Variance Estimation for Domains in the U.S. Current Employment Statistics Program

August 2005

Julie B. Gershunskaya, U.S. Bureau of Labor Statistics, Gershunskaya.Julie@bls.gov
P. Lahiri, University of Maryland, College Park, and ASA/NSF/BLS Research Fellow Program

Abstract:
The U.S. Bureau of Labor Statistics (BLS) publishes monthly estimates of employment levels, one of the key indicators of the U.S. economy, for many domains. To assess the quality of these estimates, it is important to publish their associated standard error estimates. In our simulation study, the standard design-based variance estimators of the monthly employment growth rate estimators are found to be often unstable even at a statewide industrial level where there is a sample capable of producing a good point estimates. In this paper, we develop new direct design-based,, synthetic model-based and empirical linear Bayes (ELB) variance estimators Using a Monte Carlo simulation from a real finite population, we evaluate the bias, variance, mean squared error (MSE), and coverage properties of the proposed variance estimators with respect to the randomization principle.

Keywords: composite estimation, mean squared error, Monte Carlo simulation, randomization principle.

1. Introduction

The U.S. Bureau of Labor Statistics' Current Employment Statistics (CES) program, a federal-state cooperative program, collects data each month on employment, hours, and earnings from a sample of establishments. Using the CES sample, the Bureau of Labor Statistics (BLS) publishes estimates of these economic indicators by industry at the national, state, and metropolitan statistical area (MSA) levels. See the *BLS Handbook of Methods* (2004, Chapter 2) for further details.

The BLS maintains the Longitudinal Data Base (LDB) which uses information from the Quarterly Census of Employment and Wages program. Among other items, the LDB contains monthly employment data for every U.S. business establishment covered by the Unemployment Insurance (UI) tax laws, which is virtually a census. The LDB is updated quarterly, on a lagged basis, approximately 6 to 9 months after the reference period. For the CES survey, it provides a sampling frame and the benchmark data. Information from the LDB can be used for research purposes. See Section 5 for a description of Monte-Carlo simulations based on the data from the LDB.

The CES survey uses a stratified simple random sample of the UI accounts which are clusters of establishments. The state, industrial supersector (based on a North American Industry Classification System, NAICS), and employment size class form the strata. Approximately one-third of all non-farm payroll workers are covered by the active CES sample. Optimal allocation is used to minimize the sampling variance of the statewide total private employment level. For further details, see Butani *et al.* (1997) and Werking (1997).

The CES uses a weighted link relative (WLR) estimator to produce monthly employment estimates. The WLR estimator ($\hat{Y}_t$) of the employment for month t (Y_t) is given by $\hat{Y}_t = Y_0 \hat{R}_1 \ldots \hat{R}_t$, where Y_0 is a known population employment (benchmark employment) in a certain month in the past, and $\hat{R}_t$ is an estimator of R_t, the employment growth rate for month t. The employment growth rate is estimated by, $\hat{R}_t = \sum w_j y_{j,t} \big/ \sum w_j y_{j,t-1}$, where y_{jt} and w_j are the employment and sampling weight for the jth establishment for month t, the sum being taken over all establishments reporting nonzero employment for both months t and t-1 in the sample belonging to the population of interest. This is essentially a standard ratio estimator and hence is approximately unbiased under the randomization principle (see Cochran 1977). For further details on the CES survey methods, see *Current Employment Statistics Manual* (2001, Chapter 7), Butani *et al.* (1997), and Harter *et al.* (2003).

Variance estimation for the CES program was discussed in Wolter *et al.* (1998). The BLS uses a balanced half sample replication (BHS) method for variance estimation. Such BHS variance estimation performs well at the national level. However, BHS variances are very unstable at the state by industrial supersectors or lower levels. For these levels the repeatedly grouped BHS (RGBHS) (Rao and Shao, 1996) method is used. In this paper, we evaluate the randomization properties of the ratio components $\hat{R}_t$ of the WLR estimator and their associated variance estimators at the state level by industrial supersectors using a design-based Monte Carlo simulation study. The ratios appear to perform well in terms of both design-bias and design-variance. The BHS and RGBHS variance estimators have good bias properties. However, they are very unstable in terms of their design-based variances.

In Section 2, we consider a delta method for variance estimation of the monthly employment growth rates. The formula is not readily available from a textbook in sampling since we deal with a nonstandard sampling design resulted from the fact that at the time of sampling only the cluster membership and not the industry membership of an establishment is considered. The delta method has computational advantages over the competing BHS or RGBHS replication methods. Like the BHS and RGBHS variance estimators, the new variance estimator is approximately design-unbiased but unstable. However, our simulation suggests that the delta method is less unstable than both the BHS and RGBHS methods. Similar comparison was made earlier by Krewski and Rao (1981). In Section 3, we consider a robust model-based synthetic variance estimator. A robust linear empirical Bayes(ELB), a compromise between the direct and synthetic methods, is considered in Section 4. In the process of deriving the ELB, we obtain an estimator of the variance of our delta variance estimator, a factor needed in computing the ELB. This is again not available from a standard text. A Monte Carlo simulation study is undertaken in Section 5. In terms of the design-bias property, as expected, the direct variance estimators perform better than either the synthetic or the ELB variance estimators. On the other hand, overall the design-variance of the synthetic method is smaller than that of the direct or the ELB method. The design-based mean squared error (MSE) criterion is a convenient way to compare a design-biased estimator with a design-unbiased estimator. The ELB turns out to be the best in terms of the design based MSE criterion.

2. A Delta Method

In this section, we use a delta method to develop a new variance formula that captures the CES sampling design. To this end, we introduce some notation. We focus on the estimation for the industries within a state. Let the subscript *i* denote the industry under study. The CES has a stratified cluster design. Within each state, a UI account, a collection of establishments, represents a cluster. Most establishments in a UI account belong to a certain, dominant, industry, which is used to define a UI selection industry stratum. The UI accounts are stratified by this dominant industry and the employment size class. In general, establishments from other industries may be a part of a given UI account. Thus, after the sample is drawn, for the purpose of estimation related to an industry, establishments that belong to other industries are removed and establishments that belong to the industry under study are included in the estimation regardless of the industry from which they were selected as a part of a UI cluster.

Let y_{hlkt} and w_{hlk} be the month *t* employment and the associated sampling weight for the *k*th establishment of the *l*th UI account belonging to the *h*th stratum.. Note that the sampling weight does not change over time. Let *a* denote an independent sample from which a given establishment is selected as part of a UI cluster. The ratio estimator of the monthly employment growth is given by

$$\hat{R}_{it} = \frac{\sum_{s_{it}} w_{hlk} y_{hlkt}}{\sum_{s_{it}} w_{hlk} y_{hlkt-1}} = \frac{\sum_{a=1}^{A} \hat{\eta}_{y;at}}{\sum_{a=1}^{A} \hat{\eta}_{y;at-1}},$$

where $\hat{\eta}_{y;at}$ is the estimator of the employment total $\eta_{y;at}$ from *a*th independent sample and s_{it} is a set of establishments that report positive employment in both months *t*-1 and *t*.

It is possible that we do not have establishments belonging to the industry of interest *i* for every stratum from each independent sample. In calculating the cluster total, we simply add up the employment of all the establishments of the industry *i* (the theory is valid by treating employment as zero for other industries).

Throughout the paper, we use E_d and V_d to denote the design-based expectation and variance. We use E_m and V_m to denote the expectation and variance with respect to an assumed model. An estimator of $V_d(\hat{R}_{it})$, the design-based variance of $\hat{R}_{it}$, is given by

$$\hat{V}_{it} \approx \frac{1}{\hat{\eta}_{t-1}^2} \sum_{a=1}^{A} \sum_{h=1}^{H_a} N_{ha}^2 (1 - f_{ha}) \frac{s_{e;hat}^2}{n_{ha}},$$

where

$$s_{e;hat}^2 = \frac{\sum_{l \in s_{ha}} (\eta_{e;hlat} - \bar{\eta}_{e;hat})^2}{n_{ha} - 1},$$

$$\bar{\eta}_{e;hat} = \frac{1}{n_{ha}} \sum_{l \in s_{ha}} \eta_{e;hlat}, \quad \text{and}$$

$$\eta_{e;hlat} = \eta_{y;hlat} - \hat{R}_{it} \eta_{y;hlat-1},$$

$\eta_{y;hlat}$ is the total of the variable *y* at time *t* for the *l*th cluster of the *h*th stratum selected in s_{ha}, the *h*th stratum from the *a*th sample; N_{ha}, n_{ha}, f_{ha} are the population size, sample size and the sampling fraction for the *h*th stratum from *a*th sample respectively.

As expected, in our simulation $\hat{V}_{it}$ performed well in terms of the design-bias. However, $\hat{V}_{it}$ is very unstable in terms of the design-variance. This may be due to the small sample used to estimate the finite population variance for an industry. It is interesting to note that Cho et al. (2002) and Huff et al. (2002) noticed similar instability of the BHS variance estimator based on exploratory data analyses and not on a design-based simulation.

3. A Synthetic Method

In this section, we attempt to obtain a variance estimator that is expected to perform well in terms of its design-variance. We need a suitable model to compensate for the small sample problem mentioned in Section 3. It is well-known that certain salient design features can be captured by using an appropriate model. In fact, for many commonly used sampling designs, standard design-based variance estimators can be produced from a model-based approach (see, e.g. Lohr 1999). But, a model-based approach is flexible enough to produce alternate variance estimators that may be desired in an attempt to rectify some deficiencies of the design-based estimator.

The above discussions motivate us to obtain an estimator of $V_d(\hat{R}_{it})$ by estimating $V_m(\hat{R}_{it})$, the variance of $\hat{R}_{it}$ with respect to a working model. We call this a synthetic variance estimator since this approach implicitly assumes that $V_d(\hat{R}_{it}) = V_m(\hat{R}_{it})$. This is an assumption that is likely to fail.

Let $\mathbf{y}_{it-1}$, $\mathbf{y}_{it}$ denote vectors of sample observations for industry *i* at months *t-1* and *t*, respectively. We have explored several models for $\mathbf{y}_{it}$ given $\mathbf{y}_{it-1}$ and finally decide to choose the following model. We recognize that this model could be improved and so we are calling it a working model.

Model 1:

Conditional on $\mathbf{y}_{it-1}$, observations in $\mathbf{y}_{it}$ are uncorrelated with

$$E_m[y_{hlkt} \mid \mathbf{y}_{it-1}] = R_{it} y_{hlkt-1},$$

$$V_m[y_{hlkt} \mid \mathbf{y}_{it-1}] = \sigma_t^2 y_{hlkt-1}.$$

Using the variance decomposition formula,

$$V_m\left(\hat{R}_{it}\right) = E_m V_m\left(\hat{R}_{it} \mid \mathbf{y}_{it-1}\right) + V_m E_m\left(\hat{R}_{it} \mid \mathbf{y}_{it-1}\right),$$

we now derive the variance of $\hat{R}_{it}$ under Model 1.

Note that the second term in the variance decomposition is zero:

$$V_m E_m\left(\hat{R}_{it} \mid \mathbf{y}_{it-1}\right) = 0.$$

For the first term, we have:

$$E_m V_m\left(\hat{R}_{it} \mid \mathbf{y}_{it-1}\right) = \sigma_t^2 E_m \left[\frac{\sum_{s_{it}} w_{hlk}^2 y_{hlkt-1}}{\left(\sum_{s_{it}} w_{hlk} y_{hlkt-1} \right)^2} \right].$$

Thus, we have the following synthetic estimator of $V_d(\hat{R}_{it})$:

$$\hat{V}_{it}^S = \hat{\sigma}_t^2 \frac{\sum_{s_{it}} w_{hlk}^2 y_{hlkt-1}}{\left(\sum_{s_{it}} w_{hlk} y_{hlkt-1} \right)^2},$$

where

$$\hat{\sigma}_t^2 = \frac{1}{\sum_{s_t} w_{hlk}} \sum_{s_t} w_{hlk} \frac{(y_{hlkt} - \hat{R}_{it} y_{hlkt-1})^2}{y_{hlkt-1}}$$ and

s_t contains observations from all industries.

In our simulation, $\hat{V}_{it}^S$ performs much better than the corresponding design-based estimator $\hat{V}_{it}$ in terms of the design-variance. However, it is worse than $\hat{V}_{it}$ in terms of the design-bias property. Essentially, this approach avoids the problem of estimating the individual industry variances and thereby improves on the design-variance property at the expense of increasing the design-bias.

4. A ELB Method

In this section, we take adavantage of the good design-bias property of the design-based estimator and good design-variance property of the synthetic estimator in proposing a new ELB estimator. We propose to achieve this objective by considering a two-level model that accounts for the sampling variability of $\hat{V}_{it}$ and the variability of a model that attempts to link the design-based estimator to a corresponding alternate variance estimator (e.g., $\hat{V}_{it}^S$).

We concentrate on a fixed month, so from now on we shall drop the subscript *t*. Let $\mathbf{Y}_i$ denote the corresponding finite population, and $V_i(\mathbf{Y}_i) \equiv V_i$ denote true design-based variance of $\hat{R}_i$.

We consider the following two-level model :

Model 2:

Level 1: $E_d[\hat{V}_i \mid \mathbf{Y}_i] = V_i$ and $V_d[\hat{V}_i \mid \mathbf{Y}_i] = \sigma_i^2(\mathbf{Y}_i)$.

Level 2: $E_m[V_i] = \xi_i$ and $V_m[V_i] = \delta^2$.

We use Level 1 to incorporate the sampling distribution of $\hat{V}_i$ and Level 2 to link V_i to ξ_i , an auxiliary information about V_i. Level 2 essentially implies the existence of a superpopulation model that generates the finite population. Note that Model 2 is a robust model since we do not need any distributional assumptions beyond the specifications of the first two moments.

In order to stabilize the variance, we take a log transformation. Define $u_i = \ln(\hat{V}_i)$ and consider the following model on the transformed variable u_i:

Model 3:

Level 1: $E_d[u_i \mid \mathbf{Y}_i] = \theta_i(\mathbf{Y}_i)$ and $V_d[u_i \mid \mathbf{Y}_i] = \gamma_i^2(\mathbf{Y}_i)$.

Level 2: $E_m[\theta_i(\mathbf{Y}_i)] = \mu_i$ and $V_m[\theta_i(\mathbf{Y}_i)] = \tau^2$.

Note that we can view both Model 2 and Model 3 as robust Bayesian models. In both cases, Level 1 and Level 2 can be treated as the sampling and the prior distributions respectively. We can also treat them as robust mixed models where random effects are introduced through the specification of a superpopulation for the finite population. Similar models were considered in Ghosh and Lahiri (1987) and Ghosh and Meeden (1997) .

We are interested in estimating (in the Bayesian approach) or predicting (in the classical prediction approach) $V_i \equiv V_i(\mathbf{Y}_i)$. Here we adopt the Bayesian approach. We shall first estimate $\theta_i \equiv \theta_i(\mathbf{Y}_i)$.

We define the model MSE (same as the integrated Bayes risk) of an arbitrary estimator $\hat{\theta}_i$ of θ_i as $MSE(\hat{\theta}_i) = E(\hat{\theta}_i - \theta_i)^2$, where the expectation is taken with respect to the marginal distribution of Model 3. The linear Bayes estimator (LB) that minimizes *MSE* in the class of all linear estimators is given by

$$\hat{\theta}_i^{LB}(\boldsymbol{\varphi}) = B_i u_i + (1 - B_i)\mu_i,$$

where

$$B_i \equiv B_i(\boldsymbol{\varphi}) = \frac{\tau^2}{\tau^2 + \psi_i},$$

$$\psi_i = E_m\left[\gamma_i^2(\mathbf{Y}_i)\right],$$

$$\boldsymbol{\varphi} = (\psi_i, \mu_i, \tau^2)'.$$

In practice, $\boldsymbol{\varphi}$ is unknown and needs to be estimated from the data. We assume the following relationship between μ_i and the synthetic estimator $\hat{V}_i^S : \mu_i = \beta x_i$, where $x_i = \ln[\hat{V}_i^S]$ and β is an unknown slope.

We estimate ψ_i from the sample by $\hat{\psi}_i = \frac{\hat{V}_d[\hat{V}_i \mid \mathbf{Y}_i]}{\hat{V}_i^2}$, where $\hat{V}_d[\hat{V}_i \mid \mathbf{Y}_i]$ is a design-based estimator of $V_d[\hat{V}_i \mid \mathbf{Y}_i]$ (Lahiri and Gershunskaya, 2005).

For estimation of τ^2 and β, we use the method given in Fay and Herriot (1979). Plugging in the estimator $\hat{\boldsymbol{\varphi}} = (\hat{\psi}_i, \hat{\beta}, \hat{\tau}^2)$ for $\boldsymbol{\varphi}$, we obtain the following empirical linear Bayes (ELB) estimator of θ_i:

$$\hat{\theta}_i = \hat{\theta}_i^{ELB} = \hat{\theta}_i^{LB}(u_i; \hat{\boldsymbol{\varphi}}).$$

Finally, we take the reverse transformation to get an estimator of our parameter of interest:

$$\hat{V}_i^{LB} = \exp(\hat{\theta}_i^{LB}).$$

Remark: If both Level 1 and Level 2 of Model 2 are normal, then we can use the log-normal mean formula to obtain the exact LB of V_i. It is given by

$$\hat{V}_i^{LB} = \hat{V}_i^{LB}(\boldsymbol{\varphi}) = \exp(\hat{\theta}_i^{LB} + \frac{1}{2}\psi_i B_i).$$

Plugging in $\hat{\boldsymbol{\varphi}}$ for $\boldsymbol{\varphi}$ in $\hat{V}_i^{LB}(\boldsymbol{\varphi})$ we obtain the following ELB:

$$\hat{V}_i^{ELB} = \hat{V}_i^{LB}(\hat{\boldsymbol{\varphi}}).$$

Some bias corrections can be made using Chambers and Dorfman (2003) or Lahiri (2005). However, we shall not pursue this approach here since our simulation results provide an evidence of nonnormality of Level 1.

5. Monte Carlo Simulation

The main focus of this section is to evaluate the performances of different variance estimators of the ratios $\hat{R}_{it}$ used to estimate monthly relative employment changes, in terms of bias, variance and mean squared error with respect to the randomization principle. We draw 10,000 independent samples from the universe data set for the State of Alabama using a design which approximates the CES sampling design. These 10,000 simulated samples allow us to compute the statistics needed in Tables 1-3: relative bias $RB\left[\hat{V}\right] = E_d\left[\hat{V} - V\right] / V$, coefficient of variation $CV\left[\hat{V}\right] = \sqrt{V_d\left[\hat{V}\right]} / V$, relative root MSE $RRMSE\left[\hat{V}\right] = \sqrt{CV^2\left[\hat{V}\right] + RB^2\left[\hat{V}\right]}$, and confidence intervals based on the assumption of normality for an estimator $\hat{V}$ of variance V.

Note (Table 1) that the point estimators themselves, the ratios $\hat{R}_{it}$, are highly efficient, the highest CV being 2.5%. In contrast, the CV's of the variance estimators are generally very high (as high as 195.4%), showing the instability of the design-based variance estimators.

We compare bias and variance properties of three different design-based variance estimators in Table 1. Their performances are very similar: they exhibit small relative biases, however, their CV's are all very high, the BHS method being even more unstable for the domains considered here than the other two. Our delta method improves on both the BHS and RGBHS in terms of the CV.

Table 2 presents relative biases, CV's and relative root mean squared errors of the direct (based on Taylor series), synthetic and ELB variance estimators. In terms of the MSE, the ELB method performs the best among the three different variance estimators. For some of the industries, the ELB variance estimator cuts down the relative root MSE by more than half. It is interesting to note that although the ELB estimator is derived under a model it is doing a great job in terms of the design-based property.

For each of the simulated samples based on the estimated variances, we constructed the 90-percent confidence intervals under the assumption of normality. The coverage properties of the estimators are presented in Table 3. In most industries, percent of the samples covered by the interval based on the direct estimator is only slightly lower than nominal. Depending on the direction of the bias, the synthetic estimator's intervals give either under- or overcoverage. Coverage of the ELB estimator's intervals in most industries is close to that of the direct estimator's. In terms of the average length, ELB is comparable to the design-based method. However, in terms of the variability of the length, ELB is superior to the design-based method.

6. Concluding remarks

Reporting standard errors along with estimates has been a norm of any standard official publication. We observe that the ability to produce good design-based estimates does not necessarily guarantee good design-based variance estimates. In this paper, we have demonstrated that it is possible to improve on the design-based variance estimation by considering a suitable two-level robust model. We have proposed a variance estimator of the standard design-based variance estimator, a factor needed in generating our ELB variance estimates. Although, our method works reasonably well in our simulation, there is scope for further research in improving the variance estimator of the standard design-based variance estimator. Modeling is another issue that needs special attention. In spite of these possible criticisms, our paper provides a general framework to attack the important variance estimation problem. When a naïve normality-based confidence intervals of the growth rate are applied using our variance estimators, our method is quite comparable with the other methods. But, we reiterate that the problem of interval estimation is a different problem for which one needs to study the distribution of the estimated growth rate carefully in order to produce a better interval estimation procedure. The usual form of normality-based confidence interval, i.e. estimate plus or minus the margin of error is just too naïve for this purpose. More research is needed to address the problem of interval estimation.

Acknowledgement

Any opinions expressed in this paper are those of the authors and do not constitute policy of the Bureau of Labor Statistics.

References

Butani, S., Harter, R., and Wolter, K. (1997). Estimation procedures for the Bureau of Labor Statistics' Current Employment Statistics Program, *Proceedings of the Survey Research Methods Section,* American Statistical Association, 523-528.

Bureau of Labor Statistics (2004), Chapter 2, "Employment, hours, and earnings from the Establishment survey," BLS Handbook of Methods.

Bureau of Labor Statistics (2001), Chapter 7, "Estimation," *Current Employment Statistics Manual:* U.S. Bureau of Labor Statistics, Washington, DC.

Bureau of Labor Statistics (2003), "Revisions to Establishment Survey Data," *Employment Situation:* May 2003; U.S. Bureau of Labor Statistics, Washington, DC.

Butani, S., Stamas, G., and Brick, M. (1997). Sample redesign for the Current Employment Statistics Survey, *Proceedings of the Survey Research Methods Section,* American Statistical Association, 517-522.

Chambers, R.L. and Dorfman, A. H. (2003). Transformed variables in survey sampling, *SRI Methodology Working Paper M03/21.*

Cho, M., Eltinge, J., Gershunskaya, J., Huff, L., (2002). Evaluation of Generalized Variance Function Estimators for the U.S. Current Employment Survey, *Proceedings of the Survey Research Methods Section,* American Statistical Association, 534-539.

Fay, R.E. and Herriot, (1979). Estimates of Income for Small Places: an Application of James-Stein Procedure to Census Data, *Journal of American Statistical Association,* 74, 269-277.

Ghosh, M. and Lahiri, P. (1987), Robust empirical Bayes estimation of means from stratified samples, *Journal of the American Statistical Ass*ociation, Vol. 82, 1153-1162.

Ghosh, M. and Meeden, G. (1997), *Bayesian Methods for Finite Population Sampling*. London: Chapman and Hall.

Harter, R., Macaluso, M. and Wolter, K. (2003). Evaluating the Fundamentals of a Small Domain Estimator. *Survey Methodology*, 29, 63-70.

Huff, L., Eltinge, J., Gershunskaya, J. (2002). Exploratory Analysis of Generalized Variance Function Models for the U.S. Current Employment Survey, *Proceedings of the Survey*

Research Methods Section, American Statistical Association, 1519-1524.

Lahiri, P. (2005). Can a model unbiased estimator of the finite population mean be obtained for a linear model on transformed observations? *Unpublished manuscript.*

Lahiri, P., Gershunskaya, J. (2005). Variance Estimation in the Current Employment Statistics Program. *Unpublished report.*

Lahiri, P., and Wang, W. (1991). Estimation of all employee links for small domains – an application of empirical Bayes procedure, *Proceedings of the Workshop on Statistical Issues in Public Policy Analysis*, II-32-II-53 (Invited Paper).

Lohr, S.L. (1999). *Sampling: Design and Analysis,* Duxbury.

Rao, J.N.K. (2003). *Small Area Estimation*, Wiley.

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

Werking, G.S. (1997). Overview of the CES (Current Employment Statistics) redesign, *Proceedings of the Survey Research Methods Section,* American Statistical Association, 170-175.

Wolter, K., Huff, L., and Shao, J. (1998). Variance estimation for the Current Employment Statistics survey, presented at the Joint Statistical Meetings, Dallas, August 13, 1998.

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

Table 1. Relative biases and CV's of the direct variance estimators and CV's of the point estimator (in percentage)

Industry	*Relative bias*			*CV*			*CV of the Point Estimator*
	Taylor Series	*BHS*	*RGBHS*	*Taylor Series*	*BHS*	*RGBHS*	
1	-2.2	1.0	1.3	152.6	173.3	158.3	2.5
2	-3.0	-2.3	-2.3	73.3	100.0	75.9	1.5
3	-0.6	0.0	0.2	47.2	73.4	49.7	0.5
4	-1.6	0.7	1.1	48.0	73.9	51.3	0.5
5	-1.6	0.4	-0.2	91.2	125.8	93.9	1.0
6	-4.0	-4.4	-3.9	29.0	64.7	32.5	0.6
7	-10.2	-7.1	-8.4	195.4	221.7	198.7	1.2
8	4.5	7.0	7.3	94.3	104.1	97.3	1.4
9	1.9	3.5	2.8	59.4	92.6	62.7	0.7
10	-18.5	-18.1	-17.2	192.2	202.1	198.7	1.0
11	-3.2	-3.2	-2.7	38.5	75.4	42.5	0.6
12	1.2	1.2	1.9	34.2	69.0	37.6	0.8
13	0.5	3.2	1.9	46.0	89.7	49.6	1.7

Table 2. Relative biases, CV's, and relative root MSE of the direct, synthetic, and ELB estimators (in percentage)

Industry	*Relative bias*			*CV*			*Relative root MSE*		
	Direct	*Synthetic*	*ELB*	*Direct*	*Synthetic*	*ELB*	*Direct*	*Synthetic*	*ELB*
1	-2.2	66.8	-14.5	152.6	44.7	68.1	152.6	80.4	69.6
2	-3.0	-43.0	-28.4	73.3	14.4	27.5	73.3	45.3	39.5
3	-0.6	12.2	-11.0	47.2	31.1	29.5	47.2	33.4	31.4
4	-1.6	-9.4	-15.1	48.0	28.2	31.2	48.0	29.8	34.7
5	-1.6	144.6	1.6	91.2	65.2	53.3	91.3	158.6	53.3
6	-4.0	29.2	-7.4	29.0	32.3	23.4	29.3	43.5	24.5
7	-10.2	-7.5	-35.9	195.4	25.6	57.6	195.6	26.7	67.9
8	4.5	-1.1	-20.4	94.3	28.7	57.6	94.4	28.8	61.1
9	1.9	180.3	13.6	59.4	73.6	54.6	59.5	194.7	56.3
10	-18.5	-28.4	-41.6	192.2	18.9	39.7	193.1	34.2	57.5
11	-3.2	63.1	-4.5	38.5	41.4	30.8	38.6	75.5	31.1
12	1.2	-31.9	-10.4	34.2	17.5	20.4	34.2	36.4	22.9
13	0.5	69.5	1.4	46.0	44.0	36.8	46.0	82.2	36.8

Table 3. Coverage probability with average length and CV of length (in parenthesis) for different methods (nominal: 90%)

Industry	*Direct*	*Synthetic*	*ELB*	*True*
1	88.4 (0.073, 54.7)	95.4 (0.108, 12.8)	89.7 (0.074, 33.5)	91.4 (0.084)
2	89.0 (0.046, 31.2)	79.0 (0.036, 12.1)	84.1 (0.041, 17.6)	90.4 (0.049)
3	89.8 (0.016, 21.8)	91.1 (0.017, 13.2)	88.1 (0.015, 16.0)	90.2 (0.016)
4	88.5 (0.015, 23.1)	87.6 (0.015, 14.4)	86.4 (0.014, 18.0)	90.5 (0.015)
5	89.5 (0.031, 34.8)	98.3 (0.051, 12.8)	91.2 (0.032, 24.7)	90.3 (0.033)
6	88.9 (0.019, 14.8)	93.1 (0.022, 12.0)	88.5 (0.019, 12.6)	90.0 (0.020)
7	88.7 (0.032, 59.0)	91.3 (0.037, 13.3)	86.3 (0.030, 31.9)	92.8 (0.039)
8	84.9 (0.041, 50.8)	88.6 (0.045, 13.9)	84.2 (0.038, 38.5)	90.2 (0.045)
9	89.3 (0.024, 27.5)	99.2 (0.040, 12.5)	91.5 (0.025, 24.0)	90.2 (0.024)
10	89.3 (0.025, 53.1)	90.7 (0.027, 12.6)	86.9 (0.024, 23.2)	95.3 (0.032)
11	89.7 (0.018, 18.7)	95.7 (0.023, 12.1)	89.5 (0.018, 15.7)	90.2 (0.018)
12	89.9 (0.025, 15.5)	82.0 (0.021, 12.3)	87.9 (0.024, 11.1)	89.9 (0.025)
13	89.2 (0.055, 22.3)	96.2 (0.072, 12.4)	89.7 (0.055, 18.4)	90.0 (0.056)