Costs	William J. Fowler, Jr.
98-01	Collection of Public School Expenditure Data: Development of a Questionnaire	Stephen Broughman
98-04	Geographic Variations in Public Schools' Costs	William J. Fowler, Jr.
1999-02	Tracking Secondary Use of the Schools and Staffing Survey Data: Preliminary Results	Dan Kasprzyk

No.	Title	NCES contact

Public schools – secondary
98-09 High School Curriculum Structure: Effects on Coursetaking and Achievement in Mathematics for High School Graduates—An Examination of Data from the National Education Longitudinal Study of 1988 — Jeffrey Owings

Reform, educational
96-03 National Education Longitudinal Study of 1988 (NELS:88) Research Framework and Issues — Jeffrey Owings

Response rates
98-02 Response Variance in the 1993-94 Schools and Staffing Survey: A Reinterview Report — Steven Kaufman

School districts, public
98-07 Decennial Census School District Project Planning Report — Tai Phan
1999-03 Evaluation of the 1996-97 Nonfiscal Common Core of Data Surveys Data Collection, Processing, and Editing Cycle — Beth Young

School districts, public – demographics of
96-04 Census Mapping Project/School District Data Book — Tai Phan

Schools
97-42 Improving the Measurement of Staffing Resources at the School Level: The Development of Recommendations for NCES for the Schools and Staffing Survey (SASS) — Mary Rollefson
98-08 The Redesign of the Schools and Staffing Survey for 1999-2000: A Position Paper — Dan Kasprzyk
1999-03 Evaluation of the 1996-97 Nonfiscal Common Core of Data Surveys Data Collection, Processing, and Editing Cycle — Beth Young

Schools – safety and discipline
97-09 Status of Data on Crime and Violence in Schools: Final Report — Lee Hoffman

Software evaluation
2000-03 Strengths and Limitations of Using SUDAAN, Stata, and WesVarPC for Computing Variances from NCES Data Sets — Ralph Lee

Staff
97-42 Improving the Measurement of Staffing Resources at the School Level: The Development of Recommendations for NCES for the Schools and Staffing Survey (SASS) — Mary Rollefson
98-08 The Redesign of the Schools and Staffing Survey for 1999-2000: A Position Paper — Dan Kasprzyk

Staff – higher education institutions
97-26 Strategies for Improving Accuracy of Postsecondary Faculty Lists — Linda Zimbler

State
1999-03 Evaluation of the 1996-97 Nonfiscal Common Core of Data Surveys Data Collection, Processing, and Editing Cycle — Beth Young

Statistical methodology
97-21 Statistics for Policymakers or Everything You Wanted to Know About Statistics But Thought You Could Never Understand — Susan Ahmed

Students with disabilities
95-13 Assessing Students with Disabilities and Limited English Proficiency — James Houser

Survey methodology
96-17 National Postsecondary Student Aid Study: 1996 Field Test Methodology Report — Andrew G. Malizio
97-15 Customer Service Survey: Common Core of Data Coordinators — Lee Hoffman
97-35 Design, Data Collection, Interview Administration Time, and Data Editing in the 1996 National Household Education Survey — Kathryn Chandler

No.	Title	NCES contact
98-06	National Education Longitudinal Study of 1988 (NELS:88) Base Year through Second Follow-Up: Final Methodology Report	Ralph Lee
98-11	Beginning Postsecondary Students Longitudinal Study First Follow-up (BPS:96-98) Field Test Report	Aurora D'Amico
98-16	A Feasibility Study of Longitudinal Design for Schools and Staffing Survey	Stephen Broughman
1999-07	Collection of Resource and Expenditure Data on the Schools and Staffing Survey	Stephen Broughman
1999-17	Secondary Use of the Schools and Staffing Survey Data	Susan Wiley
2000-01	1999 National Study of Postsecondary Faculty (NSOPF:99) Field Test Report	Linda Zimbler
2000-02	Coordinating NCES Surveys: Options, Issues, Challenges, and Next Steps	Valena Plisko

Teachers

98-13	Response Variance in the 1994-95 Teacher Follow-up Survey	Steven Kaufman
1999-14	1994-95 Teacher Followup Survey: Data File User's Manual, Restricted-Use Codebook	Kerry Gruber

Teachers – instructional practices of

98-08	The Redesign of the Schools and Staffing Survey for 1999-2000: A Position Paper	Dan Kasprzyk

Teachers – opinions regarding safety

98-08	The Redesign of the Schools and Staffing Survey for 1999-2000: A Position Paper	Dan Kasprzyk

Teachers – performance evaluations

1999-04	Measuring Teacher Qualifications	Dan Kasprzyk

Teachers – qualifications of

1999-04	Measuring Teacher Qualifications	Dan Kasprzyk

Teachers – salaries of

94-05	Cost-of-Education Differentials Across the States	William J. Fowler, Jr.

Variance estimation

2000-03	Strengths and Limitations of Using SUDAAN, Stata, and WesVarPC for Computing Variances from NCES Data Sets	Ralph Lee

Violence

97-09	Status of Data on Crime and Violence in Schools: Final Report	Lee Hoffman

Vocational education

95-12	Rural Education Data User's Guide	Samuel Peng
1999-05	Procedures Guide for Transcript Studies	Dawn Nelson
1999-06	1998 Revision of the Secondary School Taxonomy	Dawn Nelson