Michael Davern
Arthur Jones Jr.
James Lepkowski
Gestur Davidson
Lynn A. Blewett

# Unstable Inferences? An Examination of Complex Survey Sample Design Adjustments Using the Current Population Survey for Health Services Research

*Statistical analysis of the Current Population Survey's Annual Social and Economic Supplement is used widely in health services research. However, the statistical evidence cited from the Current Population Survey (CPS) is not always consistent because researchers use a variety of methods to produce standard errors that are fundamental to significance tests. This analysis examines the 2002 Annual Social and Economic Supplement's (ASEC) estimates of national and state average income, national and state poverty rates, and national and state health insurance coverage rates. Findings show that the standard error estimates derived from the public use CPS data perform poorly compared with the survey design-based estimates derived from restricted internal data, and that the generalized variance parameters currently used by the U.S. Census Bureau in its ASEC reports and funding formula inputs perform erratically. Because the majority of published research (both by academics and Census Bureau analysts) does not make use of the survey design-based information available only on the internal ASEC data file, we argue that the Census Bureau ought to use alternative methods for its official ASEC reports. We also argue that for public use data the Census Bureau should produce a set of replicate weights for the ASEC or release a set of sample design variables that incorporate statistical "noise" to maintain respondent confidentiality (e.g., pseudo-primary sampling units) as other federal government surveys do. This is essential to make appropriate inferences using the ASEC data regarding statistical significance and estimate variance for health policy analysis.*

The Annual Social and Economic Supplement (ASEC) of the Current Population Survey (CPS) is used widely for health research and policy analysis because it produces timely estimates, makes its microdata available to the public soon after collection, and can be used to produce both national and state estimates of income, poverty, and health insurance coverage (Blewett et al. 2004). However, there is a major obstacle for analysts working with the CPS ASEC data that has not been adequately addressed. The standard errors of important estimates are calculated

***Michael Davern, Ph.D.,*** *is an assistant professor;* ***Gestur Davidson, Ph.D.,*** *is a senior research associate; and* ***Lynn A. Blewett, Ph.D.,*** *is an associate professor, at the School of Public Health, University of Minnesota.* ***Arthur Jones Jr., M.A.,*** *is a statistician at the U.S. Census Bureau.* ***James Lepkowski, Ph.D.,*** *is a research professor at the Institute for Social Research, University of Michigan.* *This research was supported by the Robert Wood Johnson Foundation (grant no. 038846).* *Address correspondence to Prof. Davern at the School of Public Health, University of Minnesota, 2221 University Ave., S.E., Suite 345, Minneapolis, MN 55414. Email: daver004@umn.edu*

*Inquiry* 43: 283–297 (Fall 2006). 
0046-9580/06/4303–0283
www.inquiryjournal.org

using various approaches that adjust for the complex sample design of the CPS ASEC; however, these approaches have not been evaluated in such a way as to give analysts an idea of which is the best to use. This is especially true for researchers who perform their research on the non-internal files of the U.S. Census Bureau (i.e., the public use files).[1]

The CPS ASEC is used to evaluate state and federal policies, determine trends in key economic indicators, and to monitor disparities (among many other topics). As a result of inappropriate standard error calculations, analysts may find significant levels of crowd-out, a drop in the uninsurance rate, or an increase in the public program participation rate when there may not actually be statistical significance. The data also are used in critical analyses performed by the Congressional Budget Office (CBO) to score pieces of federal health care legislation based on projected costs (Glied, Remler, and Zivin 2002). The CBO scores can impact greatly the likelihood of legislation passing Congress (Glied, Remler and Zivin 2002). Finally, the data also are used to distribute funding to states for the State Children's Health Insurance Program (SCHIP) and Title I education funding (Davern, Blewett et al. 2003). Because of the data's use for important scholarly research and official federal government purposes, such as attaching costs to legislative proposals and allocating funds for federal programs, we investigate how well various approaches perform.

Producing accurate standard errors is essential for both the scholarly research and official policy uses of the data because they indicate the precision of the estimates and the statistical significance of hypothesis tests (e.g., whether estimates of poverty differ from one year to the next, or whether one state has a higher poverty rate than another). Statistical significance provides the standard of evidence for statistical arguments, and the errors allow us to gauge our level of uncertainty associated with specific estimates. In theory, standard errors are relatively easy to compute if samples have been collected using simple random sampling. However, the CPS is based on a complex, multistage sample design and it needs to be accounted for when calculating standard errors. Failure to account for the stratification, clustering, and weighting used in the CPS generally results in serious underestimation of standard errors (Kish 1992, 1995; Lohr 2000).

We evaluate strategies used for approximating standard errors by both U.S. Census Bureau analysts in their official reports and by non-Census Bureau scholars working with the public use data. We limit our analysis to three major concepts measured as part of the CPS ASEC and used in health services research and policy: income, poverty status, and health insurance coverage. The findings from this paper should help policy analysts make informed decisions about which type of standard error estimation is most appropriate for their research and will help the Census Bureau understand the needs of the health policy research community.

## Sample Design, Standard Errors and the CPS ASEC

There are three important elements that determine the effect of the CPS complex survey sample design on standard errors: clustering, stratification, and weighting. Cluster sampling involves the grouping of the population into convenient aggregations of observations, such as people in households, households in blocks, and blocks in counties. The sample of elements is drawn from some of these clusters at the exclusion of others (Kish 1995). Stratification is also a grouping of elements, or clusters, but in this case elements or clusters are drawn from each stratum (that is, all strata are included in the sample), sometimes at different sampling rates (Kish 1995). For example, in a given strata one in 2,000 households is sampled, whereas in others one in 1,000 households is sampled. Finally, weighting is a technique for adjusting sample data to correct for design features such as oversampling and design deficiencies such as nonresponse. Base probability weights are the inverse probability of being selected into the sample. For example, if a person has a one in 1,000 probability of selection, the weight is 1,000. Weights can increase the variance of estimates when some population elements have a higher weight than others (Kish 1992). The ratio of an estimated sampling variance that takes these components into account to an estimated sampling variance that ignores clustering, weighting, and stratification is called the design effect (Kish 1995). In most cases, the standard errors calculated that take clustering, stratification, and weighting into account are larger than those that do not; the design effect therefore is usually greater than 1 for complex sample surveys.[2]

The effect of clustering is driven by the intraclass correlation coefficient ($\rho$)—which expresses the correlation between members of a sampled cluster (e.g., household), or the percentage of the total variance found between clusters—and by the size of the cluster (Kish 1995). The design effect due to clustering is determined by:

$$1 + \rho(b - 1). \tag{1}$$

Here $\rho$ is the intraclass correlation and "$b$" is the size of the cluster. In cases where the intraclass correlation is "0" the design effect is simply 1. However, when the intraclass correlation coefficient is greater than "0" the design effect due to clustering will be greater than 1.

When using information from a data set like the CPS ASEC that includes clustered observations, the intraclass correlation coefficient will vary across statistics. For example, in the CPS ASEC data everyone in a sampled housing unit is in the sample, so a housing unit is a cluster. When developing estimates for concepts that are highly correlated within a household, such as whether a person is in poverty or covered by health insurance, the intraclass correlation will be larger. For other concepts like personal income, the intraclass correlation coefficient may be lower – knowing the income of one person in the household does not provide reliable information about the earnings of other people in the household. On the other hand, knowing whether one person in the household is in poverty is highly related to whether another person in the same household is also in poverty, since entire families are assigned the same poverty status.

The design effect can be decreased under some forms of stratification (Kish 1992, 1995). Stratification can reduce the design effect when the elements or clusters within a stratum tend to be homogeneous. For example, if one stratum within a study has a group of households that are all very likely to be in poverty and another has households not likely to be in poverty, the design effect for poverty estimates will be reduced when stratification is taken into account in variance estimation.

Weights have components adjusting for differential probabilities of selection, nonresponse, and sample noncoverage (e.g., when the sample frame does not perfectly cover the population of interest). To the extent that the weights are heterogeneous, the size of the design effect can increase. Weights become heterogeneous in surveys because some elements have higher probabilities of selection than others (by design or by circumstances dictated by the sample frame), because some groups have higher response propensity than others, or because some subgroups are underrepresented by chance relative to known external population distributions (Kish 1992). Kish gives the simple formulation of "$1 + L$" (the "$L$" stands for "Loss" of sample efficiency) to approximate the effect of the sample weights on the design effect. In general, the more heterogeneity in the weights, the higher the design effect will be.

$$1 + L = (n\Sigma k_j^2)/(\Sigma k_j)^2. \tag{2}$$

Here "$n$" is the unweighted sample size, and "$k$" is the survey weight for the "$j$th" person. The numerator of this equation is the unweighted sample size multiplied by the sum of the squared weights. This total then is divided by the sum of the weights squared. The result is an approximation effect on sampling variance due to heterogeneity among weights.

Overall, weighting and clustering tend to increase the design effect and stratification decrease it. In complex sample surveys, however, the impacts of clustering and weighting tend to be larger than those of stratification, so the design effect is greater than 1.

## CPS ASEC Sample Design

Using a stratified multistage sampling design, the CPS draws a representative sample of households in each state and the District of Columbia. In the first stage of sampling, the United States is divided into primary sampling units (PSUs) that can comprise a metropolitan area, a large county, or a group of smaller adjacent counties. All PSUs that correspond to major Metropolitan Statistical Areas (MSAs) are selected into the CPS sample with certainty and are called "self-representing PSUs." The remaining PSUs (non-self-representing PSUs) outside major MSAs but within a state are grouped into strata based on labor force and other social characteristics; at least one PSU is selected from each stratum (U.S. Census Bureau 2002b).

In the second stage of drawing the sample for the CPS, groups of households in close geographical proximity (consisting of approximately four housing units) are selected from the PSUs for inclusion in the survey. Such groups of households are called "ultimate sampling units" (USUs). The CPS draws its sampled housing units from lists of addresses that are continually verified and updated by the Census Bureau for use in its decennial census operations (U.S. Census Bureau 2002b).[3] The final stage of clustering is the household level. Data on everyone within a sampled household is collected.

The probability of selecting a household into the CPS can depend on the household's state of residence, whether the household lives within an MSA, and whether it includes minority group members or children. In addition to varying household selection probabilities, certain types of households and individuals also vary in their likelihood of participating in the survey. In order to control for differential response rates, the CPS weights are adjusted to equal a set of population controls taken from the Census Bureau's annually updated population estimates (U.S. Census Bureau 2002b).

## Methods

For our analyses we use the 2002 Annual Social and Economic Supplement to the CPS.[4] We work with both the public use version of these files and the Census Bureau's internal restricted version of these files. The public use version of these files does not contain the stratifying or clustering information (e.g., USU and PSU identifiers), but the internal Census Bureau restricted versions of these data files do. We derive three estimates for all 50 states and the District of Columbia: 1) the average earned income for people over age 15; 2) the poverty rate; and 3) the health insurance coverage rate. Each of the national and state estimates are computed using four standard error estimation techniques.[5] We apply four different standard error estimators to the public use file and re-estimate one of these four with the sample design data on the internal CPS file for benchmarking purposes.

The four methods of standard error estimation are: 1) the basic "simple random sample" approach, which assumes that every sampled person is drawn independently and completely at random; 2) the Census Bureau's "generalized variance" approach, which is produced and recommended by the Census Bureau (U.S. Census Bureau 2001, 2002a); 3) the "robust variance" estimation approach (also known as the sandwich estimator, or the Huber-White estimator); and 4) a "survey design-based estimator," which uses both an identified stratum and a clustering variable. Standard error estimates computed on the public use file with these four strategies are compared to a survey design-based estimate computed on the internal CPS data using the restricted sampling variables to adjust for the complex sample design.

### *Simple Random Sample*

We use two equations to estimate the "simple random sample" standard errors. Expression 3 is used for rates and expression 4 is used for averages. Assuming that each element was selected as part of a simple random sample should produce the smallest standard errors, on average, because it does not take into account the clustering of people within sampled households, nor the clustering of households within USUs or PSUs. For the most part, we expect that the simple random sample standard errors will be smaller than the standard errors adjusted for the sample design because of the effect of clustering in the CPS sample design. However, this need not be the case. In some instances, it is possible for the design-adjusted standard error to be smaller than the simple random sample standard error (Kish 1995).

For the poverty and health insurance coverage rate estimates, the binomial is used to produce the simple random sample standard error:

$$\sigma_1 = \sqrt{P(100 - P)/n}, \tag{3}$$

where "$P$" is the "weighted" rate of insurance coverage or poverty and the "$n$" is the total number of people included in the sample used to calculate the statistic of interest.

For the continuous income variable, the standard error is computed using the following formula:

$$\sigma_2 = S_x / \sqrt{(n-1)}, \qquad (4)$$

where $S_x$ is the standard deviation of income and "$n$" is the total number of people over age 15 in the state that were included in the sample.[6]

*Generalized Variance Approach*

Currently, the Census Bureau estimates a set of "generalized variance parameters" (GVPs) that its analysts apply to adjust standard errors for the complex sample design used to collect the data (e.g., Mills 2002; DeNavas-Walt and Cleveland 2002; Proctor and Dalakar 2002). The Census Bureau makes these GVPs available to the general public and provides detailed documentation on how to use them (e.g., U.S. Census Bureau 2001, 2002a). This is important because the Census Bureau is unable to release to the general public many of the sample design variables that would be necessary to estimate standard errors because respondent confidentiality could be breached by allowing researchers to identify specific people in the data set. This generalized variance approach refers to the technique the Census Bureau uses to estimate standard errors in its various reports (Mills 2002; DeNavas-Walt and Cleveland 2002; Proctor and Dalaker 2002). It is also the procedure recommended in the Source and Accuracy Statement for the CPS ASEC (U.S. Census Bureau 2001, 2002a).

The generalized variance parameters are estimated by fitting a regression model predicting the variances for specific groups of concepts (e.g., income statistics, health insurance statistics, and poverty statistics). This regression model employs the sample design features that are not released as part of the public use file. Its parameter estimates are published in the Census Bureau's Source and Accuracy Statements and can be used to adjust standard errors computed under the simple random sample approach for the design of the CPS sample (U.S. Census Bureau 2001, 2002b). The CPS ASEC generalized variance parameters also include a correction for the state of residence of the sampled person.

We compute the generalized variance standard errors using the methodology in the Source and Accuracy Statement (U.S. Census Bureau 2001, 2002a) guidelines. Expression 5 is used to compute the standard error for the rate of poverty and health insurance coverage, and 6 is used to compute the standard error for the average income of people over age 15.

$$\sigma_3 = F\sqrt{(B/N)(P(100-P))}, \qquad (5)$$

$$\sigma_4 = F\sqrt{(B/N){S_X}^2}, \qquad (6)$$

Here $F$ denotes the state adjustment ratio used to adjust state estimates; it is available from the CPS ASEC Source and Accuracy Statement (U.S. Census Bureau 2001, 2002a). $B$ is the generalized variance parameter for the specific type of estimate being made. There are separate $B$ parameters for income, health insurance, and poverty estimates which also are available in the Source and Accuracy Statement (U.S. Census Bureau 2001, 2002a).[7] In equation 5, $P$ refers to the rate of poverty or the rate of health insurance coverage. Finally, in both formulas, $N$ is the weighted number of people appropriate to the estimate. For poverty and health insurance estimates, $N$ is the estimated number of noninstitutionalized people within the state. For the income analyses, $N$ is the estimated number of noninstitutionalized people over age 15.

*Robust Variance*

The "robust variance" estimation approach – also known as the sandwich estimator, the Huber-White estimator (SAS 1999), or the "first-order Taylor series linearization" method – is implemented using SAS version 8.2. Specifically, we use the "surveymean" procedure with states designated as subpopulations. In using these survey procedures, we declare only the survey weights among the survey features, which invokes the robust standard error estimator. Although the robust standard error estimator does not explicitly control for any of the clustering features of the CPS survey data per se in generating standard errors, we include it as one of our four standard error estimators. This is because it is common in the research literature to read that standard errors are calculated using STATA (2001), SPSS (2003), or SAS (1999) survey adjustment procedures, but no mention is made of clustering or strata adjustments. If the procedures are used by themselves, without a cluster or strata adjustment, then the robust standard error is the resulting estimator.

*Survey Design-Based Estimator*

The "survey design-based" estimator takes account of the weight, clustering, and stratification of the survey in estimating the standard errors. For our analysis, we use the survey estimator implemented in SAS version 8.2. Like the robust estimation method, the survey design-based method uses a Taylor series estimation approach. But unlike the robust estimation technique, this method explicitly controls for both stratification and clustering. In this study, we use a Taylor series survey design-based estimator to compute the variances identifying the highest (i.e., first) level of clustering (Hansen, Hurwitz, and Madow 1953; Woodruff 1971; Kalton 1977; Rust 1985). Even though this "ultimate cluster" approach to estimating the design effect is based on the sample's first stage of clustering, it does include, in expectation, any subsequent stages of variability as well.[8]

*Ideal specification on the internal file.* The preferred way of computing the variances with this procedure would be to take the actual PSUs to cluster the sampled elements from non-self-representing PSUs and USUs to cluster elements within self-representing PSUs along with the actual strata used. When working with the internal file, we use all self-representing PSUs within a state as unique strata, along with one additional stratum within each state that includes all the non-self-representing PSUs. The first-stage clusters within the self-representing CPS PSUs are the USUs, and the first-stage clusters within the non-self-representing stratum within a state are the CPS-defined PSUs. We implement this method with the internal Census Bureau data and employ it as our standard to which we compare the public use estimates.[9]

*Approximation to the ideal employing the public use file.* Because the PSU and USU variables are not released to the public, it is important to try to construct alternative methods for working with the public use file. Therefore, we also computed a survey design-based estimator on the public use file defining the strata as the lowest level of identifiable geography in the CPS data and our clustering variable as the household in which an individual lives.

We determined the lowest level of identifiable geography through geographic information available on the CPS public use file: 1) the largest 242 primary metropolitan statistical areas (PMSAs) or MSAs; 2) counties with more than 100,000 people in the 1990 census; and 3) states. PMSAs and MSAs are grouped into the same variable on the CPS ASEC public use file. When the two are different, the PMSA takes precedence. To preserve confidentiality, when a PMSA or MSA falls in multiple states, the PMSA/MSA designation may be suppressed in one (or more) of the adjoining states if the population within the PMSA/MSA in that state makes up a relatively small portion of the overall PMSA/MSA.

We impose a mutually exclusive geographic hierarchy on the CPS data that begins with the state of residence. Each sampled household is assigned to one of 51 states (including the District of Columbia). Within each state, if a sampled household resides in an identified PMSA/MSA, it is assigned a specific code (i.e., people living within the same state and PMSA/MSA are grouped together). If the sampled household does not live in an identified PMSA/MSA, but does live in an identified county within the state, it also is assigned a specific geographic code (i.e., all the people living within a state and a highly populated county that is not part of an identified PMSA/MSA). Lastly, everyone else within the state is grouped together with a specific geographic identifier (i.e., all the people living within a state but not living in an identified PMSA/MSA or identified county are grouped together). In two states (Montana and Wyoming), this residual group encompasses everyone within the state, and in three states (Delaware, New Jersey, and the District of Columbia) there is no portion within the residual category. Using this mutually exclusive geographic hierarchy, we were able to identify 311 distinct units of geography within the 2002 CPS ASEC.[10]

Finally, we implement our survey design-based estimator within the SAS statistical analysis procedures by declaring this geographic hierarchy identifier as our strata variable, and the household identifier as our clustering variable to account for the clustering of individuals within sampled households. This strategy does not take into account the higher levels of clustering (PSU and USU clustering) because this information is not available on the public use file of the CPS ASEC.

## Results

As we move from the robust standard error estimates to the survey design-based estimates and hold other things constant, we can make inferences on the effect of clustering at the various levels. The more the survey design-based estimates on the internal file diverge from the robust estimates, the larger the design effect due to clustering at the nonhousehold stage of clustering (PSU or USU). Also, the more the survey design-based estimates on the public use file diverge from the robust estimates, the larger the design effect from clustering at the household level. Finally, the more the survey design-based estimator on the public use file diverges from the survey design-based estimator on the internal files, the greater the impact of higher levels of clustering (USU and PSU clustering) on the design effect. And, as a result, more bias is introduced in standard error estimation by not having access to the clustering and stratification variables on the CPS public use file.

We present the results of our analyses in the following three tables for 2001 health insurance coverage, poverty, and income for the 50 states and District of Columbia.[11] In each table, the second column presents a state's estimated rate of insurance coverage, poverty rate, or mean income, followed in the third column by the estimated standard error calculated using the survey design method on the internal file. We treat this standard error as the base and express the standard errors calculated with the remaining methods as ratios relative to this estimate. A ratio less than 1 indicates that the standard error calculated with a particular method is smaller than the standard error calculated with the survey design-based method on the internal file of the CPS; anything over 1 implies the standard error is larger than the survey design-based estimate on the internal file. Column 4 shows our robust standard errors and column 5 shows our survey design-based estimates on the public use data relative to the internal file estimates. Column 6 shows the estimated standard error ratios for the generalized variance approach, and finally the last column shows the simple random sample (SRS) standard error ratios.

Perhaps the most startling results are the national and state estimates of health insurance coverage using the simple random sample method and the generalized variance estimators (Table 1). On average, the generalized variance estimation technique yields standard errors less than half the size (an average ratio of .44) of the standard errors calculated using the survey design-based estimator on the internal file. These are even smaller, on average, than the simple random sample approach (an average ratio of .53), which does not adjust for the complex survey sample design. As a result of these analyses, the health insurance coverage generalized variance parameter is being reevaluated by the U.S. Census Bureau.[12]

The standard errors for the state estimates of health insurance coverage showed a large effect of clustering both at the household USU and PSU levels. The simple random sample estimates were on average .53 the size of the survey design-based estimates on the internal file, while the robust standard errors were .58 the size of the survey design-based estimates on the internal file. The survey design-based estimates on the public use file were better, but they still were only on average .82 of the survey design-based estimates on the internal file. With health insurance, the survey design-based estimates on the public use file pick up some of the effect of clustering but still underestimate the standard errors when compared to the survey design-based estimates on the internal file as a result of not controlling for clustering at the USU or PSU levels.

Looking at state poverty rates (Table 2), the survey design-based estimates on the public use file were on average .79 the size of the survey design-based estimates on the internal Census Bureau files, indicating a significant impact of clustering at the USU and PSU levels. The robust estimates were only .48 the size of the survey design-based estimates on the internal Census Bureau file, indicating a large impact of clustering at the household level. The simple random sample estimates were .45 the size of the survey design-based estimates on the internal file and show a large design effect. The survey design-based estimates on the public use file pick up some of the overall effects of clustering, but still underestimate the standard errors when compared to the survey design-based estimates on the internal file.

The standard errors for the income estimates (Table 3) show the smallest impact of clustering, as the simple random sample estimates were on average .73 the size of the survey design-based

**Table 1. State health insurance coverage rates and standard error computation comparisons by year: 2001**

| | 2001 health insurance coverage estimate (%) | Survey design-based standard error on internal census file (%) | Ratio of method to survey design-based on internal census file | | | |
|---|---|---|---|---|---|---|
| | | | Robust | Survey design-based on the public use file | Generalized variance estimation | Simple random sample (SRS) |
| United States | 85.4 | .18 | .52 | .77 | .39 | .42 |
| Alabama | 86.9 | 1.20 | .53 | .72 | .43 | .48 |
| Alaska | 84.3 | 1.11 | .62 | .87 | .48 | .57 |
| Arizona | 82.1 | 1.52 | .47 | .79 | .41 | .46 |
| Arkansas | 83.9 | 1.43 | .56 | .78 | .42 | .53 |
| California | 80.5 | .68 | .48 | .77 | .42 | .44 |
| Colorado | 84.4 | 1.10 | .54 | .82 | .44 | .50 |
| Connecticut | 89.8 | .94 | .57 | .81 | .43 | .53 |
| Delaware | 90.8 | 1.11 | .52 | .82 | .40 | .51 |
| Dist. of Columbia | 87.3 | 1.02 | .74 | .92 | .54 | .70 |
| Florida | 82.5 | .87 | .48 | .73 | .39 | .46 |
| Georgia | 83.4 | 1.40 | .52 | .77 | .40 | .48 |
| Hawaii | 90.4 | .97 | .58 | .86 | .46 | .55 |
| Idaho | 84.0 | 1.65 | .44 | .66 | .36 | .42 |
| Illinois | 86.4 | .69 | .61 | .89 | .49 | .56 |
| Indiana | 88.2 | 1.05 | .54 | .80 | .40 | .50 |
| Iowa | 92.5 | .70 | .68 | .96 | .53 | .63 |
| Kansas | 88.6 | 1.05 | .55 | .75 | .43 | .52 |
| Kentucky | 87.7 | 1.30 | .49 | .71 | .38 | .47 |
| Louisiana | 80.7 | 1.61 | .53 | .76 | .39 | .51 |
| Maine | 89.7 | .80 | .74 | 1.00 | .51 | .66 |
| Maryland | 87.7 | 1.20 | .53 | .81 | .37 | .47 |
| Massachusetts | 91.8 | .71 | .68 | .90 | .49 | .61 |
| Michigan | 89.6 | .69 | .64 | .87 | .47 | .57 |
| Minnesota | 92.0 | .78 | .61 | .87 | .47 | .55 |
| Mississippi | 83.6 | 1.15 | .74 | .96 | .54 | .70 |
| Missouri | 89.8 | .77 | .76 | 1.00 | .55 | .69 |
| Montana | 86.4 | 1.41 | .53 | .77 | .41 | .50 |
| Nebraska | 90.5 | .99 | .53 | .72 | .44 | .51 |
| Nevada | 83.9 | 1.21 | .49 | .78 | .42 | .47 |
| New Hampshire | 90.6 | .94 | .59 | .80 | .42 | .52 |
| New Jersey | 86.9 | .92 | .53 | .78 | .41 | .49 |
| New Mexico | 79.3 | 1.45 | .56 | .79 | .50 | .52 |
| New York | 84.5 | .60 | .61 | .85 | .47 | .56 |
| North Carolina | 85.6 | .96 | .58 | .81 | .45 | .55 |
| North Dakota | 90.4 | 1.43 | .42 | .56 | .31 | .38 |
| Ohio | 88.8 | .68 | .64 | .89 | .48 | .57 |
| Oklahoma | 81.7 | 1.21 | .64 | .94 | .48 | .59 |
| Oregon | 87.2 | 1.22 | .50 | .76 | .40 | .48 |
| Pennsylvania | 90.8 | .71 | .53 | .79 | .39 | .48 |
| Rhode Island | 92.3 | .79 | .63 | .90 | .44 | .57 |
| South Carolina | 87.7 | 1.00 | .71 | .95 | .49 | .63 |
| South Dakota | 90.7 | 1.06 | .51 | .71 | .38 | .47 |
| Tennessee | 88.7 | .93 | .73 | .97 | .54 | .66 |
| Texas | 76.5 | .97 | .46 | .71 | .39 | .43 |
| Utah | 85.2 | 1.38 | .47 | .74 | .39 | .46 |
| Vermont | 90.4 | .79 | .81 | .99 | .53 | .66 |
| Virginia | 89.1 | 1.02 | .54 | .76 | .43 | .51 |
| Washington | 86.9 | 1.22 | .48 | .70 | .41 | .46 |
| West Virginia | 86.8 | .93 | .72 | .94 | .51 | .67 |
| Wisconsin | 92.3 | .74 | .62 | .81 | .47 | .56 |
| Wyoming | 84.1 | 1.35 | .55 | .75 | .42 | .49 |
| Average change[a] | | | .58 | .82 | .44 | .53 |

Source: 2002 Current Population Survey Annual Social and Economic Supplement.
[a] Average change is a single average where each state's change is given an equal weight.

**Table 2. State poverty rates and standard error computation comparisons by year: 2001**

| | 2001 poverty rate (%) | Survey design-based standard error on internal census file (%) | Ratio of method to survey design-based on internal census file | | | |
|---|---|---|---|---|---|---|
| | | | Robust | Survey design-based on the public use file | Generalized variance estimation | Simple random sample (SRS) |
| United States | 11.7 | .19 | .44 | .77 | .75 | .37 |
| Alabama | 15.9 | 1.35 | .49 | .83 | .89 | .47 |
| Alaska | 8.5 | 1.10 | .46 | .85 | .81 | .44 |
| Arizona | 14.6 | 1.44 | .46 | .85 | .86 | .44 |
| Arkansas | 17.8 | 1.76 | .46 | .81 | .76 | .45 |
| California | 12.5 | .68 | .41 | .77 | .77 | .37 |
| Colorado | 8.7 | 1.03 | .45 | .74 | .78 | .41 |
| Connecticut | 7.3 | 1.10 | .40 | .63 | .68 | .39 |
| Delaware | 6.7 | 1.43 | .34 | .62 | .59 | .34 |
| Dist. of Columbia | 18.2 | 2.07 | .42 | .69 | .67 | .40 |
| Florida | 12.7 | .75 | .49 | .81 | .86 | .47 |
| Georgia | 12.9 | 1.67 | .39 | .71 | .65 | .37 |
| Hawaii | 11.4 | 1.31 | .46 | .82 | .80 | .44 |
| Idaho | 11.5 | 1.52 | .41 | .71 | .73 | .39 |
| Illinois | 10.1 | .78 | .46 | .81 | .83 | .43 |
| Indiana | 8.5 | .88 | .55 | .90 | .88 | .51 |
| Iowa | 7.4 | .75 | .63 | .94 | 1.06 | .58 |
| Kansas | 10.1 | 1.15 | .47 | .75 | .81 | .45 |
| Kentucky | 12.6 | 1.46 | .44 | .74 | .73 | .42 |
| Louisiana | 16.2 | 2.10 | .37 | .64 | .61 | .36 |
| Maine | 10.3 | .98 | .60 | .89 | .90 | .54 |
| Maryland | 7.2 | .80 | .60 | .92 | .96 | .56 |
| Massachusetts | 8.9 | .91 | .52 | .83 | .85 | .49 |
| Michigan | 9.4 | .80 | .50 | .80 | .84 | .47 |
| Minnesota | 7.3 | .78 | .58 | .86 | .98 | .53 |
| Mississippi | 19.3 | 1.82 | .51 | .87 | .79 | .47 |
| Missouri | 9.7 | 1.36 | .41 | .65 | .66 | .38 |
| Montana | 13.3 | 1.59 | .48 | .76 | .79 | .44 |
| Nebraska | 9.4 | .93 | .56 | .95 | 1.01 | .54 |
| Nevada | 7.1 | .80 | .51 | .89 | .97 | .50 |
| New Hampshire | 6.4 | .78 | .59 | .88 | .92 | .52 |
| New Jersey | 8.1 | .77 | .50 | .79 | .85 | .47 |
| New Mexico | 17.9 | 1.35 | .57 | 1.02 | 1.11 | .53 |
| New York | 14.2 | .77 | .44 | .74 | .76 | .42 |
| North Carolina | 12.5 | 1.05 | .50 | .79 | .84 | .48 |
| North Dakota | 13.8 | 2.30 | .29 | .47 | .48 | .27 |
| Ohio | 10.5 | .89 | .45 | .74 | .78 | .42 |
| Oklahoma | 15.0 | 1.89 | .38 | .63 | .62 | .35 |
| Oregon | 11.8 | 1.31 | .46 | .75 | .79 | .43 |
| Pennsylvania | 9.6 | .76 | .48 | .82 | .81 | .45 |
| Rhode Island | 9.6 | 1.19 | .46 | .71 | .70 | .42 |
| South Carolina | 15.0 | 1.81 | .42 | .69 | .64 | .38 |
| South Dakota | 8.4 | 1.05 | .50 | .74 | .80 | .45 |
| Tennessee | 14.1 | 1.30 | .57 | .93 | .93 | .52 |
| Texas | 14.9 | .87 | .42 | .79 | .79 | .41 |
| Utah | 10.5 | 1.48 | .38 | .69 | .68 | .37 |
| Vermont | 9.7 | .96 | .61 | .92 | .95 | .54 |
| Virginia | 7.9 | 1.02 | .47 | .73 | .81 | .44 |
| Washington | 10.7 | 1.41 | .38 | .63 | .71 | .36 |
| West Virginia | 16.4 | 1.38 | .53 | .85 | .82 | .49 |
| Wisconsin | 7.9 | .88 | .49 | .83 | .88 | .47 |
| Wyoming | 8.6 | .89 | .62 | .99 | 1.05 | .57 |
| Average change[a] | | | .48 | .79 | .81 | .45 |

Source: 2002 Current Population Survey Annual Social and Economic Supplement.
[a] Average change is a single average where each state's change is given an equal weight.

## Table 3. Average state individual income and standard error computation comparisons by year: 2001

| | 2001 average income ($) | Survey design-based standard error on internal census file ($) | Ratio of method to survey design-based on internal census file | | | |
|---|---|---|---|---|---|---|
| | | | Robust | Survey design-based on the public use file | Generalized variance estimation | Simple random sample (SRS) |
| United States | 29,089 | 192 | .63 | .77 | .70 | .62 |
| Alabama | 24,952 | 1,128 | .85 | .85 | .79 | .75 |
| Alaska | 32,549 | 1,146 | .72 | .76 | .25 | .64 |
| Arizona | 27,659 | 1,271 | .72 | .72 | .87 | .70 |
| Arkansas | 22,448 | 848 | .86 | .86 | .74 | .83 |
| California | 30,173 | 655 | .74 | .79 | 1.05 | .74 |
| Colorado | 33,171 | 1,356 | .62 | .65 | .58 | .62 |
| Connecticut | 36,654 | 1,558 | .78 | .78 | .60 | .74 |
| Delaware | 31,758 | 1,338 | .81 | .81 | .34 | .74 |
| Dist. of Columbia | 38,417 | 3,026 | .63 | .65 | .25 | .59 |
| Florida | 27,005 | 755 | .81 | .82 | .92 | .76 |
| Georgia | 26,657 | 1,007 | .79 | .78 | 1.10 | .77 |
| Hawaii | 26,607 | 978 | .73 | .77 | .41 | .73 |
| Idaho | 24,772 | 898 | .83 | .85 | .49 | .79 |
| Illinois | 29,203 | 688 | .78 | .79 | .86 | .74 |
| Indiana | 28,398 | 915 | .84 | .84 | .91 | .86 |
| Iowa | 26,788 | 1,048 | .85 | .85 | .61 | .77 |
| Kansas | 29,147 | 1,232 | .78 | .78 | .59 | .76 |
| Kentucky | 25,639 | 1,005 | .73 | .71 | .68 | .68 |
| Louisiana | 24,617 | 1,999 | .70 | .68 | .68 | .61 |
| Maine | 25,723 | 702 | .97 | .96 | .51 | 1.01 |
| Maryland | 37,841 | 1,752 | .79 | .79 | .84 | .79 |
| Massachusetts | 35,840 | 1,505 | .76 | .78 | .77 | .73 |
| Michigan | 30,288 | 1,117 | .80 | .80 | .77 | .69 |
| Minnesota | 32,965 | 1,298 | .73 | .75 | .66 | .66 |
| Mississippi | 23,617 | 1,979 | .94 | .93 | .62 | .66 |
| Missouri | 27,780 | 1,298 | .64 | .74 | .69 | .63 |
| Montana | 21,894 | 787 | .81 | .85 | .40 | .74 |
| Nebraska | 26,935 | 836 | .73 | .72 | .48 | .74 |
| Nevada | 29,880 | 1,398 | .85 | .86 | .50 | .73 |
| New Hampshire | 34,495 | 1,390 | .78 | .76 | .39 | .78 |
| New Jersey | 35,266 | 1,311 | .84 | .86 | .80 | .75 |
| New Mexico | 22,590 | 885 | .82 | .84 | .58 | .73 |
| New York | 30,778 | 887 | .83 | .90 | .86 | .76 |
| North Carolina | 25,695 | 963 | .66 | .66 | .77 | .65 |
| North Dakota | 24,059 | 956 | .65 | .68 | .24 | .62 |
| Ohio | 28,328 | 797 | .78 | .78 | .87 | .75 |
| Oklahoma | 23,850 | 961 | .76 | .75 | .72 | .77 |
| Oregon | 27,866 | 1,236 | .76 | .76 | .65 | .71 |
| Pennsylvania | 29,319 | 772 | .74 | .75 | .82 | .74 |
| Rhode Island | 31,266 | 1,145 | .66 | .70 | .30 | .66 |
| South Carolina | 24,242 | 822 | .89 | .87 | .90 | .91 |
| South Dakota | 24,824 | 1,004 | .81 | .80 | .29 | .71 |
| Tennessee | 25,944 | 1,691 | .60 | .60 | .80 | .63 |
| Texas | 28,325 | 1,036 | .65 | .68 | .76 | .56 |
| Utah | 25,123 | 1,027 | .70 | .68 | .54 | .71 |
| Vermont | 27,128 | 801 | .76 | .79 | .28 | .76 |
| Virginia | 31,989 | 1,161 | .79 | .80 | .92 | .73 |
| Washington | 29,413 | 1,127 | .63 | .65 | .73 | .60 |
| West Virginia | 21,601 | 917 | .77 | .76 | .49 | .78 |
| Wisconsin | 29,954 | 1,020 | .73 | .72 | .68 | .67 |
| Wyoming | 23,977 | 684 | .87 | .84 | .30 | .83 |
| Average change[a] | | | .77 | .78 | .64 | .73 |

Source: 2002 Current Population Survey Annual Social and Economic Supplement.
[a] Average change is a single average where each state's change is given an equal weight.

estimates on the internal file. In addition, the income estimates show little impact of the effect of clustering at the household level, as the survey design-based estimates on the public use file and the robust standard errors produced, on average, standard errors that were .78 and .77 the size of the survey design-based estimates on the internal Census Bureau file. Thus the survey design based-estimates on the public use file and the robust estimation method would both underestimate the standard errors relative to the survey design-based estimates on the internal file. There is little difference between the two because there is not a large effect of clustering at the household level. Furthermore, the generalized variance parameters produced on average smaller standard errors than the others – estimates that were .64 the size of the survey design-based estimates on the public use file. This is largely due to adjustments from the "*F*" state factor in the generalized variance parameters. For the United States as a whole, the generalized variance parameter approach is slightly larger (as it should be) than the simple random approach. However, due to state adjustments ("*F*"s) that can be as low as .32 in Wyoming, the generalized variance parameters on average are actually smaller than the simple random simple estimates.

## Discussion

The Census Bureau's generalized variance parameter approach to standard error calculation performed erratically. From our analyses, we conclude that the poverty, health insurance, and income generalized variance estimates were too small and some of the state adjustment factors ("F's") were too extreme for variables that had smaller design effects overall (such as income). The extreme cases are shown with the health insurance estimates and income estimates where some of the generalized variance parameter standard error estimates were actually smaller, on average, than the simple random sample estimates of the standard errors. For the health insurance estimates, this was true for all of the estimates because the generalized variance parameter for health insurance coverage was too small given the large impact of sample design shown by the survey design-based estimates on the internal file. For the income estimates, this was true mainly for those states that had small adjustment factor "F's" only.

The generalized variance parameters themselves came about before statistical software incorporated survey adjustment procedures. The generalized variance parameter approach was developed as part of a concerted effort to get public data users to realize that simple random sample standard errors were not appropriate for Census Bureau data products. Although statisticians have been working to develop software for estimating reliable variance estimates from complex surveys since the 1970s, these tools were not practical for most researchers until the 1990s. The first software products suitable for analyses of complex samples were stand-alone applications such as SUDAAN and WESVAR (Brick and Morganstein 1996; Brogan 1998; Lepkowski and Bowles 1996). In the past few years, however, similar procedures have been incorporated into the three most widely used general-purpose statistical packages: SAS, Stata, and SPSS (SAS 1999; Stata 2001; SPSS 2003). With this development, it is reasonable to expect that users of public use data will employ the variance estimation procedures that are more appropriate to complex sample designs; the generalized variance parameters ought to be retired thanks to these technological advances.

Because the robust variance estimation does not take into account the intraclass correlation ($\rho$) and cluster size (i.e., clustering), it does a poor job of estimating the standard errors for the two variables examined (health insurance coverage and poverty). The largest differences between the survey design-based estimates on the internal file and the robust estimates were in the poverty estimates, and the smallest differences were in the income estimates – again, tracking the fact that the robust standard error calculations did not take into account the clustering. Therefore, researchers working with the public use data files need to include more than just the "survey" procedures in the statistical packages such as SAS, STATA and SPSS. Researchers should be explicit about how they define the strata and clusters in writing up their work since invoking the procedures for weights alone is not enough.

The survey design-based estimates employing the public use data do not reflect the large impact of clustering at the USU and PSU levels. Most of the between-cluster variation occurs within the smaller (and largely rural) PSUs within each state. These smaller PSUs represent a large share of the CPS ASEC design effect, and they are not identifiable on the public use file. This is not reassuring information for people working with the CPS ASEC public use data files, as appropriate standard error estimates for measures of central tendency, such as rates and means, are likely to be, at best, moderately biased downward.[13]

The current confusion around how to calculate appropriate standard errors has implications for policy research. In particular, the Census Bureau estimates of the number of uninsured children living in families with incomes below 200% of the federal poverty level (FPL) in each state, and the number of children in families below 200% of the FPL within each state are critical components of the formula used to allocate federal dollars for SCHIP. The SCHIP allocation formula has been shown to be rather unstable over time due to sampling error (Davern, Blewett et al. 2003). This demonstrated the lack of year-to-year stability in the formula using the Census Bureau's standard error estimates that were calculated using the generalized variance estimation approach. As we have demonstrated in this paper, these generalized variance estimates from the Census Bureau are likely an underestimation of the actual standard errors. This implies that the instability in the formula noted by Davern, Blewett et al. (2003), who used the generalized variance estimates, is worse than they calculated. This makes the alternatives proposed in the paper even more important for policymakers to consider.

In addition to funding formulas, the CPS ASEC is used to evaluate important health policies. Many states use the CPS estimates of health insurance coverage to monitor their SCHIP programs (Health and Human Services Inspector General 2004). The sample size of the CPS ASEC was expanded specifically to improve the survey for this purpose (Davern, Beebe et al. 2003). However, states that have documented a statistically significant decrease in the number of uninsured children using the CPS ASEC data (Health and Human Services Inspector General 2004) may not actually be showing a significant decrease if the standard errors were calculated correctly. This is because the standard errors the states are using to determine statistical significance are likely underestimated to the extent they rely on Census Bureau tabulations that use the generalized variance estimation approach. In addition, many academics who use the CPS ASEC data may be underestimating standard errors in the health policy research work that evaluates the effects of public program expansions during the 1990s. As a result, these analysts may find significant levels of crowd-out, a drop in the uninsurance rate, or an increase in the public program participation rate when there may not actually be statistical significance of such occurrences. Because of the crucial role the CPS ASEC plays in health policy (both in allocating funds and evaluating the success and failure of various efforts to cover the uninsured), efforts should be made to improve the estimating capability of analysts using this data for health policy purposes.

## Recommendations

We have two sets of recommendations after completing our analysis. The first is for the U.S. Census Bureau, and the second is for health services researchers using the CPS ASEC data.

### *Recommendations for the U.S. Census Bureau*

The U.S. Census Bureau should begin using alternative standard error estimation procedures in its official reports. The generalized variance parameters performed poorly in the evaluation we presented in this paper, especially for health insurance coverage. These are still the procedures used by the Census Bureau in its official poverty income and health insurance reports (Mills 2002; DeNavas-Walt and Cleveland 2002; Proctor and Dalaker 2002) as well as most other reports deriving from the CPS and the Survey of Income and Program Participation (SIPP). Given the ease of use and availability of survey design-based estimators in the major statistical packages (SAS 1999; STATA 2001; and SPSS 2003), the Census Bureau should set standards for its analysts and train them in these procedures. This would greatly improve the standard error and confidence interval estimates that are currently released in the official reports.

Our second recommendation for the Census Bureau is to produce better information, variables, and/or direction for working with the public use data files to produce appropriate standard errors. The current effort put into developing the generalized variance parameters could be allocated to producing a set of replicate weights for the ASEC, as is done for other CPS files, SIPP files, and the decennial census public use microdata sample (PUMS) files. Replicate weights would allow the Census Bureau to maintain respondent confidentiality while allowing public data users to construct appropriate standard errors from the file. The drawback to producing a set of such weights is that the process will be complicated by the recent CPS ASEC sample expansion (Davern, Beebe et al. 2003). Another possible consideration would be for the Census Bureau to construct "pseudo-PSUs" that could be used in the Taylor series survey design-based estimates employing the major statistical packages (SAS 1999; STATA 2001; SPSS 2003). This procedure is used on the National Health Interview Survey and health services researchers are able to come up with reasonable standard errors applying it in combination with the statistical package software (National Center for Health Statistics 2000). In the current climate of reducing the amount of geographic information released with public use files (due to confidentiality and disclosure concerns), the Census Bureau will not reverse its decision *not* to release the actual PSU and USU variables. However, the pseudo-PSU variable has worked well for other federal government surveys and should be considered for the CPS as well.

### *Recommendations for Health Services Researchers*

Our recommendation for what CPS public data users should *not* do is simple. First, the simple random sample approach is not appropriate and grossly underestimates the standard errors. Second, robust estimation (i.e., first-order Taylor series expansion with the CPS person weights only) does not perform well for health insurance coverage and poverty estimates. Therefore, it is not appropriate to simply use the survey procedures in statistical packages without including appropriate information regarding stratification and clustering. Third, the generalized variance parameters produced by the Census Bureau perform erratically and should not be used.

As for what analysts should do, our current recommendation is to use the survey design-based methodology applying the geographical strata and household cluster to provide an adequate, albeit downward biased, standard error estimate in cases where the intra-household correlation is high (e.g., poverty and health insurance coverage). The survey design-based estimates on the public use file can be implemented easily in popular statistical packages such as SAS, SPSS and STATA. However, in cases where the intra-household correlation is lower (e.g., income), the survey design-based estimates on the public use file do not perform any better than the robust standard errors.

This leads to our final recommendation for researchers working with the CPS public use file. They should continue to explore ways to calculate appropriate standard errors not only for the measure of central tendency presented in this paper, but also for multivariate models. This work will be critical for developing more stable inferences based on reasonable standard errors. We believe this work will help form a solid foundation upon which stable inferences can be made to evaluate health policy, estimate the various costs of health policy proposals, and appropriately fund current health care programs.

## Notes

*This paper reports the results of research and analysis undertaken by U.S. Census Bureau staff. It has undergone a more limited review than official Census Bureau publications. This report is released to inform interested parties of research and encourage discussion.*

*The authors wish to thank Daniel Weinberg, Pete Fronczek, Carole Popoff, Robin Fisher, Jeff Stratton, and Thomas Moore of the U.S. Census Bureau for their excellent reviews of this paper. We also would like to thank Katherine Swartz and two anonymous reviewers for guidance in completing this paper. Karen Soderberg at the State Health Access Data Assistance Center (SHADAC) also provided outstanding editorial help. An earlier version of this paper was presented at the American Association for Public Opinion Research meetings in Nashville, Tenn., in May 2003, and also at the Joint Statistical Meetings in San Francisco, Calif., in August 2003.*

1 A data user can apply to a Census Research Data Center (RDC) to gain access to the confidential data. A proposal must be submitted to the Census Bureau's Center for Economic Studies and the proposed access needs to be justified, including how it would benefit the Census Bureau. If the proposal is accepted, the researcher may go to

any of six RDCs to perform his/her analysis. All analyses performed are reviewed by Census Bureau staff before they are allowed to be taken from the center. The costs, travel, and time necessary to do analysis at a remote location makes this an unattractive option for most researchers.

2 This is not always the case, however. Stratification can produce significant sampling efficiencies that can overcome clustering, as seen in a few instances in the estimates. For example, Maine in the income table (Table 3) has a simple random sample standard error that is 1.01, the size of the survey design-based standard error on the internal census file.

3 A third stage of sampling is occasionally used when the USU is large.

4 We also have performed this analysis on the 2001 CPS ASEC but for presentation purposes have cut these out. They are available from the authors upon request.

5 The District of Columbia is considered a "state" for ease of presentation purposes throughout this paper.

6 Income in the CPS public use file is "top coded" and "bottom coded." If an individual's income exceeds the top coded value, it is censored and given the top code income value (same for the negative income amounts and the bottom codes). This would reduce significantly the variance and resulting standard errors of mean income. Therefore, we used the internal file income values in our "public use" analyses.

7 The 2002 generalized variance parameters (U.S. Census Bureau 2002a) were used. For health insurance coverage, the $B = 1{,}115$; for poverty, the $B = 5{,}282$; and for income the $\mathrm{B} = 1{,}249$. The state values for "$F$" are available from the U.S. Census Bureau (2002a).

8 For PSUs within a stratum that vary substantially in size (in terms of numbers of completed interviews per PSU), the first-stage sample variance component, in expectation, does include the second-stage variance (Rust 1985). Examining whether this is actually the case is beyond the scope of the current paper, the goal of which is to develop comparisons of readily accessible estimation procedures. Looking at second-stage variances requires resampling or use of sophisticated software as well as access to internal PSU and USU identifiers, none of which are easily accessible to typical CPS data users.

9 We focus our analysis on the Taylor series linearization approach and do not estimate replicate methods for two reasons. First, in most cases, the Taylor series linearization approach yields results that are comparable to replicate methods (Kish and Frankel 1974; Krewski and Rao 1981; Dippo and Wolter 1984; Weng, Zhang, and Cohen 1995; Hammer, Shin, and Porcellini 2003). Second, the Taylor series linearization is more computationally efficient than replication methods, and the procedure is available in the major statistical packages SAS, Stata, and SPSS (SAS 1999; STATA 2001; SPSS 2003).

10 We use the household identification number (H_SEQ) for the clustering variable. For the strata variable, we first identify the self-representing PSUs (HG_MSAC); then we add non-self-representing PSUs that are identified counties (GECO); then put the remaining cases into the balance or rural part of the state using the state variable (GESTFIPS). Example Stata code would be:

```
gen strataid=.
replace strataid =(gestfips*10000)+ hg_msac if hg_msac> 0
replace strataid =(gestfips*1000) + geco if geco> 0 & hg_msac==0
replace strataid = gestfips if geco == 0 & hg_msac == 0
```

11 The data contained in the CPS ASEC are retrospective for the prior calendar year. Therefore, data collected on the 2002 CPS ASEC and presented in the text and tables are for 2001. Likewise, data collected in 2001 CPS ASEC are for the calendar year 2000.

12 After being presented with these findings, the Census Bureau returned to using the 1997 value (NEW B = 2,652), and is reviewing the methodology used to recalculate the health insurance parameters. With the new value the standard errors are now 55% larger than the generalized variance estimates in Table 1.

13 Kish (1995) addressed the issue of expanding the idea of the survey design effect to non-parametric measures of central tendency such as medians (pp. 495–496) and ranges. However, this work has not been incorporated into statistical packages like SAS, Stata and SUDAAN. Future work should attempt to develop a standard way of estimating the impact of the design effect on these types of important statistics as they are widely produced in U.S. Census Bureau official reports.

## References

Blewett, L.A., M.E. Brown, K.T. Call, and M. Davern. 2004. Monitoring the Uninsured: A State Policy Perspective. *Journal of Health Politics, Policy and Law* 29(1):107–145.

Brick, J.M., and D. Morganstein. 1996. WesVarPC: Software for Computing Variance Estimates from Complex Designs. *Proceedings of the 1996 Annual Research Conference.* Washington, D.C.: U.S. Bureau of the Census.

Brogan, D. 1998. *Software for Sample Survey Data, Misuse of Standard Packages.* In *Encyclopedia of*

*Biostatistics*, vol. 5, P. Armitage and T. Colton, eds. New York: Wiley.

Davern, M., L.A. Blewett, B. Bershadsky, K.T. Call, and T. Rockwood. 2003. State Variation in SCHIP Allocations: How Much Is There, What Are Its Sources, and Can It Be Reduced? *Inquiry* 40(2):184–197.

Davern, M., T. Beebe, L.A. Blewett, and K.T. Call. 2003. Recent Changes to the Current Population Survey: Sample Expansion, Health Insurance Verification and State Health Insurance Coverage Estimates. *Public Opinion Quarterly* 67(4): 603–626.

DeNavas-Walt, C., and R. Cleveland. 2002. Money Income in the United States: 2001. Washington, D.C.: U.S. Census Bureau.

Dippo, C.S., and K.M. Wolter. 1984. A Comparison of Variance Estimators Using the Taylor Series Approximation. *ASA Proceedings of the Section on Survey Research Methods.* Arlington, Va.: American Statistical Association.

Glied, S., D.K. Remler, and J.G. Zivin. 2002. Inside the Sausage Factory: Improving Estimates of the Effects of Health Insurance Expansion Proposals. *Milbank Quarterly* 80(4):603–636.

Hammer, H., H. Shin, and L. Porcellini. 2003. A Comparison of Taylor Series and JK1 Resampling Methods for Variance Estimation. Hawaii International Conference on Statistics and Related Fields. Honolulu, Hawaii, June 5–8.

Hansen, M.H., W. Hurwitz, and W. Madow. 1953. *Sample Survey Methods and Theory.* New York: Wiley and Sons.

Health and Human Services Inspector General. 2004. SCHIP: States' Progress in Reducing the Number of Uninsured Kids. Washington, D.C: Office of the Inspector General, Health and Human Services. Available at: http://www.oig.hhs.gov/oei/reports/oei-05-03-00280.pdf

Kalton, G. 1977. Practical Methods for Estimating Survey Sampling Errors. *Bulletin of the International Statistical Institute* 47(3):495–514.

Kish, L. 1992. Weighting for Unequal $P_i$. *Journal of Official Statistics* 8(2):183–200.

———. 1995. *Survey Sampling*, Wiley Classics Library Edition. New York: Wiley and Sons.

Kish, L., and M.R. Frankel. 1974. Inference from Complex Samples. *Journal of the Royal Statistical Society* B(36):1–37.

Krewski, D., and J.N.K. Rao. 1981. Inference from Stratified Samples: Properties of Linearization, Jackknife and Balanced Repeated Replication Methods. *Annals of Statistics* 9:1010–1019.

Lepkowski, J.M., and J. Bowles, 1996. Sampling Error Software for Personal Computers. *Survey Statistician* 35:10–17.

Lohr, S. 2000. *Sampling: Design and Analysis.* Pacific Grove, Calif.: Duxbury Press.

Mills, R. 2002. Health Insurance Coverage in the United States for 2001. Washington, D.C.: U.S. Census Bureau.

National Center for Health Statistics. 2000. Design and Estimation for the National Health Interview Survey, 1995-2004. *Vital and Health Statistics Series* 2(130):1–41.

Proctor, B., and J. Dalaker. 2002. Poverty in the United States: 2001. Washington, D.C.: U.S. Census Bureau.

Rust, K. 1985. Variance Estimation for Complex Estimators in Sample Surveys. *Journal of Official Statistics* 1(4):381–397.

SAS. 1999. *Documentation for SAS Version 8.* Cary, N.C.: SAS Institute, Inc.

SPSS. 2003. *Correctly and Easily Compute Statistics for Complex Sampling.* Chicago, Ill.: SPSS Inc. Available at: http://www.spss.com/complex_samples/

STATA. 2001. *Reference Manual.* College Station, Tex.: STATA Press.

U.S. Census Bureau. 2001. Source and Accuracy of the Data for the March 2001 Current Population Survey Microdata File. Washington, D.C.: U.S. Census Bureau.

———. 2002a. Source and Accuracy of the Data for the March 2002 Current Population Survey Microdata File. Washington, D.C.: U.S. Census Bureau.

———. 2002b. Current Population Survey: Design and Methodology. Technical Paper #63RV. Washington, D.C.: U.S. Census Bureau.

Weng, S.S., F. Zhang, and M.P. Cohen. 1995. Variance Estimates Comparison by Statistical Software. *ASA Proceedings of the Section on Survey Research Methods.* Arlington, Va.: American Statistical Association.

Woodruff, R.S. 1971. A Simple Method for Approximating the Variance of a Complicated Estimate. *Journal of the American Statistical Association* 66(334):411–414.