Community Tracking Study

Comparison of Selected Statistical Software Packages for Variance Estimation in the CTS Surveys

Elizabeth Schaefer
Frank Potter
Stephen Williams
Nuria Diaz-Tena
James D. Reschovsky
Gary Moore

600 Maryland Avenue, SW
Suite 550
Washington, DC 20024
www.hschange.org

Technical Publication No.

40

May 2003

AUTHORS

Elizabeth Schaefer, Center for Studying Health System Change

Frank Potter, Mathematica Policy Research

Stephen Williams, Mathematica Policy Research

Nuria Diaz-Tena, Mathematica Policy Research

James D. Reschovsky, Center for Studying Health System Change

Gary Moore, Social & Scientific Systems

Center for Studying Health System Change
600 Maryland Ave., SW, Suite 550, Washington, DC 20024
www.hschange.org

Mathematica Policy Research, Inc.
P.O. Box 2393, Princeton, NJ 08543-2393
www.mathematica-mpr.com

Social & Scientific Systems, Inc.
8757 Georgia Ave., Silver Spring, MD 20910
www.s-3.com

ACKNOWLEDGEMENTS

We would like to express our great appreciation to Lucy Hsu and Bryan Sayer of Social & Scientific Systems for their programming support and expertise, as well as helpful comments on an earlier draft. We also thank Joy Grossman of HSC for her review of an earlier draft.

This is one in a series of technical documents that have been done as part of the Community Tracking Study being conducted by the Center for Studying Health System Change (HSC), which is funded exclusively by The Robert Wood Johnson Foundation and is affiliated with Mathematica Policy Research, Inc.

CONTENTS

Appendices

CHAPTER 1

INTRODUCTION

This report addresses the suitability of some commonly used statistical software packages for the analysis of the Community Tracking Study (CTS) surveys of households and physicians. The CTS is a national study of changes in the health care system and the effects of those changes on people. Funded by the Robert Wood Johnson Foundation, the study is being conducted by the Center for Studying Health System Change (HSC). Two of the main ongoing data collection efforts in the CTS are the Household Survey and the Physician Survey, each of which has been conducted three times (1996-97, 1998-99, and 2000-01); data collection is currently underway for the 2003 Household Survey. Both surveys have samples distributed throughout the U.S. and can generate estimates for the nation as well as for selected individual communities. The Household Survey includes about 60,000 people, and the Physician Survey includes about 12,000 physicians.

The purpose of this report is to allow CTS data users to make an informed choice about the software that they use to analyze the CTS survey data. In the past, user's guides for the CTS data files have recommended SUDAAN because of its ability to more fully capture the complex sample design of the CTS. However, some CTS data users have expressed an interest in using other statistical software packages because those packages are more familiar to them, are able to do more types of statistical analyses, or are available to them at a lower cost. In order to be responsive to that, HSC worked with Mathematica Policy Research, Inc. (MPR) to produce this report. MPR is the organization that has been responsible for survey sample design and related estimation issues since the inception of the CTS.

The choice of software matters because the CTS surveys, like most large national surveys, use complex sampling techniques rather than simple random sampling. Although the calculation of sampling variance is easiest with a simple random sample, a complex sample design is necessary to achieve specific analytic goals at minimal cost. In contrast to simple random sampling, complex sample design can include stratification, multistage cluster sampling, and unequal sampling rates. All of these features, along with nonresponse adjustments and poststratification of the weights, affect the sampling variance and influence the way variances should be calculated. If complex survey data are analyzed as if they were from a simple random sample, the sampling errors will typically be understated, affecting the significance tests and precision statements. An additional concern is that not all software with complex survey analysis capabilities accommodates complex sample designs to the same degree. Some of the commonly used statistical software packages (e.g., SAS and Stata) have the ability to correctly handle many simple and complex sampling designs but not all complex designs. Furthermore, the types of sample designs that can be accommodated often depend on the statistical routine being used.

In the sections that follow, we review the CTS design, some of the commonly used statistical software packages available, and methods for computing variances for the estimates from complex surveys. Following that, we summarize results from the CTS data, comparing variance estimates using the different algorithms available in alternative software packages. That comparison allows us to identify situations where some other software packages besides SUDAAN can be used to obtain reasonable estimates of the sampling variances (or at least “conservative” estimates, by which we mean estimates that decrease the likelihood of finding a statistically significant result).

Chapter 6 summarizes our conclusions and provides recommendations for researchers. For those who choose to analyze the CTS data with statistical software other than SUDAAN, Appendices B and C indicate which sampling variables to use and how to obtain them.

CHAPTER 2

CTS DESIGN AND FEATURES AFFECTING VARIANCE ESTIMATION

The CTS surveys rely on a complex sample design rather than simple random sampling. To help with understanding subsequent chapters of this report, this chapter provides a summary of the design of the CTS household and physician surveys and the types of estimates that can be calculated from them. Following that is a discussion identifying the specific design features in the CTS that affect variance estimation.

2.1. SUMMARY OF THE CTS DESIGN

The first three rounds of the CTS surveys (1996-97, 1998-99, and 2000-01) have had a sample design that consists of two independent samples: the *site sample* and the *supplemental sample*. The site sample comes from 60 randomly selected communities (*sites*) in the 48 contiguous states in the U.S. It reflects the Community Tracking Study's focus on local health care markets, since health care delivery is primarily local. The supplemental sample is a much smaller sample, drawn from throughout the contiguous U.S. Its purpose is to increase the precision of national estimates with a relatively small increase in sample size. The supplemental sample is only about 10 percent as large as the site sample.

The following sections describe the selection of the site and supplemental samples and the types of estimates for which they can be used. More detailed information is available in the user's guides for the public use data files and in the survey methodology reports, which are currently available for the 1996-97 and 1998-99 surveys and are listed in the references section of this report. The documentation for the 2000-01 surveys is forthcoming and will be available as technical publications on the HSC Web site (www.hschange.org).

2.1.1. Site Sample

Table 2.1 describes how the site sample was selected for the 1996-97 surveys. Only minor changes were made for later years, and those changes are described briefly in the text below. The first stage of sample selection was to select the 60 sites, and the second stage was to select individual households and physicians from those sites. Understanding the sampling discussion requires knowing that *stratification* is the partitioning of the sampling units (i.e., sites, households, and physicians) into groups (*strata*) prior to sample allocation and selection. Stratification is discussed in more detail in Section 2.2.

As shown in Table 2.1, before selection, sites first needed to be defined and placed into three strata based on site type (metropolitan or not) and size: large metropolitan sites, small metropolitan sites, and nonmetropolitan sites. Metropolitan sites generally conform to the metropolitan statistical areas (MSAs) defined by the Office of Management and Budget. The nonmetropolitan sites conform to the economic areas defined by the Bureau of Economic Analysis. A metropolitan site was considered "large" if the 1992 population was greater than 200,000. Within these three strata, 48 large metropolitan sites, three small metropolitan sites, and nine nonmetropolitan sites were selected with probability proportional to size (site population).[1] Within the stratum of large metropolitan sites, there are nine *certainty sites*, which were selected with certainty because of their size and/or significance.[2]

Within the group of 48 large metropolitan sites, 12 were randomly selected to be the *high-intensity* sites. Because greater precision was desired for estimates in the high-intensity sites, they have larger samples of households and physicians than the other sites (called the *low-intensity* sites). The high-intensity sites are also central to the qualitative component of the CTS, which consists of intensive biennial case studies in those 12 markets. The larger sample sizes in the high-intensity sites allow more precise estimates, which can then be used to inform the case studies.

The second stage of sampling involved selection of individual households[3] and physicians within each of the 60 sites. Stratification was used in a few cases. In the Physician Survey, physicians were stratified by patient care classification (primary care or specialist), and primary care physicians (PCPs) were oversampled. In the Household Survey, stratification was imposed only in the high-intensity sites, as described in Table 2.1. For both surveys, simple random sampling was used within each stratum (or site, if there was no stratification within the site).

In the Household Survey, the members of each household were divided into *family insurance units (FIUs)*, which contain the household members typically covered under a family insurance policy. Detailed information was obtained on the adults in each FIU and, if there were any children, one child in the FIU. In FIUs with more than one child, a child was randomly selected to be included in the sample.

For the 1998-99 and 2000-01 surveys, the 60 sites remained the same, as did their designations of high-intensity and low-intensity. However, for the second stage of sampling, substrata were defined within the sites in order to obtain a specific allocation between those in the population for the prior survey and those new to the survey population (i.e., newly-formed households or physicians new to the profession). The substrata were defined also to allow oversampling of respondents from the prior survey. This sample allocation was designed to improve statistical precision for cross-sectional and change estimates, to ensure complete coverage of the survey populations, and to minimize survey costs. It also allows analysis of physician-level change for a panel of the Physician Survey.

---

[1] *Implicit stratification* by region was used for sites in the large and small metropolitan strata and for states in the nonmetropolitan stratum. Stratification was *implicit* in that the sites and states were selected using a sequential selection procedure in which geography was a factor in the ordering of the sites and states. This approach contrasts to selection within explicitly constructed geographic strata.
[2] For more information on site selection, see Metcalf et al. (1996).
[3] The selection of households occurred through selection of phone numbers for the *telephone sample* (which is a Random Digit Dial, or RDD, sample) and selection of housing units for the *nontelephone sample* (which is also known as the *field* sample). The purpose of the nontelephone sample was to be able to calculate estimates that are representative of everyone in the U.S., not just people with telephones.

2.1.2. National Supplemental Sample

Table 2.2 describes how the national supplemental sample was selected. Because the sample was essentially a stratified simple random sample drawn from throughout the contiguous U.S., the selection of individual households[4] and physicians occurred at the first stage of sample selection. Stratification was used for both surveys, although the strata differed. Simple random sampling was used within each stratum. As in the site sample described above, interviewed households were divided into FIUs, and only one randomly selected child from each FIU was in the sample. Only minor changes were made to sample selection for 1998-99 and 2000-01. Specifically, as in the site sample, further substrata were defined in order to obtain a specific allocation between those in the population for the prior survey and those new to the survey population. The substrata were defined also to allow oversampling of respondents from the prior survey.

2.1.3. Types of Estimates

Table 2.3 indicates how the site sample and supplemental sample can be used separately and together to make national and site-specific estimates. For national estimates, the combined sample will provide the most precise estimates. However, for certain types of analyses, it may be preferable to use the other samples to make national estimates; the user's guides listed in the references of this report provide a discussion of those types of analyses. For site-specific estimates, the site sample observations in each site are augmented with the observations from the supplemental sample that are also in that site. Note that the site-specific estimates for the low-intensity sites will be less precise because of the smaller sample sizes for those sites.

2.2. CTS DESIGN FEATURES AFFECTING VARIANCE ESTIMATION

Many features of the CTS design have implications for variance estimation. This section identifies those features and discusses how their disadvantages with respect to variance estimation are counter-balanced by advantages in achieving other survey goals.

2.2.1. Clustering

The CTS, by definition, is community-based, and the sites were therefore defined to correspond to local health care market areas. As a consequence, the CTS site sample has observations clustered geographically in the 60 sites.[5] The advantage of clustering the sample in sites is that it produces sufficient observations in each site to analyze individual local markets and control for market characteristics in multivariate analyses. The disadvantage is the effect of clustering on variances for national estimates. Because observations within a cluster are typically more similar than observations from different clusters, they tend to exhibit a positive intra-cluster correlation, which reduces survey precision and needs to be accounted for in variance estimation. Note that the degree to which precision is affected by clustering is not the same for each estimate; instead, it depends on the intra-cluster correlation for the measure and subsample used for the estimate.

---

[4] The national supplement did not include a nontelephone sample.
[5] In the Household Survey, observations are also clustered within households and family insurance units.

2.2.2. Stratification

Stratification is the partitioning of sampling units into groups (*strata*) prior to sample allocation and selection. It is a feature of most large-scale surveys and performs several important functions. Those functions include ensuring adequate sample size for important study populations and optimally allocating sample for surveys in which some groups exhibit more variability in responses or are more costly to survey. Stratification is also a useful tool for ensuring adequate dispersion of the sample. Since stratification is a departure from simple random sampling, to receive the benefits from it, the variance estimation algorithm needs to account for it. Stratification is used in many parts of the CTS sample selection, as indicated in Tables 2.1 and 2.2. For example, it is used in selecting the national supplemental sample to ensure that all regions of the U.S. are proportionally represented.

2.2.3. Unequal Selection Probabilities and Without-Replacement Sampling

For the site sample, the first sampling stage (i.e., selection of the 60 sites) uses *unequal selection probabilities*. Specifically, instead of assigning the same selection probability to all sites (which in this case are the *primary sampling units*, or PSUs), the sites are selected with probabilities proportional to size (PPS). PPS sampling at the first stage is used to obtain sampling weights for the final-stage units that are nearly equal and is therefore important for increasing the efficiency (i.e., greater precision) for survey estimates for the final-stage units. Essentially all major surveys use PPS sampling of the first-stage units.

Since the CTS sites were selected from a relatively small frame (176 large MSAs and 118 small MSAs), variance estimation can take advantage of the improved precision from *without-replacement* (WOR) sampling. For the metropolitan noncertainty sites, the sampling rate was large enough that variance estimation should include a finite population correction factor (a function of inclusion and joint inclusion probabilities). In these cases, the use of the *with-replacement* (WR) variance estimation assumption would tend to overstate the true variance.

2.2.4. Unequal Weights

Unequal weighting indirectly affects variance estimation, and there are multiple sources of unequal weights in the CTS: oversampling, nonresponse adjustment, the selection of only one child per family in the Household Survey, and errors on the sampling frame.

Oversampling. In most multi-stage surveys, the primary sampling units (PSUs) are selected with probability proportional to size, and then an equal number of elementary units (e.g., physicians in a physician survey) are selected within each PSU. This procedure results in basically equal selection probabilities for all elementary units and in equal sampling weights. In the CTS, however, there are multiple examples of oversampling within the PSUs (i.e., the sites). In the Physician Survey, PCPs are oversampled, and both surveys have oversampling in the high-intensity sites relative to the low-intensity sites. In addition, there is oversampling of the respondents from the prior survey year, which reduces the cost per interview and improves precision for change or trend estimates. Oversampling tends to increase precision for some estimates (e.g., site-specific estimates in the high-intensity sites) but decrease precision for other estimates (e.g., national estimates from the combined sample).[6]

Nonresponse adjustment. Unequal weights also result from the adjustment to the sampling weights that is made to account for variation in survey response rates across different groups of people. Even for initial samples that have equal sampling weights, nonresponse adjustment introduces unequal weighting. The objective of the nonresponse adjustment is to reduce the potential for bias from survey nonresponse, but it is achieved at the expense of a modest increase in variance. For the CTS weights, nonresponse adjustments utilize weighting class adjustments for the Household Survey and propensity modeling for the Physician Survey. Weight trimming is used to reduce the effect of a few extreme weights, and weights are post-stratified so that the estimated national and major domain totals agree with known counts.

One child per family. In the Household Survey, the members of the selected households are divided into families, and only one child per family is included in the survey. When two or more children are in a family, the child weight is inflated to account for this sampling.

Sampling frame errors. Misclassification and other errors in the sampling frames can result in unequal weighting for specific estimates. For example, the misclassification of the location of physician practice in the Physician Survey results in increased variation in weights for site-specific estimates.

2.2.5. Combining Samples

As discussed above, the CTS surveys consist of two independent samples: the clustered site sample and the unclustered supplemental sample. Table 2.3 shows that each sample produces an independent estimate of national statistics, and the two samples can also be combined to obtain a single national estimate. Combining the samples produces the most precise estimates, but the calculation of variances needs to account for the two different designs (the multi-stage design of the site sample and the single-stage stratified random sampling of the supplemental sample). In addition, combining multiple years of data also requires variance estimation procedures to accommodate the overlap between the samples from each year.

[6] The types of estimates that can be made from the various samples in the CTS are discussed below and shown in Table 2.3.

Table 2.1
Selection of Site Sample in CTS Household Survey and Physician Survey

First stage: Selecting 60 sites Sites selected for the 1996-97 surveys are used for all rounds.			**Second stage: Selecting households or physicians within each site** Describes 1996-97 surveys. See text for changes in later years.	
Strata for Site Size and Type	**Select Sites**	**High-Intensity or Low-Intensity**	**Household Survey: Select Households**	**Physician Survey: Select Physicians**
Large metropolitan sites (48 sites)	Sites implicitly stratified by region and selected without replacement using probability proportional to site population. Eight sites were large enough, and one site considered important enough, to be selected with certainty, and those nine sites are called the *certainty sites*. The other sites are the *noncertainty* sites.	Twelve sites randomly selected to be the *high-intensity* sites, which means that they have larger second-stage samples (i.e., larger samples of households and physicians) than the *low-intensity* sites. The purpose of having larger samples in high-intensity sites is greater precision for those sites.	<u>High-intensity sites</u>: Two samples *Telephone sample*: Within each site, two county strata formed: the county containing the central city of the MSA vs. the other counties. Within the central city county, strata formed using local (sub-county level) income distribution or race/ethnicity distribution, depending on the composition of the county. Simple random sample of telephone numbers within each stratum. *Nontelephone sample*: [About two percent of sample.] Within each site, sub-county level geographic clusters sampled with probability proportional to size. Housing units selected within each cluster and screened for eligibility. <u>Low-intensity sites</u>: Simple random sample of telephone numbers within each site. No stratification and no nontelephone sample.	Physicians stratified by specialty (PCP or specialist). PCPs oversampled. Simple random sample within each stratum.
Small metropolitan sites (3 sites)	Sites implicitly stratified by region and selected without replacement using probability proportional to site population.	All small metropolitan sites are low-intensity.	Simple random sample of telephone numbers within each site. No stratification and no nontelephone sample.	Physicians stratified by specialty (PCP or specialist). PCPs oversampled. Simple random sample within each stratum.
Nonmetropolitan sites (9 sites)	States implicitly stratified by region and nine states selected without replacement using probability proportional to state population in nonmetropolitan areas. Within each state, one nonmetropolitan area selected using probability proportional to population in the area.	All nonmetropolitan sites are low-intensity.	Simple random sample of telephone numbers within each site. No stratification and no nontelephone sample.	Physicians stratified by specialty (PCP or specialist). PCPs oversampled. Simple random sample within each stratum.

Table 2.2
Selection of National Supplement Sample
in CTS Household Survey and Physician Survey

Selecting Households or Physicians from Anywhere in the Contiguous U.S. Describes 1996-97 surveys. See text for changes in later years.	
Household Survey: Selecting Households	**Physician Survey: Selecting Physicians**
Five strata: one for nonmetropolitan areas and four for metropolitan areas (defined by the four Census regions). Simple random sample of telephone numbers within each stratum. No nontelephone sample.	Twenty strata, defined by 10 geographic categories and two physician specialty classifications (PCP or specialist). Simple random sample within each stratum.

Table 2.3
Samples and Estimates
from the CTS Household Survey and Physician Survey

Sample	**Definition of Sample**	**Estimates for Nation**[a]	**Estimates for Individual Sites**
Site sample	Sample chosen within the 60 CTS sites only (see Table 2.1).	✔	
National supplement	Sample chosen from throughout the contiguous U.S. (see Table 2.2).	✔	
Augmented site sample	The entire site sample combined with the subsample of the national supplement that falls within the boundaries of the 60 CTS sites.	✔	✔
Combined sample	The entire site sample combined with the entire national supplement.	✔	

[a] For each year of each survey, national estimates can be made from either the site sample or the augmented site sample and in some cases from both.

CHAPTER 3

TECHNIQUES FOR ESTIMATING SAMPLING VARIANCE FROM COMPLEX SAMPLE DESIGNS

The sampling variance of an estimate derived from survey data for a statistic (such as a total, a mean or proportion, or a regression coefficient) is a measure of the random variation among estimates of the same statistic computed over repeated implementation of the same sample design with the same sample size on the same population. The sampling variance is a function of the population characteristics, the form of the statistic, and the nature of the sampling design. The two general forms of statistics are linear combinations of the survey data (e.g., a total) and nonlinear combinations of the survey data. Nonlinear combinations include the ratio of two estimates (e.g., a mean or a proportion in which both the numerator and the denominator are estimated) and more complex combinations such as regression coefficients. For linear estimates with simple sample designs (such as a stratified or unstratified simple random sample) or complex designs (such as stratified multi-stage designs), explicit equations are available to compute the sampling variance. For the more common nonlinear estimates with simple or complex sample designs, explicit equations are not generally available and various approximations or computational algorithms are used to provide an essentially unbiased estimate of the sampling variance.

There are two primary forms of sampling variance estimators for complex sample designs: the procedures based on the Taylor series linearization of the nonlinear estimator using explicit sampling variance equations and the procedures based on forming pseudo-replications of the sample. Within the class of pseudo-replications procedures, the balanced repeated replication (BRR) procedure, the jackknife procedure, and the bootstrap procedure are most widely used or discussed.[7] The discussion here will be limited to the Taylor series linearization procedure, BRR, and bootstrap procedures.[8]

This chapter concludes with a section discussing the appropriateness of the different sampling variance estimation techniques for the CTS surveys.

3.1. TAYLOR SERIES LINEARIZATION PROCEDURE

The Taylor series linearization procedure is based on classical statistical method in which a nonlinear statistic can be approximated by a linear combination of the components within the statistic. The accuracy of the approximation is dependent on the sample size and the complexity of the statistic. For most commonly used nonlinear statistics (such as ratios, means, proportions, and regression coefficients), the linearized form has been developed and has good statistical properties. Once a linearized form of an estimate is developed, the explicit equations for linear estimates can be used to estimate the sampling variance. Because the explicit equations can be used, the sampling variance can be estimated using many of the features of the sampling design (e.g., finite population corrections, stratification, multiple stages of selection, and unequal

---

[7] Wolter (1985).

[8] The jackknife procedure is not discussed because of its inherent similarity to BRR.

selection rates within strata). This is the basic variance estimation procedure used in SUDAAN, SAS, Stata, and some other software packages to accommodate many simple and complex sampling designs.

3.2. BALANCED REPEATED REPLICATION PROCEDURE

The balanced repeated replication (BRR) procedure is designed for use with stratified multi-stage sample designs in which two primary sampling units are selected with replacement in each stratum. The full sample of primary sampling units is divided into equal-sized half-samples (pseudo-replicates), and the sampling variance is estimated by computing the variation among the survey estimates calculated for each half-sample. The process for forming the half-samples is constrained to ensure a "balance" among the half-samples. The BRR procedure was developed by the Census Bureau for the estimation of sampling variances before the availability of sophisticated high-speed computers for large national surveys. For some estimates for small subpopulations, the BRR procedure could not compute correct estimates of the sampling variances. To account for this, a modified BRR procedure (Fay's method) is commonly used in which the full sample is used with differential weighting of the half-samples[9].

For sampling designs using simple stratified random samples, without-replacement sample selection with high sampling rates, or certainty selection of primary sampling units, the BRR procedure is not directly appropriate and adaptations are required to produce unbiased sampling variance estimates.[10] In addition, BRR, like other pseudo-replication methods, requires an initial expenditure of effort in forming the replicates, computing a separate set of weights for each replicate, and applying all the nonresponse and poststratification adjustments independently to each replicate. On the other hand, the BRR approach does not require the development of a linearized form of the estimator, so sampling variances can be computed for some forms of complex nonlinear estimates or non-smooth estimators that either cannot be or have not been incorporated in software using the Taylor series linearization procedure.

3.3. BOOTSTRAP PROCEDURE

Whereas the BRR procedure was developed because of the lack of sophisticated high-speed computers, the bootstrap procedure has become more prominent with the increasing availability of such computers. In the classical statistical setting and assuming a simple random sample, the basic bootstrap procedure is to select some number (B) of subsamples, each consisting of a sample of n elements selected with replacement from the original sample. For each of the B subsamples, an estimate is derived from the data, and the variance of the B estimates is the bootstrap variance estimate. Typically, several hundred bootstrap subsamples are used. For any given size of the original sample, a larger value of B results in an estimated variance that is closer to the true variance of the estimate.

For sample surveys, the natural extension of the bootstrap procedure to a stratified sampling design is to select B subsamples independently in each stratum. Even when accounting for the original sampling strata, bootstrapping results in a biased estimate of the sampling variance for

---

[9] Judkins (1990).
[10] Rao and Shao (1996); Rao and Shao (1999).

both simple and complex sample designs.[11] Various complex adaptations of bootstrap procedures have been proposed to remedy the problem for simple survey designs, and these procedures can result in unbiased variance estimates for linear survey estimates. In addition, some of these adaptations can produce unbiased variance estimates for some nonlinear statistics when the statistics are linearized using the Taylor series approximation.[12] The proposed methods of using the bootstrap procedure with sample survey data are still being studied, and no specific method has been fully accepted because none has been shown to be consistently best. In other words, each of the various methods result in biased estimates in different situations.

3.4. VARIANCE ESTIMATION PROCEDURES AND THE CTS DESIGN

Among the variance estimation procedures discussed in this chapter, software using the Taylor series linearization procedure and explicit sampling variance equations offers the best capabilities to account for the complexity of the sample design of the CTS surveys. As discussed above, BRR requires development of replicate weights and is limited in its ability to handle certain CTS design features, such as the dual sample design (i.e., the site and supplemental samples), without-replacement sampling with unequal selection probabilities, high selection rates for primary sampling units, as well as the selection of some primary sampling units with certainty. Because of these and other limitations, the BRR variance estimation procedure was not considered appropriate for estimating the sampling variances for the CTS surveys. Although some forms of bootstrap procedures have shown general equivalence in some situations to the Taylor series procedures using explicit sampling variance equations and BRR procedures, no single bootstrap method seems to be fully accepted by the survey research community, and software reflecting the current methods is not readily available.

---

[11] Sitter (1992).
[12] Sitter (1992); Rao, Wu, and Yue (1992).

CHAPTER 4

STATISTICAL SOFTWARE AND THE CTS DATA

Although there are a number of statistical software packages, only some have the ability to perform analysis that takes into account complex survey design.[13] All of these software packages, as well as those that lack complex survey data analysis capabilities, will generate weighted point estimates (e.g., the estimate of a proportion) that are virtually identical when the correct analysis weights are used. The difference comes in the estimation of standard errors (the square root of the sampling variance) because the packages vary in their capability to accommodate alternative complex sampling situations.

In this chapter, we review four of the commonly used software packages that have at least some complex survey capabilities (SUDAAN, Stata, SAS, and WesVar).[14] These packages vary in terms of their ability to incorporate components of the CTS sample design. They also vary in the number of statistical procedures available that accommodate complex sampling structures. Table 4.1 summarizes the information in this chapter.

4.1. SUDAAN (Version 8)

SUDAAN (Research Triangle Institute, www.rti.org/sudaan/) is a software package designed for the analysis of correlated data from either complex or multi-stage surveys or from clinical or pharmacology experimental studies. SUDAAN uses the first-order Taylor series linearization and design-based variance equations to compute sampling variance estimates. SUDAAN can also estimate variances using pseudo-replication procedures (BRR and jackknife).[15] Unlike the other statistical packages, SUDAAN can accommodate the major features of the CTS design and generate correct standard errors for both national and site-specific estimates (as indicated in Table 4.1).

SUDAAN's routines for complex survey data include descriptive statistics, linear regression, logit (dichotomous, multinomial, and ordered), survival analysis, and log-linear (Poisson) regression. Some of the regression procedures also allow for analysis of longitudinal data using generalized estimating equation methods. SUDAAN's weakness relative to Stata is that Stata has more routines for some forms of multivariate analysis of complex survey data.

[13] See www.fas.harvard.edu/~stats/survey-soft for basic information on statistical software packages for survey data analysis.

[14] Other statistical software packages designed for the analysis of complex survey data include Bascula from Statistics Netherlands, CENVAR and VPLX from U.S. Bureau of the Census, CLUSTERS from University of Essex, Epi Info from Centers for Disease Control, Generalized Estimation System (GES) from Statistics Canada, IVEware (beta version) from University of Michigan, and PC CARP from Iowa State University.

[15] This discussion concerns only the Taylor series linearization method because it is the preferred approach to variance estimation for the CTS data, as was explained in Chapter 3.

4.2. Stata (Version 8)

Stata (Stata Corporation, www.stata.com) is a software package that is primarily for the analysis and graphing of data from a simple random sample. It has been expanded to handle simple and some complex survey designs (the *svy-* commands). Stata uses the Taylor series linearization method with design-based sampling variance equations for calculating variances, and it has routines for computing variances using bootstrap and jackknife procedures.[16]

Because Stata can accommodate with-replacement sampling, it can be used with the CTS data for site-specific estimates (for the Household Survey only) and national estimates from the supplemental sample (for both the Household Survey and the Physician Survey). In other words, for those types of estimates, Stata can fully accommodate all features of the sample design and provides variance estimates identical to those generated by SUDAAN.

Note, however, that for CTS estimates that require assuming without-replacement sampling at the first stage (i.e., national estimates from the site sample, augmented site sample, or combined sample), Stata cannot accommodate the CTS design as fully as SUDAAN. Although Stata can compute sampling variances for some without-replacement sampling, it cannot accommodate the type of without-replacement sampling in the CTS (i.e., two-stage survey design using without-replacement sampling at the first stage and sampling at the second stage). As will be shown in Chapter 5, these limitations cause Stata to typically generate conservative (larger) variance estimates than SUDAAN for CTS national estimates from the site sample, augmented site sample, and combined sample.

Stata includes a broader range of statistical procedures that accommodate complex survey data than the other software packages discussed here. Procedures that incorporate complex design features include descriptive statistics, linear regression, instrumental variables regression, censored and interval regression, negative binomial regression, logit (dichotomous, multinomial, and ordered), probit (dichotomous and ordered), Poisson regression, and Heckman selection models.

Unlike the other packages discussed here, Stata currently fails to generate variance estimates when there is a single observation in a stratum. In the CTS, this situation can be encountered frequently, particularly when a small subpopulation is being examined. When this occurs, the Stata user must either drop the observations in the strata with single observations or manually create a new stratum by combining the strata with one observation to other strata. Neither of these procedures is recommended. Dropping observations can potentially result in biased point and variance estimates, and combining strata can potentially result in biased variance estimates.

[16] This discussion concerns only the Taylor series linearization method because it is the preferred approach to variance estimation for the CTS data, as was explained in Chapter 3.

4.3. SAS (Version 8)

SAS (SAS Institute Inc., www.sas.com) has been expanded to include some of the features needed for complex survey data analysis. The SAS procedures use the Taylor series linearization procedure with design-based sampling variance equations for calculating variances.

In accommodating features of the CTS design, SAS has the same capabilities and limitations as Stata. Because SAS can accommodate with-replacement sampling, it can be used with the CTS data for site-specific estimates (for the Household Survey only) and national estimates from the supplemental sample (for both the Household Survey and the Physician Survey). Like Stata, SAS can compute sampling variances for some without-replacement sampling, but it cannot accommodate the type of without-replacement sampling in the CTS. Therefore, SAS cannot accommodate the CTS design as fully as SUDAAN for CTS estimates that require assuming without-replacement sampling at the first stage (i.e., national estimates from the site sample, augmented site sample, or combined sample). For those types of estimates, using SAS typically results in conservative (larger) variances estimates than SUDAAN.

SAS survey data analysis capabilities are relatively new, and only a limited number of statistical procedures are currently available. Version 8 has two such procedures, *SURVEYMEANS* and *SURVEYREG,* for descriptive statistics and linear regression, respectively. It does not have procedures for estimation of logit models for complex survey data analysis.

4.4. WesVar (Version 4)

WesVar (Westat, www.westat.com/wesvar/) uses pseudo-replication methods (BRR and jackknife procedures) for calculating variances. As discussed in Chapter 3, pseudo-replication methods are not suited to the CTS surveys because of several elements of the design (dual sample design, without-replacement sampling with unequal selection probabilities, high selection rates of PSUs, and certainty selection of PSUs). Researchers have suggested modifications to the standard BRR procedure to adapt it to design elements like those used in the CTS.[17]

[17] Rao and Shao (1996); Rao and Shao (1999).

Table 4.1
Use of Major Statistical Software Packages with the CTS Data

Software Capabilities	Statistical Software Packages			
	SUDAAN (Version 8)	**Stata** (Version 8)	**SAS** (Version 8)	**WesVar** (Version 4)
Method for calculating variances	Taylor series linearization with design-based sampling variance equations[a]	Taylor series linearization with design-based sampling variance equations[a]	Taylor series linearization with design-based sampling variance equations	Pseudo-replication methods
Major CTS sample design features that cannot be accommodated	None	Without-replacement selection of PSUs; high sampling rate of PSUs[b]	Without-replacement selection of PSUs; high sampling rate of PSUs	Without-replacement selection of PSUs; high sampling rate of PSUs; certainty PSUs; stratified simple random samples[c]
CTS estimates: Are major features of sample design accommodated?[d]				
National est., site sample	yes	no	no	no
National est., aug. site sample	yes	no	no	no
National est., combined sample	yes	no	no	no
National est., supp. sample	yes	yes	yes	no
Site est., aug. site sample	yes	yes	yes	no
Statistical procedures available for complex sample design				
Descriptive statistics	yes	yes	yes	yes
Linear regression	yes	yes	yes	yes
Logit models	yes	yes	no	yes
Other regression models	yes	yes	no	no

PSU = primary sampling unit

[a] SUDAAN and Stata also offer pseudo-replication approaches (not discussed here because the Taylor series linearization procedure is the preferred approach to variance estimation for the CTS data, as explained in Chapter 3).
[b] Stata requires at least two observations in each stratum. This can create difficulties when analyzing small subsamples in the CTS data.
[c] Pseudo-replication procedures also require replicates with separate sets of weights, which are not provided in the CTS data files.
[d] See Chapter 2 for definitions of samples.

CHAPTER 5

RESULTS FROM ANALYZING CTS DATA WITH DIFFERENT STATISTICAL SOFTWARE

As discussed in Chapter 4, currently only SUDAAN is able to accommodate the major components of the CTS design (i.e., selection of sites with unequal probability and without replacement). Using the current variance estimation procedures in Stata and SAS with the CTS sample design for national estimates is equivalent to assuming with-replacement sampling at the first stage instead of without-replacement. This chapter compares the standard error estimates that result from using without-replacement (SUDAAN) and with-replacement (Stata and SAS) estimation assumptions to analyze the CTS data.[18]

5.1. TECHNICAL INFORMATION ON COMPARISONS

The estimates used for these comparisons all come from the CTS Household Survey (person-level estimates, from the 2000-01 survey unless otherwise noted) and Physician Survey (physician-level estimates from 2000-01 only). For the combined sample (which is defined in Chapter 2), we calculated national estimates and their associated standard errors using SUDAAN for without-replacement estimation and Stata for with-replacement estimation.[19] There is no reason to expect that our conclusions would differ if we used either the site sample or the augmented site sample instead of the combined sample to calculate the national estimates.

Information on the without-replacement (SUDAAN) and with-replacement (Stata) sampling variance estimation variables that were used in calculating the estimates is provided in Appendix A. For users of the CTS public use and restricted use data files, information on obtaining or constructing the sampling variables for the first three rounds of both surveys is provided in Appendices B and C.

The measure that we use for comparing the results is the relative difference (RelDiff) between the estimated standard errors under the two estimation assumptions:

$$\text{RelDiff} = \frac{(se_{WR} - se_{WOR})}{se_{WOR}} \times 100$$

where se_{WR} is the standard error estimate using the with-replacement assumption and se_{WOR} is the standard error estimate using the without-replacement assumption. In words, the relative difference is the percentage by which the with-replacement (Stata) estimate is larger or smaller than the without-replacement (SUDAAN) estimate.

---

[18] Results from WesVar are not included because it uses pseudo-replication methods to calculate variance estimates (see Chapter 4).

[19] Because SAS and Stata currently have similar same capabilities for variance estimation, comparing without-replacement standard errors from SUDAAN and with-replacement standard errors from Stata is equivalent to comparing without-replacement standard errors from SUDAAN and with-replacement standard errors from SAS.

We used both descriptive and multivariate results for the comparison. For the descriptive results, we chose a set of commonly used variables for which to calculate percentage estimates (i.e., the weighted percentage of persons or physicians with an attribute).[20] We used the full sample and multiple subsamples from each survey, as listed in Table 5.1. For the multivariate results, we used the full sample and one subsample to estimate four multivariate models for each survey.[21]

In each table of results (Tables 5.2 through 5.7), a row is included showing the percentage of estimates for which the relative difference is negative. A negative relative difference means that with-replacement estimation results in a standard error estimate that is smaller than the without-replacement standard error estimate, which means a higher probability of finding a result statistically significant. This increases the likelihood that a Type I error will occur (i.e., rejecting the null hypothesis when it is true).

5.2. RESULTS FOR DESCRIPTIVE ANALYSIS

For the descriptive results, we first examined the relative differences in the standard errors for the full sample and the subsamples. Then we investigated whether there was any relationship between the relative difference and the size of the point estimate or the size of the (sub)sample. Statistical theory says that in general the sampling variance using the with-replacement estimation assumption will be greater than the variance using the without-replacement assumption. As discussed below, this is true to varying extents for all of the samples that we examined except Hispanics.

5.2.1. Descriptive Results by Population Type for the Household Survey

For the Household Survey data, we computed person-level estimates for the full population and three subpopulations: Hispanics, people in low-income households,[22] and uninsured people. Table 5.2 shows a frequency distribution of the relative differences, as well as the mean and median relative differences, for each population. For the full population, the modal relative difference between the with-replacement and the without-replacement standard errors is between 0 and 10 percent, and the median is 9.4 percent, which indicates that the with-replacement estimation assumption tends to produce larger standard errors. However, for 14 percent of the estimates (18 of 125 estimates), the with-replacement estimate of the standard error was smaller than the without- replacement estimate. The results are similar for the two subpopulations of low-income people and the uninsured, although for nearly a third of the estimates of the uninsured, with-replacement standard error estimates are smaller than the without-replacement estimates. The results are noticeably different for the Hispanic subpopulation. For that group, almost two-thirds of the relative differences are negative, and the median is -4.2 percent, which indicates that the standard errors using the with-replacement assumption tend to be smaller than when using the without-replacement assumption.

[20] See Appendix D for a list of the survey questions from which the estimates were derived.
[21] The multivariate models are described in Appendix D.
[22] The low-income population consisted of people in families with family income less than 200 percent of the federal poverty level.

The fact that the results for the Hispanic sample differ from the results for the other samples is likely attributable to the uneven distribution of the Hispanic sample across sites. Three sites (Miami, Phoenix, and Orange County, which are all high-intensity sites) account for 40 percent of the Hispanic sample, with Miami alone accounting for nearly 20 percent.[23] In contrast, neither the low-income sample nor the uninsured sample is clustered in only a few sites. In addition, two sites with a high proportion of Hispanic people (Miami and West Palm Beach, Florida) had a relatively high probability of being selected as CTS sites but a very low probability of both being selected (the joint inclusion probability). The fact that both are in the sample makes variance estimates less stable when using the Yates-Grundy-Sen without-replacement variance estimator, which is the without-replacement variance estimator used in SUDAAN.[24]

5.2.2. Descriptive Results by Population Type for the Physician Survey

For the Physician Survey, we used the full population and five subpopulations for the analysis of relative differences. The subpopulations are physicians in practices with a high proportion of revenue from managed care (more than 40 percent), physicians in solo or two-physician practices, physicians in group practice, primary care physicians (PCPs), and specialists. As shown in Table 5.3, for the full population of physicians, the median relative difference between the with-replacement and without-replacement standard errors is 21.4 percent, indicating that with-replacement estimation will tend to produce standard error estimates that are too large.

The technical reason that the relative differences for the full population in the Physician Survey are so much larger than for the full population in the Household Survey relates to the proportion of the total variance that comes from variation between sites as opposed to variation within sites. Compared to the Household Survey, the Physician Survey has a larger proportion of the total variance that is due to variance between sites. Because the comparison of with-replacement and without-replacement considers the contribution of the variance between primary sampling units (i.e., sites) under the two assumptions and the Physician Survey has a larger proportion of the total sampling variance from the between-site component of variance, the effect of using with-replacement estimates instead of without-replacement is more pronounced for the Physician Survey.

The results for all the subpopulations reflect those for the full population, although there is significant variation by subpopulation. The median relative difference ranges from 28.0 percent for the physicians in group practices to only 4.3 percent for physicians in solo or two-physician practices. Similarly, the percentage of estimates where the with-replacement estimate is smaller than the without-replacement estimate ranges from 3 percent to 44 percent. With 44 percent of the estimates showing a negative relative difference, the subpopulation of physicians in solo or two-physician practices is noticeably different from the other subpopulations.

It is possible that clustering is affecting these results in the same way that it appears to do for the Household Survey. Specifically, the subpopulation for which the with-replacement estimation assumption tends to overstate the standard errors the least (i.e., the physicians in solo or two-

---

[23] High-intensity sites were assigned sample sizes that were approximately four times larger than those for the low-intensity sites. See Chapter 2 for more information on the site sample.

[24] Yates and Grundy (1953); Sen (1953).

physician practices) is also the subsample with pronounced clustering, with three high-intensity sites (Miami, Newark, and Phoenix) accounting for 17 percent of that sample. In contrast, the other subsamples are not as clustered (i.e., not as concentrated in a small number of sites).

5.2.3. Descriptive Results by Size of Estimate

Because the estimates for this analysis are percentage estimates (the weighted percentage of persons or physicians with an attribute), the size of the point estimate can affect the relative differences in the standard error estimates. Specifically, percentage estimates essentially follow a binomial distribution, and the variance for a binomial variable is greatest for estimates near 50 percent and decreases as the estimates tend to zero or to 100 percent. Therefore, for both the Household Survey and the Physician Survey, we categorized the relative differences by the magnitude of the estimates. However, we found no consistent patterns.[25]

5.2.4. Descriptive Results by Size of Sample

Because statistical precision increases with sample size, we also wanted to examine whether the relative differences varied with sample size. We had a variety of sample sizes from the set of estimates in this analysis; within each of the samples listed in Table 5.1, sample size varied across estimates because not every survey question had the same population. We investigated the relative differences in the standard errors by sample size, but again, no consistent pattern emerged.[26]

5.3. RESULTS FOR MULTIVARIATE ANALYSIS

As discussed below, the multivariate results for the full samples are similar to the descriptive results. For the full populations in both surveys, the estimated standard errors calculated using the with-replacement estimation assumption tend to be larger than those calculated using without-replacement. Separate multivariate analyses were also done for the two subpopulations (Hispanics in the Household Survey and physicians in solo or two-physician practices in the Physician Survey) that had descriptive results noticeably different from the results for the corresponding full survey populations; the multivariate results also differ between these subpopulations and the full populations.

5.3.1. Multivariate Results for the Household Survey

Table 5.4 provides the results from four multivariate models for the full sample in the Household Survey. The relative differences tend to be positive, with the medians ranging between 7.8 percent and 13.6 percent for the four models. The results are generally consistent with those for the descriptive estimates for the same population (i.e., all persons) in Table 5.2.

Because the descriptive results for the Hispanic subpopulation (in Table 5.2) were so different from the results for the full population and the other subpopulations, we investigated whether

---

[25] Tables with these results are available from the authors upon request.
[26] Tables with these results are available from the authors upon request.

that difference persisted for multivariate estimates. Table 5.5 shows the results of estimating the same multivariate models for the Hispanic subpopulation. As with the descriptive results, the relative differences for the Hispanic subpopulation are very different from those for the full population (in Table 5.4), showing a much greater tendency for the with-replacement standard error estimate to be smaller than the without-replacement estimate. For example, the percentage of relative differences that are negative ranges from 0% to 25% for the full population, whereas that percentage ranges from 42% to 100% for the Hispanic subpopulation. As discussed above in the section on the descriptive results for the Household Survey, it is possible that clustering is one reason that results for the Hispanic subpopulation differ from those for the full population and the other subpopulations.

5.3.2. Multivariate Results for the Physician Survey

Table 5.6 provides the results from four multivariate models for the full population in the Physician Survey. The relative differences tend to be positive, with the medians ranging between 9.1 percent and 16.5 percent for the four models. The relative differences are generally smaller than those for the descriptive estimates for the same population (i.e., all physicians) in Table 5.3.

Because the descriptive results for the sample of physicians in solo and two-physician practices (in Table 5.3) were so different from the results for the full population and the other subpopulations, we investigated whether that difference persisted for multivariate estimates. Table 5.7 shows the results of estimating the same multivariate models for the subpopulation of physicians in solo or two-physician practices. As with the descriptive results, the relative differences for this subpopulation are noticeably different than those for the full population (in Table 5.6), showing a much greater tendency for the with-replacement standard error estimate to be smaller than the without-replacement estimate. For example, the percentage of relative differences that are negative ranges from 5% to 16% for the full population, whereas that percentage ranges from 15% to 42% for the subpopulation of physicians in solo or two-physician practices. As discussed above in the section on the descriptive results for the Physician Survey, it is possible that clustering is one reason that results for this subpopulation differ from those for the full population and the other subpopulations.

5.4. SUMMARY

Statistical theory says that the sampling variance using the with-replacement estimation assumption will be greater than the sampling variance using the without-replacement assumption. With some exceptions, this appears to be true for data from the CTS Household Survey and Physician Survey. The exceptions revealed by this analysis are the subpopulation of Hispanics in the Household Survey and the subpopulation of physicians in solo and two-physician practices in the Physician Survey.

There are undoubtedly other subpopulations that do not follow statistical theory's general prediction, but it was beyond the scope of this report to identify them. One sample characteristic that might affect the bias is the extent to which the sample is clustered in just a few sites. Clustering is inherent to the design of the CTS, where the 12 high-intensity sites have approximately 45 to 48 percent of the full sample and the remaining sample is spread across the 48 low-intensity sites. Note, however, that although the bias might vary with the amount of clustering, additional investigation would be necessary to determine more fully which sample characteristics determine the bias that can be expected.

Researchers should also note that the degree of overstatement or understatement of standard errors for any particular estimate cannot be known with certainty without specifically calculating the estimates using both with-replacement and without-replacement assumptions. This analysis shows only the nature and level of the bias that will tend to occur across a set of estimates for a specific population.

As discussed more in the following chapter, the fact that the results based on with-replacement estimation tend to differ from those based on without-replacement estimation means that researchers should be cautious when using software that assumes with-replacement sampling. For (sub)populations where the with-replacement estimates tend to overstate the standard errors, there is a decreased likelihood of finding a result to be statistically significant, which decreases the probability of making a Type I error (rejecting the null hypothesis when it is true). However, there is also an increased likelihood of finding that a result is not statistically significant, which increases the probability of making a Type II error (accepting the null hypothesis when it is false). In these cases, with-replacement estimation can be considered to yield "conservative" results because the probability of a Type I error, which researchers typically regard as a more serious concern, is reduced. Nevertheless, since this analysis suggests that the bias that can be expected from with-replacement estimation can vary markedly by subpopulation, the effect of using with-replacement estimation instead of without-replacement for some subpopulations is an increase (possibly substantial) in the likelihood of making a Type I error.

Table 5.1
Sample Sizes and Number of Estimates for Descriptive Results
for Without-Replacement and With-Replacement Analysis

Survey and Sample	Number of Estimates in Descriptive Analysis[a]	Sample Size
CTS Household Survey		
All	125	59,725
Hispanic	112	6,397
Low-income[b]	123	14,428
Uninsured	102	6,462
CTS Physician Survey		
All	31	12,406
In high managed care revenue practices[c]	31	6,219
In solo and two-physician practices	25	4,292
In group practices	25	3,593
Primary care physicians (PCPs)	26	7,673
Specialists	26	4,733

[a] The number of estimates varies across the subpopulation because (1) some variables were not appropriate for the subpopulation, (2) the relative standard error of the estimate was greater than 30 percent, or (3) Stata could not produce an estimate of the standard error because there was at least one stratum with only one PSU (one variable for the Household Survey and four variables for the Physician Survey).
[b] People with family incomes less than 200 percent of the federal poverty level.
[c] Physicians in practice with more than 40 percent of income from managed care.

Table 5.2
Descriptive Results for Without-Replacement and With-Replacement Analysis
Using the CTS Household Survey, by Subpopulation

Summary of Relative Differences[a]	All Households	Subpopulations		
		Hispanics	Low-Income	Uninsured
Distribution of relative differences (number of estimates)				
WR 50% or more larger than WOR	2			
WR 40%–50% larger than WOR	4	2	3	1
WR 30%–40% larger than WOR	6	2	3	2
WR 20%–30% larger than WOR	16	6	7	7
WR 10%–20% larger than WOR	28	12	34	20
WR 0%–10% larger than WOR	51	19	51	40
WR 0%–10% smaller than WOR	15	29	2	29
WR 10%–20% smaller than WOR	3	28	3	3
WR 20%–30% smaller than WOR		14		
Average relative difference	11.8%	–2.9%	8.4%	6.2%
Median relative difference	9.4%	–4.2%	8.1%	6.1%
Percentage of estimates with negative relative difference	14%	63%	20%	30%

WR = with replacement
WOR = without replacement

[a] The relative difference compares the standard error estimates using with-replacement (WR, se_{WR}) and without-replacement (WOR, se_{WOR}) assumptions.

$$\text{RelDiff} = \frac{(se_{WR} - se_{WOR})}{se_{WOR}} \times 100$$

Table 5.3
Descriptive Results for Without-Replacement and With-Replacement Analysis
Using the CTS Physician Survey, by Subpopulation

Summary of Relative Differences[a]	**All Physicians**	**Subpopulations**				
		High Managed Care Revenue	**Solo and Two-Physician Practice**	**Group-Practice**	**PCPs**	**Specialists**
Distribution of relative differences (number of estimates)						
WR 50% or more larger than WOR		1		3		
WR 40%–50% larger than WOR	2	1		1		2
WR 30%–40% larger than WOR	7	2	1	8	2	5
WR 20%–30% larger than WOR	8	6	1	6	6	3
WR 10%–20% larger than WOR	8	15	7	3	5	12
WR 0%–10% larger than WOR	3	5	5	3	9	3
WR 0%–10% smaller than WOR	2	1	3	1	2	1
WR 10%–20% smaller than WOR	1		7		2	
WR 20%–30% smaller than WOR			1			
Average relative difference	21.5%	17.6%	1.7%	27.6%	11.7%	20.7%
Median relative difference	21.4%	13.6%	4.3%	28.0%	10.8%	18.7%
Percentage of estimates with negative relative difference	10%	3%	44%	4%	15%	4%

WR = with replacement
WOR = without replacement

[a] The relative difference compares the standard error estimates using with-replacement (WR, se_{WR}) and without-replacement (WOR, se_{WOR}) assumptions.

$$\text{RelDiff} = \frac{(se_{WR} - se_{WOR})}{se_{WOR}} \times 100$$

Table 5.4
Multivariate Results for Without-Replacement and With-Replacement Analysis
Using the CTS Household Survey, by Model

Summary of Relative Differences[a] for Standard Errors of Regression Coefficients	Dependent Variable and Model			
	Number Ambulatory Visits, Linear	**Cost Concerns Affected Seeking of Medical Care, Linear**	**Health Status, Linear**	**Health Plan Rating, Logit**
Distribution of relative differences (number of estimates)				
WR 50% or more larger than WOR		1(165%)		
WR 40%–50% larger than WOR				
WR 30%–40% larger than WOR		4		1
WR 20%–30% larger than WOR	1	4	1	3
WR 10%–20% larger than WOR	2	6	3	4
WR 0%–10% larger than WOR	6	6	3	13
WR 0%–10% smaller than WOR	2	4		1
WR 10%–20% smaller than WOR	1			1
WR 20%–30% smaller than WOR				
Average relative difference	5.0%	20.3%	11.0%	9.6%
Median relative difference	7.8%	13.6%	11.5%	8.2%
Percentage of estimates with negative relative difference	25%	16%	0%	9%

WR = with replacement
WOR = without replacement

[a] The relative difference compares the standard error estimates of the regression coefficients using with-replacement (WR, se_{WR}) and without-replacement (WOR, se_{WOR}) assumptions.

$$\text{RelDiff} = \frac{(se_{WR} - se_{WOR})}{se_{WOR}} \times 100$$

Table 5.5
Multivariate Results for Without-Replacement and With-Replacement Analysis
Using the CTS Household Survey, by Model,
for Subpopulation of Hispanics

Summary of Relative Differences[a] for Standard Errors of Regression Coefficients	Dependent Variable and Model			
	Number Ambulatory Visits, Linear	Cost Concerns Affected Seeking of Medical Care, Linear	Health Status, Linear	Health Plan Rating, Logistic
Distribution of relative differences (number of estimates)				
WR 50% or more larger than WOR		1(104%)		
WR 40%–50% larger than WOR				
WR 30%–40% larger than WOR				
WR 20%–30% larger than WOR				
WR 10%–20% larger than WOR		3		1
WR 0%–10% larger than WOR		7		10
WR 0%–10% smaller than WOR	3	6	6	10
WR 10%–20% smaller than WOR	3	2	1	1
WR 20%–30% smaller than WOR				
Average relative difference	-9.4%	2.0%	-6.7%	-0.4%
Median relative difference	-8.3%	0.1%	-6.7%	-0.4%
Percentage of estimates with negative relative difference	100%	42%	100%	50%

WR = with replacement
WOR = without replacement

[a] The relative difference compares the standard error estimates of the regression coefficients using with-replacement (WR, se_{WR}) and without-replacement (WOR, se_{WOR}) assumptions.

$$\text{RelDiff} = \frac{(se_{WR} - se_{WOR})}{se_{WOR}} \times 100$$

Table 5.6
Multivariate Results for Without-Replacement and With-Replacement Analysis
Using the CTS Physician Survey, by Model

Summary of Relative Differences[a] for Standard Errors of Regression Coefficients	Dependent Variable and Model			
	Hours of Charity, Linear	**Income, Linear**	**Career Satisfaction, Logit**	**Charity Care, Logit**
Distribution of relative differences (number of estimates)				
WR 50% or more larger than WOR				
WR 40%–50% larger than WOR		1		
WR 30%–40% larger than WOR		1		2
WR 20%–30% larger than WOR	2	2	4	3
WR 10%–20% larger than WOR	6	6	11	11
WR 0%–10% larger than WOR	11	2	12	4
WR 0%–10% smaller than WOR		1	5	1
WR 10%–20% smaller than WOR	1			
WR 20%–30% smaller than WOR				
Average relative difference	9.3%	18.7%	9.3%	14.8%
Median relative difference	9.1%	16.5%	9.3%	13.5%
Percentage of estimates with negative relative difference	5%	8%	16%	5%

WR = with replacement
WOR = without replacement

[a] The relative difference compares the standard error estimates of the regression coefficients using with-replacement (WR, se_{WR}) and without-replacement (WOR, se_{WOR}) assumptions.

$$\text{RelDiff} = \frac{(se_{WR} - se_{WOR})}{se_{WOR}} \times 100$$

Table 5.7
Multivariate Results for Without-Replacement and With-Replacement Analysis
Using the CTS Physician Survey, by Model,
for Subpopulation of Physicians in Solo or Two-Physician Practices

Summary of Relative Differences[a] for Standard Errors of Regression Coefficients	Dependent Variable and Model			
	Hours of Charity, Linear	**Income, Linear**	**Career Satisfaction, Logit**	**Charity Care, Logit**
Distribution of relative differences (number of estimates)				
WR 50% or more larger than WOR				
WR 40%–50% larger than WOR				
WR 30%–40% larger than WOR			1	
WR 20%–30% larger than WOR			1	
WR 10%–20% larger than WOR	3		9	5
WR 0%–10% larger than WOR	4	6	3	6
WR 0%–10% smaller than WOR	5	1	7	2
WR 10%–20% smaller than WOR		1	2	
WR 20%–30% smaller than WOR				
Average relative difference	2.9%	2.3%	5.5%	7.9%
Median relative difference	1.7%	6.1%	9.0%	8.8%
Percentage of estimates with negative relative differences	42%	25%	39%	15%

WR = with replacement
WOR = without replacement

[a] The relative difference compares the standard error estimates of the regression coefficients using with-replacement (WR, se_{WR}) and without-replacement (WOR, se_{WOR}) assumptions.

$$\text{RelDiff} = \frac{(se_{WR} - se_{WOR})}{se_{WOR}} \times 100$$

CHAPTER 6

CONCLUSIONS AND RECOMMENDATIONS FOR RESEARCHERS

Complex survey designs are used in order to achieve analytic goals in the most cost efficient manner and because surveys are often designed to meet multiple analytic goals. Complex survey designs can include clustering, stratification, multiple stages of selection, and different rates of selections for subpopulations. These deviations from simple random sampling will affect the calculation of variances and standard errors. As a result, one implication of complex sampling is that researchers who use the data need to use specialized software that will generate correct variance estimates.

The design of the CTS surveys is more complex than many, in part because it has the dual goals of being able to make both site-specific and national estimates. At present, analysis of CTS data requires software that uses the Taylor series linearization approach with explicit design-based equations for the sampling variance. As described in Chapter 3, software packages that rely solely on balanced repeated replications (BRR) or jackknife procedures are not appropriate for the CTS sample design. The only statistical software package that is capable of accommodating all of the major features of the CTS design at this time is SUDAAN. Its use for analysis of CTS data is strongly urged. However, we realize that there are situations in which use of the SUDAAN software is not practical or where the set of statistical procedures available in SUDAAN might not meet the researcher's analytic needs. In this chapter, we offer some advice for researchers who may find themselves in these types of situations. Because it is impossible to anticipate every type of situation a researcher may face, though, we strongly suggest that researchers discuss their analytic and software choices with statisticians versed in statistics for complex surveys.

Recommendations are summarized in Table 6.1.

6.1. NO ABILITY TO ACCOUNT FOR COMPLEX SAMPLE DESIGN

This section discusses options for those who are considering using the CTS data but lack access to software that takes into account any part of the complex sample design, either in general or for the specific statistical procedures desired.

6.1.1. No Access to Software that Accommodates Any Complex Sample Designs

Situation: I lack access to any software that accommodates complex sample designs.

Practically all statistical software packages will produce equivalent point estimates (e.g., means and proportions, regression coefficients, population totals), assuming the same weight variable is used. However, as long as the user is interested in the precision of statistical estimates (which should be nearly always), statistical packages that fail to accommodate complex survey sampling to analyze CTS data should never be used for the final results (although they may be acceptable for exploratory or preliminary analysis). The use of statistical routines that assume simple random sampling will produce variance estimates that are almost always too small. This will increase the likelihood of rejecting the null hypothesis when it is true (Type I error).

There are two situations where one can obtain approximate or correct standard errors for estimates from the CTS surveys without the benefit of specialized statistical software.

First, the user guides developed for the 1996-97 and 1998-99 CTS household and physician surveys contain standard error look-up tables. These provide approximate standard errors for means and proportions estimated on various subpopulations. However, it should be emphasized that these provide only approximate standard errors, typically based on a sample of estimates. The effects of complex survey design on the precision of statistical estimates can often vary considerably from variable to variable and from subgroup to subgroup. In addition, the look-up tables cannot be used to obtain approximate standard errors of coefficients from multivariate models. The user guides are available as technical publications on the HSC Web site at www.hschange.org. Standard error tables were discontinued after the 1998-99 CTS survey user guides were published. Therefore, they can only be used for approximate standard errors for estimates generated from the 1996-97 and 1998-99 CTS surveys. Sample design and weighting changes in the subsequent CTS surveys preclude their applicability for estimates from more recent data.

Second, CTSonline is a Web-based table generator available on the HSC Web site (www.hschange.org) that may preclude the need for any statistical software. Based on user-supplied specifications, CTSonline provides proportions (continuous variables are categorized) for many of the variables from the CTS Physician Survey, both for the full population and for some subgroups. CTSonline for the CTS Household Survey is expected to be available in the summer of 2003. Correct standard errors generated by SUDAAN for these estimates are also provided. Researchers can use these standard errors to test whether there are significant differences across subgroups or over time (using the appropriate statistical formula) but should be cautioned that these tests will only give approximate results because the tests do not account for any correlation between the estimates that was caused by the survey design.

6.1.2. No Complex Data Analysis Capability for Preferred Statistical Procedure

> *Situation: I want to use a particular statistical procedure but cannot find a statistical software package that has the capability of performing this procedure while accounting for the use of data from a complex survey.*

This situation is most often likely to occur when the researcher wishes to use a sophisticated multivariate estimation procedure. First, researchers should check the latest capabilities of software packages, especially those with survey data analysis capability, since new releases of software programs often add new capabilities.[27] In addition, although it carries some risk, one can occasionally find user-written routines that extend the capabilities of existing software packages. Also, developers of statistical software packages at times make beta versions of new subroutines available as part of the testing process. If these options are unavailable, it is

[27] Descriptions of survey data analysis software are available from the Survey Research Methods Section of the American Statistical Association at www.fas.harvard.edu/~stats/survey-soft/.

advisable to consult with a statistician familiar with analysis of complex survey data who can suggest either alternative approaches or estimation strategies.

6.2. ASSUMING WITH-REPLACEMENT SAMPLING

> *Situation: I want to use a statistical package (such as Stata or SAS) that accommodates the with-replacement design assumption for estimating the sampling variances but not the type of without-replacement design assumption needed for some CTS estimates (either because I lack access to SUDAAN or because I wish to use a statistical procedure not available in SUDAAN).*

SUDAAN has been the software recommended for use with the CTS data because it can accommodate both with-replacement (WR) sampling and the type of without-replacement (WOR) sampling assumption that is appropriate for the CTS sample design. Some other popular statistical packages (e.g., Stata, SAS) have complex survey capabilities for certain statistical routines (such as the *svy* procedures in Stata), but these packages are less useful than SUDAAN for analysis of the CTS data. Therefore, using these packages to analyze the CTS data means assuming WR sampling. This section discusses the types of estimates from the CTS data for which using the WR assumption can be considered an option.

If you do decide to use WR estimation for your analysis of CTS data, see Appendices B and C for information on the sampling variables that you should use.

6.2.1. Site-Specific Estimates

As Table 6.1 indicates, if you only wish to make estimates for specific CTS sites using the Household Survey, the WR specifications in these alternative statistical packages will provide correct standard errors, identical to those obtained using the WR specification in SUDAAN.

In contrast to the Household Survey, site estimates from the Physician Survey require the equal-probability WOR specification to take advantage of the high selection rate of physicians in specific sites. Therefore, using software other than SUDAAN to make site estimates for the Physician Survey should not be considered without investigating the effect of using WR estimation specifically for site estimates and specifically for the (sub)population of interest. Unfortunately, such investigation was beyond the scope of this report.

The reason that the variance estimation assumption is different for the two surveys is the following. For the Household Survey, even though sample selection within each site was done WOR, the sample selected represents such a small proportion of the frame that WR estimation is an appropriate representation of the sampling (the finite population adjustment would have a negligible effect). For the Physician Survey, sample selection within each site was also done WOR. Because the Physician Survey sample is a large enough proportion of the frame in some sites, the WOR variance estimation assumption is used in order to take advantage of the high sampling rate and the finite population adjustment.

6.2.2. National Estimates [28]

As shown in Table 6.1, using the WOR assumption available in SUDAAN is the preferred method for making national estimates from the CTS data because it best reflects the sample design for the CTS surveys. However, if you wish to make national estimates using a software package that cannot accommodate the WOR sampling in the CTS (e.g., Stata and SAS), you should be aware of the results from Chapter 5 of this report, which are summarized in the following two sections.

6.2.2.1. National Estimates for the Full Population

The results in Chapter 5 of this report show that the WR assumption generally produces standard errors that are conservative (i.e., they tend to be larger). This means that using the WR assumption will tend to decrease the probability of finding a result to be statistically significant, which decreases the probability of making a Type I error (rejecting the null hypothesis when it is true). It also tends to increase the probability of making a Type II error (accepting the null hypothesis when it is false), which for researchers is generally less of a concern than a Type I error. However, you should keep in mind that the results in Chapter 5 also show that a sizeable proportion of standard errors estimated using WR assumptions are smaller than those estimated under WOR assumptions. This raises the possibility that there are some situations where use of WR design assumptions increases the likelihood of making a Type I error beyond nominal levels.

For the Physician Survey, as a general statement, differences in standard errors obtained using WR and WOR sampling variance estimation assumptions were smaller for coefficient estimates in multivariate models than for simple estimates of means or proportions. Researchers should note, however, that using multivariate analysis rather than means or proportions does not necessarily reduce the differences between standard errors obtained under WR and WOR assumptions. The effect on the WR-WOR differences in the standard errors from using multivariate analysis is likely to be a function of which variables are included in the multivariate model and the population being studied.

6.2.2.2. National Estimates for Subpopulations

When analysis was conducted on various subpopulations, the likelihood that standard errors using a WR design were smaller than those obtained under WOR assumptions varied considerably. This was particularly evident in the Household Survey for Hispanics and in the Physician Survey for physicians in solo or two-physician practices. Although we suspect this is most likely to occur for subgroups that are heavily clustered in certain sites, analysis to confidently diagnose the reason for the disparities we found across subgroups was beyond the scope of this report. As a result, researchers should avoid doing analysis that uses WR assumptions for the subpopulations we identified as problematic (Hispanics in the CTS Household Survey and physicians in the CTS Physician Survey in solo or two-physician

[28] This discussion concerns national estimates only from the combined sample, site sample, and augmented site sample. The national supplement is excluded because its uses an as independent sample are limited to special cases. The samples are defined in Chapter 2.

practices). This applies to estimates of means and proportions as well as multivariate analysis. Moreover, since we did not systematically investigate all possible subpopulations, researchers are generally cautioned about using WR assumptions in analyses of any subpopulation not specifically examined for this report. The use of WR assumptions in analysis of certain subpopulations could result in a substantially higher probability of Type I errors (rejecting the null hypothesis when it is true). If uncertain how to proceed, researchers should consult with a statistician who is familiar with analysis of complex survey data.

Table 6.1
Summary of Recommendations

Estimate (Geographic Area)	Population	Types of Sample Designs that Software Can Accommodate: Simple Random Sample	Complex Sample Design: Software accommodates WR sampling but not the WOR sampling assumption appropriate for the CTS surveys [Stata, SAS] — Household Survey	Complex Sample Design: Software accommodates WR sampling but not the WOR sampling assumption appropriate for the CTS surveys [Stata, SAS] — Physician Survey	Complex Sample Design: Software accommodates WR sampling and the WOR sampling assumption appropriate for the CTS surveys [SUDAAN]
Site-specific	Full population	no[b]	yes	no[d]	yes[f]
	Subpopulation	no[b]	yes	no[d]	yes[f]
National[a]	Full population	no[b,c]	acceptable (with caution)	acceptable (with caution)	yes[g]
	Subpopulation	no[b,c]	not advisable[e]	not advisable[e]	yes[g]

WR = with replacement
WOR = without replacement

[a] National estimates from combined sample, site sample, or augmented site sample.

[b] For 1996-97 and 1998-99, the user's guides provide standard error look-up tables for use when calculating means and proportions (not useful for multivariate estimates).

[c] CTSonline has standard errors for proportions calculated for selected variables.

[d] Optimal calculation of standard errors for site-specific estimates from the Physician Survey requires the WOR estimation assumption. Determining the effect of using the WR sampling variance estimation assumption for Physician Survey site-specific estimates was beyond the scope of this report.

[e] The effect of using the WR estimation assumption instead of WOR can vary greatly from one subpopulation to another. Use of WR estimation for analysis of a subpopulation is advisable only if investigation of WR and WOR estimation specifically for that subpopulation indicates that the WR estimates tend to be close to the WOR estimates.

[f] For site-specific estimates, use the WR assumption for the Household Survey and the *equal*-probability WOR assumption for the Physician Survey.

[g] For national estimates, use the *unequal*-probability WOR assumption for both the Household Survey and the Physician Survey.

REFERENCES

Judkins, D.R., “Fay’s Method for Variance Estimation,” *Journal of Official Statistics*, Vol. 6, No. 3, pp. 223-239 (1990).

Metcalf, C.E., et al., *Site Definition and Sample Design for the Community Tracking Study*, Technical Publication No. 1, Center for Studying Health System Change, Washington, D.C. (October 1996).

Rao, J.N.K. and J. Shao, “On Balanced Half-Sample Variance Estimation in Stratified Random Sampling,” *Journal of the American Statistical Association*, Vol. 91, pp. 343-348 (1996).

Rao, J.N.K. and J. Shao, “Modified Balanced Repeated Replication for Complex Survey Data,” *Biometrika*, Vol. 86, No. 2, pp. 403-415 (1999).

Rao, J.N.K., C.F. Wu, and K. Yue, “Some Recent Work on Resampling Methods for Complex Surveys,” *Survey Methodology,* Vol. 18, pp. 209-217 (1992).

Sen, A.R., “On the Estimate of the Variance in Sampling with Varying Probabilities,” *Journal of the Indian Society of Agricultural Statistics*, Vol. 5, pp. 119-27 (1953).

Sitter, R.R., “A Resampling Procedure for Complex Survey Data,” *Journal of the American Statistical Association*, Vol. 87, pp. 755-765 (1992)

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

Yates, F. and P.M. Grundy, “Selection Without Replacement from Within Strata with Probability Proportional to Size,” *Journal of the Royal Statistical Society Series B,* Vol. 15, pp. 253-61 (1953).

Survey Methodology Reports: CTS Household Survey

Community Tracking Study Household Survey Methodology Report, 2000-01, forthcoming Technical Publication, Center for Studying Health System Change, Washington, D.C.

Strouse, Richard, Barbara Carlson, and John Hall, *Community Tracking Study Household Survey Methodology Report, 1998-99 (Round Two)*, Technical Publication No. 34, Center for Studying Health System Change, Washington, D.C. (March 2002).

Strouse, Richard, et al., *Community Tracking Study Household Survey Methodology Report, 1996-97 (Round One)*, Technical Publication No. 15, Center for Studying Health System Change, Washington, D.C. (November 1998).

Survey Methodology Reports: CTS Physician Survey

Diaz-Tena, Nuria, et al., *Community Tracking Study Physician Survey Methodology Report, 2000-01,* Technical Publication No. 38, Center for Studying Health System Change, Washington, D.C. (May 2003).

Potter, Frank, et al., *Community Tracking Study Physician Survey Methodology Report, 1998-99 (Round Two),* Technical Publication No. 32, Center for Studying Health System Change, Washington, D.C. (November 2001).

Keil, Linda, et al., *Community Tracking Study Physician Survey Methodology Report, 1996-97 (Round One),* Technical Publication No. 9, Center for Studying Health System Change, Washington, D.C. (October 1998).

User's Guides for the Public Use and Restricted Use Data Files: CTS Household Survey

Community Tracking Study Household Survey Public Use File: User's Guide, 2000-01, Technical Publication No. 41, Center for Studying Health System Change, Washington, D.C. (May 2003).

Community Tracking Study Household Survey Restricted Use File: User's Guide, 2000-01, Technical Publication No. 43, Center for Studying Health System Change, Washington, D.C. (May 2003).

Community Tracking Study Household Survey Public Use File: User's Guide, 1998-99, Technical Publication No. 21, Center for Studying Health System Change, Washington, D.C. (June 2001, revised September 2002).

Community Tracking Study Household Survey Restricted Use File: User's Guide, 1998-99, Technical Publication No. 23, Center for Studying Health System Change, Washington, D.C. (June 2001, revised September 2002).

Community Tracking Study Household Survey Public Use File: User's Guide, 1996-97, Technical Publication No. 7, Center for Studying Health System Change, Washington, D.C. (June 1998, revised July 2000).

Community Tracking Study Household Survey Restricted Use File: User's Guide, 1996-97, Technical Publication No. 17, Center for Studying Health System Change, Washington, D.C. (December 1999, revised July 2000).

User's Guides for the Restricted Use Data Files: CTS Physician Survey

Community Tracking Study Physician Survey Restricted Use File: User's Guide, 2000-01, forthcoming Technical Publication, Center for Studying Health System Change, Washington, D.C.

Community Tracking Study Physician Survey Restricted Use File: User's Guide, 1998-99, Technical Publication No. 27, Center for Studying Health System Change, Washington, D.C. (July 2001, revised December 2001)

Community Tracking Study Physician Survey Restricted Use File: User's Guide, 1996-97, Technical Publication No. 12, Center for Studying Health System Change, Washington, D.C. (October 1998, revised October 2001)

HSC Technical Publications are available on the HSC Web site (www.hschange.org)

APPENDIX A

Calculation of Sampling Variances in this Analysis

APPENDIX A

Calculation of Sampling Variances in this Analysis

This appendix documents the SUDAAN and Stata specifications that were used in calculating the estimates of the sampling variances for this report. For more detailed information about the sampling variables and other specification issues, see Appendices B and C of this report and the user's guides for the CTS public and restricted use data files, which are listed in the References section.

Table A.1
SUDAAN Specifications Used for Estimates in this Analysis
(National Estimates from the Combined Sample)

SUDAAN Statements	**Household Survey**	**Physician Survey**	**Description**
DESIGN=	UNEQWOR	UNEQWOR	UNEQWOR indicates that the first stage units should be treated as though selected without replacement and with unequal probabilities within strata.
DDF=	6500	2900	Because the CTS design includes the selection of some primary sampling units with certainty and the use of a stratified simple random sample as a supplement, the default denominator degrees of freedom in SUDAAN is not an accurate estimate. This specification overrides the default denominator degrees of freedom.
NEST	PSTRATA PPSU SECSTRA NFSUX	PSTRATA PPSU SECSTRA NFSU	PSTRATA and PPSU indicate first stage stratification and the first stage sampling units, respectively. SECSTRA and NFSUX indicate second stage stratification and the second stage sampling units, respectively.
TOTCNT	PSTRTOT3 _ZERO_ _MINUS1_ _ZERO_	PSTRTOT3 _ZERO_ _MINUS1_[29] _ZERO_	PSTRTOT3 indicates to SUDAAN: (a) the frame counts at the first stage of sample selection for noncertainty metropolitan sites, or (b) to compute the sampling variance based on the second stage stratification and units. The term *_MINUS1_* indicates to SUDAAN that variance estimation uses the with-replacement sampling assumption at the second stage. The term *_ZERO_* is a reserved SUDAAN keyword to denote that the corresponding variable in the NEST statement does not contribute to the sampling variance.
JOINTPROB	P1X P2X P3X P4X P5X P6X P7X	P1X P2X P3X P4X P5X P6X P7X	P1X – P7X indicate single inclusion probabilities for each site and joint inclusion probabilities for each possible pair of sites within the noncertainty metropolitan strata of the site sample. For certainty sites, the stratum of nonmetropolitan sites, and the supplemental sample, P1X equals 1.0 and P2X – P7X are not applicable and are assigned missing values.
WEIGHT	WTPER4	WTPHY4	Weights for national estimates from the combined sample.

[29] The SUDAAN specifications for the Physician Survey for 1996-97 and 1998-99 use NFRAME instead of *_MINUS1_* . NFRAME indicates the second stage frame counts for without-replacement selection at the second stage. The 2000-01 survey uses *_MINUS1_* because it has little effect on the standard error estimates (compared to using NFRAME) and simplifies analysis of multiple years of the survey.

Table A.2
Stata Specifications Used for Estimates in this Analysis
(National Estimates from the Combined Sample)

Stata Statements	**Household Survey**	**Physician Survey**	**Description**
strata	STRATAWR	STRATAWR	Stratification variable. Defined using the SUDAAN sampling parameters PSTRATA and SECSTRA. (See Appendices B and C for the definitions.)
psu	PSUWRX	PSUWR	Identifies the sampling unit. Defined using the SUDAAN sampling parameters NFSUX and PPSU for the Household Survey and NFSU and PPSU for the Physician Survey. (See Appendices B and C for the definitions.)
pweight	WTPER4	WTPHY4	Weights for national estimates from the combined sample.

NOTE: In cases where there was only one observation in a stratum, Stata failed to compute a variance estimate. These situations were not included in the analysis.

APPENDIX B

Calculating Sampling Variances in the CTS Household Survey Public Use and Restricted Use Data Files

APPENDIX B

Calculating Sampling Variances in the CTS Household Survey Public Use and Restricted Use Data Files

Because SUDAAN is the software package best able to accommodate the design of the CTS Household Survey, a full explanation of how to calculate estimates with SUDAAN has been included in the user's guides for the public use and restricted use data files for all years. Accordingly, the data files for all years include the variables necessary for specifying the sample design in SUDAAN.

For survey analysis software other than SUDAAN, the appropriate sampling variables and an explanation of how to use them are provided with the public use and restricted use data files only for the 2000-01 Household Survey. Consequently, the purpose of this appendix is to summarize how to use the other software packages discussed in this report (i.e., Stata and SAS, both of which assume with-replacement sampling) to calculate estimates with the CTS Household Survey data for all years. First, we indicate which sampling variables to use for which types of estimates. Then we provide instructions for constructing the sampling variables that were not provided on the public use and restricted use data files for 1996-97 and 1998-99.

Sampling Variables for Survey Software Using the With-Replacement (WR) Sampling Assumption for Variance Estimation

Table B.1 shows which variables to use for weights, stratification, and sampling units for each type of estimate (see Chapter 2 for the definitions of the samples).

The weight variables are provided on the data files for all three years (except for 1996-97 national estimates from the augmented site sample, for which no weights were developed). However, for national estimates from the site sample, combined sample, and augmented site sample, the stratification and analysis unit variables are provided only for 2000-01; they need to be constructed for 1996-97 and 1998-99, using the definitions in Tables B.2 and B.3.

Note that for Household Survey site-specific estimates and national estimates from the supplemental sample, the sampling variables are the same as used for SUDAAN, and so they are already on the data files for all years. Both Stata and SAS produce the same standard error estimates as SUDAAN. This is because the sample design for those two types of estimates can be accommodated equally well by all three software packages.

Constructing Sampling Variables for the CTS Household Survey

As indicated in Table B.1, if you would like to use software other than SUDAAN for making national estimates from the site sample, combined sample, or augmented site sample for the 1996-97 and/or 1998-99 CTS Household Survey, then you will need to construct the variables for stratification and sampling units from the sampling variables for SUDAAN that are already on the data files. The definitions are provided in Tables B.2 and B.3 below, and they are the same definitions that were used to create the variance estimation variables for 2000-01.

Tables B.4 and B.5 provide sample counts that allow you to check whether you have constructed the sampling variables correctly.

Table B.1
Sampling Variables for Variance Estimation
Using the With-Replacement Sampling Assumption
for the CTS Household Survey

	Site-Specific Estimates (Augmented Site Sample)	**National Estimates (National Supp. Only)**	**National Estimates (Site Sample Only)**	**National Estimates (Combined Sample)**	**National Estimates (Augmented Site Sample)**
Weights[a] Stata: pweight SAS: weight	WTPER1	WTPER3	WTPER2	WTPER4	WTPER5[b]
Stratification Stata : strata SAS: stratum	SITE_STR	STRATUM	STRATAWR	STRATAWR	PSTRHWR
Sampling unit Stata: psu SAS: cluster	FSUX	NFSUX	PSUWRX	PSUWRX	PPSUHWRX
Availability of variables	Variables are the same as used for estimation with SUDAAN and are available on the public use and restricted use data files for all years of the Household Survey.		All variables are available on the public use and restricted use data files for the 2000-01 Household Survey. For the 1996-97 and 1998-99 surveys, only the weight variables are available (except for WTPER5 and WTFAM5 for 1996-97). Data users can construct the stratification and sampling unit variables from the SUDAAN variables that are already on the data files using the definitions shown in Tables B.2 and B.3.		

[a] The weights for family-level analysis are WTFAM1, WTFAM2, WTFAM3, WTFAM4, and WTFAM5.
[b] Weights for national estimates from the augmented site sample were not developed for 1996-97.

Table B.2
Definitions of STRATAWR and PSUWRX
for National Estimates from the Site Sample and Combined Sample
in the CTS Household Survey

SITE	PSTRATA	STRATAWR	PSUWRX
1 – 60	1 - 9	(pstrata * 10) + secstra	nfsux
1 – 60	10 – 18	pstrata * 10	ppsu
1 – 60	19	pstrata * 10	nfsux
1 – 60	20	pstrata * 10	ppsu
0	30	(pstrata * 10) + secstra	nfsux

Table B.3
Definitions of PSTRHWR and PPSUHWRX
for National Estimates from the Augmented Site Sample
in the CTS Household Survey

SITEID	PSTRATAH	PSTRHWR	PPSUHWRX
0	n.a.	n.a.	n.a.
1 – 60	1 – 9	(pstratah * 10) + secstrah	nfsuhx
1 – 60	10 – 18	pstratah * 10	ppsuh
1 – 60	19	pstratah * 10	nfsuhx
1 – 60	20	pstratah * 10	ppsuh

n.a. = not applicable (because observations with SITEID = 0 are not in augmented site sample)

Table B.4
Sample Counts for STRATAWR and PSUWRX
for National Estimates from the Site Sample and Combined Sample
in the CTS Household Survey

STRATAWR	PSUWRX	SITE	Household Survey Sample Counts		
			1996-97	1998-99	2000-01
11	[varies]	1 Boston	71	68	99
12	[varies]	1 Boston	187	160	174
13	[varies]	1 Boston	1,742	1,766	1,828
19	[varies]	1 Boston	24	13	56
20	[varies]	34 Philadelphia	569	530	606
30	[varies]	46 Washington DC	551	558	558
40	[varies]	15 Baltimore	527	520	516
50	[varies]	33 New York City	483	491	537
60	[varies]	20 Detroit	562	525	585
70	[varies]	17 Chicago	573	551	516
80	[varies]	22 Houston	546	520	542
90	[varies]	27 Los Angeles	462	516	497
100	101	32 Nassau	662	620	550
	102	8 Newark	2,311	2,263	2,282
	103	28 Middlesex	572	555	565
	104	48 Worcester	586	583	579
	105	16 Bridgeport	548	506	541
110	111	38 Rochester	658	705	786
	112	12 Syracuse	2,363	2,184	2,277
	113	35 Pittsburgh	544	512	526
120	121	5 Lansing	2,291	2,258	2,307
	122	30 Minneapolis	648	607	605
	123	43 St. Louis	590	627	682
	124	4 Indianapolis	2,451	2,274	2,291
	125	29 Milwaukee	524	487	557
130	131	2 Cleveland	2,217	2,116	2,138
	132	18 Columbus	557	532	625
	133	23 Huntington	568	556	548
	134	25 Knoxville	577	545	501
140	141	21 Greensboro	506	471	516
	142	3 Greenville	2,436	2,574	2,280
	143	14 Augusta	563	542	484
	144	13 Atlanta	538	488	416

(continued)

Table B.4 (continued)

STRATAWR	PSUWRX	SITE	Household Survey Sample Counts		
			1996-97	1998-99	2000-01
150	151	47 West Palm Beach	423	434	479
	152	7 Miami	2,031	2,065	2,115
	153	44 Tampa	499	437	546
	154	42 Shreveport	565	557	561
160	161	6 Little Rock	2,644	2,465	2,525
	162	24 Killeen	579	561	517
	163	39 San Antonio	565	540	577
170	171	11 Seattle	2,043	1,792	1,931
	172	36 Portland	557	619	663
	173	41 Santa Rosa	541	512	535
	174	40 San Francisco	431	402	394
	175	31 Modesto	606	615	638
	176	37 Riverside	574	621	617
	177	9 Orange County	2,101	2,057	2,171
180	181	45 Tulsa	588	638	611
	182	19 Denver	558	501	503
	183	26 Las Vegas	481	510	495
	184	10 Phoenix	2,263	2,310	2,090
190	191	57 Eastern Maine	633	605	594
	192	58 Eastern North Carolina	592	540	604
	193	54 Northern Georgia	511	498	465
	194	52 West Central Alabama	606	593	658
	195	53 Central Arkansas	770	723	786
	196	56 Northeast Indiana	565	558	574
	197	55 Northeast Illinois	564	545	572
	198	59 Northern Utah	811	853	937
	199	60 Northwest Washington	611	590	645
200	201	50 Terre Haute	553	493	538
	202	51 Wilmington	541	498	474
	203	49 Dothan	558	619	652
301	[varies]	0 Supplemental Sample	785	820	857
302	[varies]	0 Supplemental Sample	1,120	1,058	984
303	[varies]	0 Supplemental Sample	1,735	1,684	1,515
304	[varies]	0 Supplemental Sample	1,042	1,076	1,018
305	[varies]	0 Supplemental Sample	1,393	1,344	1,314
Total			60,446	58,956	59,725

Table B.5
Sample Counts for PSTRHWR and PPSUHWRX
for National Estimates from the Augmented Site Sample
in the CTS Household Survey

PSTRHWR	PPSUHWRX	SITE	Household Survey Sample Counts	
			1998-99	2000-01
11	[varies]	1 Boston	68	99
12	[varies]	1 Boston	160	174
13	[varies]	1 Boston	1,766	1,828
16	[varies]	1 Boston	83	99
19	[varies]	1 Boston	13	56
20	[varies]	34 Philadelphia	625	706
30	[varies]	46 Washington DC	687	691
40	[varies]	15 Baltimore	581	567
50	[varies]	33 New York City	582	645
60	[varies]	20 Detroit	623	686
70	[varies]	17 Chicago	691	649
80	[varies]	22 Houston	625	613
90	[varies]	27 Los Angeles	719	660
100	101	32 Nassau	689	608
	102	8 Newark	2,308	2,315
	103	28 Middlesex	596	600
	104	48 Worcester	588	587
	105	16 Bridgeport	517	552
110	111	38 Rochester	723	811
	112	12 Syracuse	2,189	2,283
	113	35 Pittsburgh	566	572
120	121	5 Lansing	2,272	2,322
	122	30 Minneapolis	677	661
	123	43 St. Louis	661	727
	124	4 Indianapolis	2,312	2,328
	125	29 Milwaukee	534	600
130	131	2 Cleveland	2,167	2,184
	132	18 Columbus	576	654
	133	23 Huntington	575	559
	134	25 Knoxville	562	516

(continued)

Table B.5 (continued)

PSTRHWR	PPSUHWRX	SITE	Household Survey Sample Counts	
			1998-99	2000-01
140	141	21 Greensboro	491	539
	142	3 Greenville	2,597	2,298
	143	14 Augusta	547	494
	144	13 Atlanta	547	484
150	151	47 West Palm Beach	454	508
	152	7 Miami	2,105	2,137
	153	44 Tampa	495	589
	154	42 Shreveport	569	571
160	161	6 Little Rock	2,478	2,539
	162	24 Killeen	565	523
	163	39 San Antonio	589	616
170	171	11 Seattle	1,832	1,977
	172	36 Portland	658	714
	173	41 Santa Rosa	518	543
	174	40 San Francisco	443	429
	175	31 Modesto	629	653
	176	37 Riverside	679	672
	177	9 Orange County	2,102	2,215
180	181	45 Tulsa	653	623
	182	19 Denver	585	576
	183	26 Las Vegas	534	522
	184	10 Phoenix	2,374	2,141
190	191	57 Eastern Maine	620	605
	192	58 Eastern North Carolina	558	629
	193	54 Northern Georgia	518	498
	194	52 West Central Alabama	597	658
	195	53 Central Arkansas	745	807
	196	56 Northeast Indiana	571	580
	197	52 Northeast Illinois	549	574
	198	59 Northern Utah	862	946
	199	60 Northwest Washington	593	650
200	201	50 Terre Haute	496	541
	202	51 Wilmington	506	481
	203	49 Dothan	623	659
Total			55,417	56,343

APPENDIX C

Calculating Sampling Variances in the CTS Physician Survey Restricted Use Data Files

APPENDIX C

Calculating Sampling Variances in the CTS Physician Survey Restricted Use Data Files

Because SUDAAN is the software package best able to accommodate the design of the CTS Physician Survey, a full explanation of how to calculate estimates with SUDAAN has been included in the user's guides for the restricted use data file for all years.[1] Accordingly, the restricted use data file for each year includes the variables necessary for specifying the sample design in SUDAAN.

For survey analysis software other than SUDAAN, the appropriate sampling variables and an explanation of how to use them are provided with the restricted use data file only for the 2000-01 Physician Survey. Consequently, the purpose of this appendix is to summarize how to use the other software packages discussed in this report (i.e., Stata and SAS, both of which assume with-replacement sampling) to calculate estimates with the CTS Physician Survey data for all years. First, we indicate which sampling variables to use for which types of estimates. Then we provide instructions for constructing the sampling variables that were not provided on the restricted use data files for 1996-97 and 1998-99.

Sampling Variables for Survey Software Using the With-Replacement (WR) Sampling Assumption for Variance Estimation

Table C.1 shows which variables to use for weights, stratification, and sampling units for each type of estimate (see Chapter 2 for the definitions of the samples).

The weight variables are provided on the data files for all three years. However, for national estimates from the combined sample and augmented site sample, the stratification and analysis unit variables are provided only for 2000-01; they need to be constructed for 1996-97 and 1998-99, using the definitions in Tables C.2 and C.3.

For Physician Survey site-specific estimates, only SUDAAN can estimate variances correctly because of the high sampling rates of physicians in some sites. Investigation of how Stata and SAS variance estimates (for site-specific estimates) differ from SUDAAN estimates was beyond the scope of this report, and so we currently cannot provide guidance for using any other statistical software packages besides SUDAAN.

Note that for national estimates from the supplemental sample, the sampling variables are the same as used for SUDAAN, and so they are already on the data files for all years. Both Stata and SAS produce the same standard error estimates as SUDAAN. This is because the sample design variance estimation assumption for those estimates can be accommodated equally well by all three software packages.

---

[1] Because of confidentiality concerns about not revealing the identities of the survey respondents, sampling variables are not included on the public use versions of the Physician Survey data files.

Constructing Sampling Variables for the CTS Physician Survey

As indicated in Table C.1, if you would like to use software other than SUDAAN for making national estimates from the site sample, combined sample, or augmented site sample for the 1996-97 and/or 1998-99 CTS Physician Survey, then you will need to construct the variables for stratification and sampling units from the sampling variables for SUDAAN that are already on the data files. The definitions are provided in Tables C.2 and C.3 below. The variance estimation variables for 2000-01 were created from those definitions.

Tables C.4 and C.5 provide sample counts that allow you to check whether you have constructed the sampling variables correctly.

Table C.1
Sampling Variables for Variance Estimation
Using the With-Replacement Sampling Assumption
for the CTS Physician Survey

	Site-Specific Estimates (Augmented Site Sample)	**National Estimates (National Supp. Only)**	**National Estimates (Site Sample Only)**	**National Estimates (Combined Sample)**	**National Estimates (Augmented Site Sample)**
Weights Stata: pweight SAS: weight	WTPHY1	WTPER3	WTPHY2[a]	WTPHY4	WTPHY5
Stratification Stata : strata SAS: stratum	Not available.	NSTRATA	STRATAWR	STRATAWR	PSTRAWR
Sampling unit Stata: psu SAS: cluster	Not available.	NFSU	PSUWR	PSUWR	PPSUAWR
Availability of variables	Not available.[b]	Variables are the same as used for estimation with SUDAAN and are available on the restricted use data files for all years of the Physician Survey.	All variables are available on the restricted use data file for the 2000-01 Physician Survey. For the 1996-97 and 1998-99 surveys, only the weight variables are available. Data users can construct the stratification and sampling unit variables from the SUDAAN variables that are already on the data files using the definitions shown in Tables C.2 and C.3.		

[a] Weights for national estimates from the site sample are provided only for 1996-97.
[b] For Physician Survey site-specific estimates, only SUDAAN can estimate variances correctly. Investigation of how Stata and SAS variance estimates (for site-specific estimates) differ from SUDAAN estimates was beyond the scope of this report, and so we currently cannot provide guidance for using any other statistical software packages besides SUDAAN.

Table C.2
Definitions of STRATAWR and PSUWR
for National Estimates from the Site Sample and Combined Sample
in the CTS Physician Survey

Survey Year	PSTRATA	SECSTRA	STRATAWR	PSUWR
1996-97	1 - 9	all values	(pstrata * 10) + secstra	nfsu
	10 – 18	all values	pstrata * 10	ppsu
	19	all values	pstrata * 10	nfsu
	20	all values	pstrata * 10	ppsu
	30	21	311	nfsu
		22	312	
		23	321	
		24	322	
		25	331	
		26	332	
		27	341	
		28	342	
		29	351	
		30	352	
		31	361	
		32	362	
		33	371	
		34	372	
		35	381	
		36	382	
		37	391	
		38	392	
		39	401	
		40	402	

(continued)

Table C.2 (continued)

Survey Year	PSTRATA	SECSTRA	STRATAWR	PSUWR
1998-99 and 2000-01	1 – 9	1 or 2	(pstrata * 10) + secstra	nfsu
	1	3	11	
	1	4	12	
	2	3	21	
	2	4	22	
	3	3	31	
	3	4	32	
	4	3	41	
	4	4	42	
	5	3	51	
	5	4	52	
	6	3	61	
	6	4	62	
	7	3	71	
	7	4	72	
	8	3	81	
	8	4	82	
	9	3	91	
	9	4	92	
	10 – 18	all values	pstrata * 10	ppsu
	19	all values	pstrata * 10	nfsu
	20	all values	pstrata * 10	ppsu
	30	11 or 13	311	nfsu
		12 or 14	312	
		21 or 23	321	
		22 or 24	322	
		31 or 33	331	
		32 or 34	332	
		41 or 43	341	
		42 or 44	342	
		51 or 53	351	
		52 or 54	352	
		61 or 63	361	
		62 or 64	362	
		71 or 73	371	
		72 or 74	372	
		81 or 83	381	
		82 or 84	382	
		91 or 93	391	
		92 or 94	392	
		101 or 103	401	
		102 or 104	402	

Table C.3
Definitions of PSTRAWR and PPSUAWR
for National Estimates from the Augmented Site Sample
in the CTS Physician Survey

Survey Year	**SUBGRP**	**ASTRATA**	**ASECSTRA**	**PSTRAWR**	**PPSUAWR**
1996-97	A	1 – 9	all values	(astrata * 10) + asecstra	afsu
		10 – 18	all values	astrata * 10	apsu
		19	all values	astrata * 10	afsu
		20	all values	astrata * 10	apsu
	B, C, D	n.a.	n.a.	n.a.	n.a.
1998-99 and 2000-01	A or C	1 – 9	1 or 2	(astrata * 10) + asecstra	afsu
		1	3	11	
		1	4	12	
		2	3	21	
		2	4	22	
		3	3	31	
		3	4	32	
		4	3	41	
		4	4	42	
		5	3	51	
		5	4	52	
		6	3	61	
		6	4	62	
		7	3	71	
		7	4	72	
		8	3	81	
		8	4	82	
		9	3	91	
		9	4	92	
		10 – 18	all values	astrata * 10	apsu
		19	all values	astrata * 10	afsu
		20	all values	astrata * 10	apsu
	B or D	n.a.	n.a.	n.a.	n.a.

n.a. = not applicable (because only observations with SUBGRP = A in 1996-97 and SUBGRP = A or C in 1998-99 and 2000-01 are used for national estimates from the augmented site sample)

Table C.4
Sample Counts for STRATAWR and PSUWR
for National Estimates from the Site Sample and Combined Sample
in the CTS Physician Survey

STRATAWR	PSUWR	SITE	Physician Survey Sample Counts		
			1996-97	1998-99	2000-01
11	PHYSID	1 Boston	414	408	356
12	PHYSID	1 Boston	225	182	176
21	PHYSID	34 Philadelphia	75	92	97
22	PHYSID	34 Philadelphia	45	48	40
31	PHYSID	46 Washington DC	104	90	90
32	PHYSID	46 Washington DC	62	45	38
41	PHYSID	15 Baltimore	89	92	93
42	PHYSID	15 Baltimore	50	50	46
51	PHYSID	33 New York City	92	53	90
52	PHYSID	33 New York City	45	39	43
61	PHYSID	20 Detroit	83	85	102
62	PHYSID	20 Detroit	48	37	39
71	PHYSID	17 Chicago	91	78	99
72	PHYSID	17 Chicago	49	41	36
81	PHYSID	22 Houston	83	92	94
82	PHYSID	22 Houston	58	47	41
91	PHYSID	27 Los Angeles	72	60	94
92	PHYSID	27 Los Angeles	41	36	35
100	101	32 Nassau	128	139	119
	102	8 Newark	549	567	493
	103	28 Middlesex	155	140	126
	104	48 Worcester	130	132	147
	105	16 Bridgeport	129	150	144
110	111	38 Rochester	129	124	135
	112	12 Syracuse	388	398	370
	113	35 Pittsburgh	143	141	135
120	121	5 Lansing	307	322	332
	122	30 Minneapolis	147	136	136
	123	43 St. Louis	137	130	138
	124	4 Indianapolis	520	496	454
	125	29 Milwaukee	144	131	127
130	131	2 Cleveland	518	516	482
	132	18 Columbus	136	136	135
	133	23 Huntington	89	114	112
	134	25 Knoxville	120	117	121

(continued)

Table C.4 (continued)

STRATAWR	PSUWR	SITE	Physician Survey Sample Counts		
			1996-97	1998-99	2000-01
140	141	21 Greensboro	144	150	152
	142	3 Greenville	402	372	387
	143	14 Augusta	123	120	134
	144	13 Atlanta	149	147	155
150	151	47 West Palm Beach	111	118	130
	152	7 Miami	446	435	492
	153	44 Tampa	123	133	124
	154	42 Shreveport	120	118	132
160	161	6 Little Rock	373	342	353
	162	24 Killeen	98	104	102
	163	39 San Antonio	120	145	135
170	171	11 Seattle	524	498	509
	172	36 Portland	125	130	133
	173	41 Santa Rosa	116	122	126
	174	40 San Francisco	124	143	158
	175	31 Modesto	100	101	111
	176	37 Riverside	127	99	129
	177	9 Orange County	506	538	404
180	181	45 Tulsa	129	130	133
	182	19 Denver	140	139	143
	183	26 Las Vegas	111	121	113
	184	10 Phoenix	493	465	491
190	99952	52 West Central Alabama	27	26	33
	99953	53 Central Arkansas	105	107	127
	99954	54 Northern Georgia	109	109	117
	99955	55 Northeast Illinois	93	93	91
	99956	56 Northeast Indiana	60	76	81
	99957	57 Eastern Maine	121	121	128
	99958	58 Eastern North Carolina	94	105	112
	99959	59 Northern Utah	92	99	136
	99960	60 Northwest Washington	107	109	102
200	201	50 Terre Haute	56	70	73
	202	51 Wilmington	86	101	104
	203	49 Dothan	61	66	73

(continued)

Table C.4 (continued)

STRATAWR	PSUWR	SITE	Physician Survey Sample Counts		
			1996-97	1998-99	2000-01
311	PHYSID	0 Supplemental Sample	45	27	29
312	PHYSID	0 Supplemental Sample	37	39	41
321	PHYSID	0 Supplemental Sample	54	44	37
322	PHYSID	0 Supplemental Sample	55	57	63
331	PHYSID	0 Supplemental Sample	58	45	49
332	PHYSID	0 Supplemental Sample	44	67	64
341	PHYSID	0 Supplemental Sample	79	50	55
342	PHYSID	0 Supplemental Sample	73	82	85
351	PHYSID	0 Supplemental Sample	69	43	55
352	PHYSID	0 Supplemental Sample	61	63	75
361	PHYSID	0 Supplemental Sample	77	59	60
362	PHYSID	0 Supplemental Sample	67	69	70
371	PHYSID	0 Supplemental Sample	52	40	48
372	PHYSID	0 Supplemental Sample	60	57	61
381	PHYSID	0 Supplemental Sample	51	41	45
382	PHYSID	0 Supplemental Sample	64	60	62
391	PHYSID	0 Supplemental Sample	67	60	57
392	PHYSID	0 Supplemental Sample	75	67	71
401	PHYSID	0 Supplemental Sample	70	50	61
402	PHYSID	0 Supplemental Sample	60	68	80
Total			12,528	12,304	12,406

Table C.5
Sample Counts for PSTRAWR and PPSUAWR
for National Estimates from the Augmented Site Sample
in the CTS Physician Survey

PSTRAWR	**PPSUAWR**	**Physician Survey Sample Counts**		
		1996-97	**1998-99**	**2000-01**
11	PHYSID	404	398	340
12	PHYSID	217	188	180
21	PHYSID	74	95	98
22	PHYSID	44	61	54
31	PHYSID	103	93	94
32	PHYSID	62	60	53
41	PHYSID	89	102	96
42	PHYSID	49	55	49
51	PHYSID	90	74	99
52	PHYSID	43	61	65
61	PHYSID	81	85	101
62	PHYSID	48	47	51
71	PHYSID	90	88	109
72	PHYSID	49	63	56
81	PHYSID	83	95	99
82	PHYSID	56	52	48
91	PHYSID	69	70	108
92	PHYSID	39	61	66
100	101	127	141	124
	102	498	510	433
	103	151	142	123
	104	122	125	130
	105	119	140	126
110	111	124	126	135
	112	378	372	338
	113	139	144	136
120	121	292	282	284
	122	146	139	133
	123	130	140	139
	124	497	473	414
	125	143	136	124
130	131	503	488	440
	132	133	134	131
	133	83	99	97
	134	112	109	106

(continued)

Table C.5 (continued)

PSTRAWR	PPSUAWR	Physician Survey Sample Counts		
		1996-97	1998-99	2000-01
140	141	138	139	139
	142	395	350	350
	143	118	115	121
	144	148	153	165
150	151	102	108	114
	152	432	412	456
	153	119	130	121
	154	112	102	111
160	161	360	323	315
	162	94	94	83
	163	115	134	124
170	171	507	480	482
	172	125	132	137
	173	110	113	117
	174	109	132	146
	175	96	94	103
	176	123	102	127
	177	497	533	395
180	181	120	115	119
	182	136	141	137
	183	109	126	120
	184	479	456	469
190	99952	26	24	28
	99953	101	103	121
	99954	104	103	105
	99955	89	86	78
	99956	55	69	70
	99957	113	106	109
	99958	93	94	95
	99959	80	78	108
	99960	97	105	95
200	201	55	66	63
	202	78	93	91
	203	59	61	66
Total		10,881	10,920	10,659

APPENDIX D

Detailed Information on Means and Multivariate Models

APPENDIX D

Detailed Information on Means and Multivariate Models

Survey questions used for calculating means

Household Survey (2000-01 person-level estimates)

Gender (Section A)
Age (Section A)
Highest grade completed (Section A)
Whether covered by employer-sponsored health insurance (question b1a)
Whether covered by private insurance purchased directly (question b1b)
Whether covered by private insurance through someone not in household (question b1c)
Whether covered by Medicare (question b1d)
Whether covered by Medicaid (question b1e)
Whether covered by Medicare supplemental or Medigap policy (question b59)
Whether ever enrolled in an HMO (question b901)
Willingness to accept limited choice of physicians and hospitals to save money (question b951)
Insurance type (questions in Section B)
Whether private insurance plan has network of providers (questions in Section B)
Previous type of insurance (questions in Section B)
Number of overnight hospital stays (question c121)
Number of emergency room visits without hospital admission (question c221)
Number of doctor visits (question c311)
Number of visits to non-physician medical professionals (question c331)
Whether person had any mental health visits (question c511)
Whether person had a flu shot (question c531)
Whether person had a mammogram (question c611)
Whether person put off getting needed medical care (question c821)
Whether person has usual source of care (question d101)
Type of place for usual source of care (question d111)
Trust that doctor will put medical needs first (question d321)
Rating of doctor's explanation (question e321)
Assessment of whether more likely to take risks than average (question e521)
Whether doctor advised person to quite smoking (question e671)
Satisfaction with choice of primary care physician (questions in Section E)
Satisfaction with choice of specialists (questions in Section E)
Satisfaction with family's health care (questions in Section E)
Whether last doctor visit was for check-up (questions in Section E)
Time spent waiting in office until seen by medical professional (questions in Section E)
General health status (questions in Section E)
SF-12 physical component summary score (questions in Section E)
Whether person worked for pay last week (question f111)

Employer type (question f201)
Hourly wage (questions in Section F)
Employer size in number of employees at all locations (questions in Section F)
Employer industry (questions in Section F)
Offerings of, eligiblity for, and coverage by employer-sponsored insurance (questions in Section F)
Whether employer offers multiple insurance plans (questions in Section F)
Annual family income (question g10)
Race/ethnicity (questions in Section G)

Physician Survey (2000-01)

Gender
Whether physician provides patient care in more than one practice (question A4)
Satisfaction with overall career in medicine (question A19)
Board certification status of physician (questions in Section A)
Whether physician is full owner, part owner, or not an owner of practice (question C1)
Number of outside owners of practice (questions in Section C)
Physician’s practice type (questions in Section C)
Effect of practice guidelines on physician’s practice of medicine (question D4A)
Effect of practice profiles on physician’s practice of medicine (question D4B)
Effect of patient satisfaction surveys on physician’s practice of medicine (question D4C)
Appropriateness of complexity/severity of patients’ conditions for which PCPs are expected to provide care without referral (question D8)
Freedom to make clinical decisions (question F1C)
Ability to provide high quality care (question F1D)
Ability to make clinical decisions in patients’ best interest (question F1E)
Ability to obtain non-emergency hospital admissions (question F8C)
Ability to obtain high quality outpatient mental health services (question F8G)
Acceptance of new Medicaid patients in physician’s practice (question F9B)
Whether physician is eligible for bonuses (questions in Section H)
Whether physician has fixed salary or compensation based on time worked, or whether physician is compensated on some other basis (questions in Section H)
Whether practice profiles are risk adjusted (questions in Section H)
Whether compensation is affected by measures of quality of care (questions in Section H)
Whether compensation is affected by patient satisfaction surveys (questions in Section H)
Whether compensation is affected by physician’s own productivity (questions in Section H)
Whether compensation is affected by practice profiling (questions in Section H)

Multivariate Models

Household Survey (1998-99 or 2000-01 data, depending on model; person-level analysis)

Number of ambulatory visits in the previous year (i.e., visits with doctor or medical professional or visits to emergency room without subsequent hospital admission)

Independent variables
- Gender
- Race and ethnicity
- Family income as percentage of federal poverty line
- Self-report of overall health status
- Type of health insurance

Whether person postponed or did not get medical care during the previous year because of concern about cost

Independent variables
- Age
- Annual family income
- Per capita family income
- Gender
- Eduction
- Self-report of overall health status
- Propensity to take risks
- Family composition
- Large metropolitan area or small metropolitan area or non-metropolitan area
- Insured or uninsured
- Proximity to community health center, federally qualified health center, health care for the homeless program, migrant health program, or public housing program
- Proximity to public hospital
- Proximity to emergency room
- Number of physicians in local area

Health status (SF-12 physical component summary score[1])

Independent variables
- Whether person is in HMO
- Family income as percentage of federal poverty line
- Age
- Gender

---

[1] See Ware, J.E., M. Kosinski, and S.D. Keller, *How to Score the SF-12 Physical and Mental Health Summary Scales*, Second Edition, Boston, MA: The Health Institute, New England Medical Center (1995).

Whether person gives health plan a high rating

Independent variables

- Whether in a managed care plan with gatekeeping
- Chronic conditions
- Family income as a percentage of federal poverty line
- Education
- Self-report of overall health status
- Age
- Gender
- Marital status
- Propensity to take risks
- Willingness to accept restrictions on choice of providers to reduce costs
- Private insurance coverage from employer or purchased directly or provided by others

Physician Survey (2000-01 data)

Number of hours in previous month spent providing charity care

Independent variables

- Gender
- Foreign medical school graduate
- Number of years in practice
- Practice type
- Primary specialty category
- Full-/part-owner or not an owner of practice
- Hours in previous week spent in direct patient care
- Percent of patient care practice revenue that comes from Medicaid
- Percent of patient care practice revenue that comes from managed care

Net income from practice of medicine (after expenses but before taxes). Excludes approximately 3,000 full owners of solo practices.

Independent variables

- Gender
- Foreign medical school graduate
- Number of years in practice
- Practice type
- Primary specialty category
- Board certification status
- Medical doctor or osteopath

Whether physician was “very satisfied” with overall career in medicine

Independent variables

- Gender
- Foreign medical school graduate
- Number years in practice
- Practice type
- Full-/part-owner or not an owner of practice
- Fixed or variable compensation
- Percent of patient care practice revenue that comes from managed care
- Net income from practice of medicine (after expenses but before taxes)
- Ability to obtain services (e.g., non-emergency hospital admission, referrals to high-quality specialists)
- Freedom to make clinical decisions that meet patients’ needs
- Ability to make clinical decisions in best interest of patients without possibility of reducing income
- Ability to provide high-quality care to all patients
- Adequacy of time to spend with patients during typical visit
- Ability to communicate sufficiently with other physicians to deliver high-quality care
- Ability to maintain continuing relationships with patients over time that promote the delivery of high-quality care
- Number of weeks worked in previous year
- Number of hours in previous week spent in medically-related activities
- Number of hours in previous month spent providing charity care
- Census region
- Metropolitan area or non-metropolitan area

Whether physician provided any charity care in previous month

Independent variables

- Gender
- Foreign medical school graduate
- Number of years in practice
- Practice type
- Primary specialty category
- Full-/part-owner or not an owner of practice
- Board certification status
- Hours in previous week spent in direct patient care
- Percent of patient care practice revenue that comes from Medicaid
- Percent of patient care practice revenue that comes from managed care