Comparison of Variance Estimation Methods Using PPI Data October 2010

Helen Chen, Andy Sadler
U.S. Bureau of Labor Statistics
2 Massachusetts Avenue NE, Washington DC 20212 U.S.A.
sadler.andy@bls.gov

Abstract
The Producer Price Index (PPI) collects price data from domestic producers of commodities and publishes monthly indexes on average price changes received by those producers at all stages of processing. PPI samples employ a two-stage design where establishments are selected in the first stage and unique items are selected in the second stage. In this paper we review the research results from the PPI variance estimation study. The objective of the study was to determine the best method of variance estimation appropriate for PPI data. Historical data from eleven NAICS industries were used to create simulation frames, from which simulation samples were drawn and estimated variances calculated. The replication methods compared were the Balanced Repeated Replication (BRR), Fay's BRR, Jackknife and the Bootstrap. The Bootstrap method was recommended for the PPI program by the study.

Key Word s: NAICS**;** Balanced Repeated Replication; Fay's Method; Stratified Jackknife; Bootstrap.

1. Introduction

Recently, the PPI chartered a team to determine the best method of variance estimation appropriate for PPI data. The team selected eleven NAICS industries for inclusion in the study. Variances for 1-month and 12-month index percentage change estimates were calculated. The replication methods studied were: the Balanced Repeated Replication (BRR or standard BRR), Fay's method of BRR, the stratified Jackknife method (JK) and the Bootstrap method. This paper reviews the major findings of the PPI variance estimation study.

The Producer Price Index (PPI) of the Bureau of Labor Statistics (BLS) is a family of indexes that measure the average change over time in the prices received by domestic producers of goods and services. PPIs measure price change from the perspective of the seller. More than 100,000 price quotations per month are organized into three sets of PPIs: (1) Stage-of-processing indexes, (2) commodity indexes, and (3) indexes for the net output of industries and their products. The stage-of processing structure organizes products by class of buyer and degree of fabrication. The commodity structure organizes products by similarity of end use or material composition. The entire output of various industries is sampled to derive price indexes for the net output of industries and their products. PPIs for the net output of industries and their products are grouped according to the North American Industry Classification System (NAICS).

2. Sampling

The PPI uses the BLS sample and research database known as the Longitudinal Database (LDB) as the source of frame information for most of the industries sampled. The LDB contains U.S. business frame records representing all U.S. non-farm industries, with the exception of some sole proprietors. The LDB consists of all covered employers under the Unemployment Insurance (UI) Tax System. The frame information used to cluster establishments on the LDB is the Employer Identification Number (EIN).

The 6-digit NAICS industries are sampled using a two-stage design. First-stage sample units are selected in the Washington office from a list of establishments and clusters of establishments whose primary production is thought to be in a given 6-digit NAICS industry. The final or second-stage sample units are then selected during data collection at the location of the sampled establishment. The second-stage units are unique items, products, or services for which the respondent is to report prices monthly for 5-7 years.

The first-stage sample units are selected systematically with probability proportional to a measure of size. The measure of size is usually employment when the most common source of frame information, the Longitudinal Database, is used for sampling. The measure of size is thought to correlate with revenue, which is collected directly from a sampled unit and used in weights of items in index calculation. The second-stage sample units are selected in the field at the location of the establishment selected in the first stage.

3. Index Estimation

3.1 Calculating lowest level cell index

We calculate the long-term price relative for each item as follows:

$$LTR_{k,c,t} = \frac{P_{k,c,t}}{P_{k,c,0}}$$

where $LTR_{k,c,t}$ = long-term price relative of item *k* with a good price in cell *c* at time *t*, $p_{k,c,t}$ = price of item *k* in cell *c* at time *t*, $p_{k,c,0}$ = price of item *k* in cell *c* in base period 0. Missing price relatives are estimated as follows:

$$LTR_{j,c,t} = LTR_{j,c,t-1} \frac{\sum_{k=1}^{n_g} w_{k,c} LTR_{k,c,t}}{\sum_{k=1}^{n_g} w_{k,c} LTR_{k,c,t-1}}$$

where $LTR_{j,c,t}$ = estimated missing long-term price relative for item j, $j = n_{g+1},\ldots,n_c$, n_g is the number of items with a good price in cell *c* at time *t*, n_c is the total number of items in cell *c* at time *t*, $LTR_{j,c,t-1}$ = long-term price relative of item *j* in cell *c* at time *t*-1, $LTR_{k,c,t}$ = long-term price relative for item *k* which has a good price, *k* = 1,…, n_g, $LTR_{k,c,t-1}$ = long-term price relative of item *k* in cell *c* at time *t*-1, $w_{k,c}$ = weight of item *k* in cell *c*. The cell aggregate is calculated as follows:

$$CA_{c,t} = \sum_{i=1}^{n_c} w_{i,c} LTR_{i,c,t}$$

where $CA_{c,t}$ = cell aggregate for cell *c* at time *t*, $w_{i,c}$ = weight of item *i* in cell *c*, $LTR_{i,c,t}$ = long-term price relative of item *i* in cell *c* at time *t*, n_c = total number of items in cell *c*. The cell index is calculated as follows: $I_{c,t} = \frac{CA_{c,t}}{CA_{c,t-1}} I_{c,t-1}$, where $I_{c,t}$ = index for cell *c* at time *t*, $I_{c,t-1}$ = index for cell *c* at time *t*-1, $CA_{c,t}$ = cell aggregate for cell *c* at time *t*, $CA_{c,t-1}$ = cell aggregate for cell *c* at time *t*-1.

3.2 Percent change for an index

The following formula calculates the percentage change for an index using cell index as an example:

$$PC_{c,t,t-m} = \left(\frac{I_{c,t}}{I_{c,t-m}} - 1\right)100$$

where $PC_{c,t,t-m}$ = percentage change of index for cell *c* from time *t-m* to time *t*, m is equal to 1month, 3 months, or 12 months, $I_{c,t}$ = index for cell *c* at time *t*, and $I_{c,t-m}$ = index forcell *c* at time *t-m*. In the PPI study, we computed estimates of I and PC for each timeperiod and calculated the variance of PC using the replication methods to be discussed.

4. Variance Estimation Methods for Item-based Industry Indexes

Replication was used to estimate the variance of industry cell indexes in the PPI study. In replication, we calculate the estimate of the index from the full sample as well as a number of subsamples (replicates). Replicates, i.e., subsets of primary sampling units (PSUs) of a sample, are formed, the sampling weights of the PSUs are adjusted, and replicate index estimates are calculated in the same way that the full sample estimate is calculated. The variation among the replicate estimates is used to estimate the variance for the full sample.

4.1 The Bootstrap Method

4.11 Forming bootstrap variance strata

Strata for variance estimation were set up in the following way: We placed all probability establishments in a single variance stratum. The probability establishments served as PSUs within the single stratum. Certainty establishments with more than 1 item were put into their own stratum. The items of certainty establishments served as variance PSUs within the stratum. Certainty establishments with only one item were paired with another single-item certainty establishment, if one existed, or with a multi-item certainty establishment to form a variance stratum. We collapsed a maximum of three single item certainty establishments into any one variance stratum. If there was only one certainty establishment with one item we placed this establishment in every replicate.

4.12 Replicate formation and bootstrap weights

Define n_h as the total number of variance PSUs in a variance stratum h. Let m^*_{hi} be the number of times that the variance PSU i in variance stratum h is selected in the bootstrap procedure. A replicate is formed by drawing m_h variance PSUs with replacement from the n_h variance PSUs in variance stratum *h*, so that $\sum_{i=1}^{n_h} m^*_{hi} = m_h$.

Note $m^*_{hi} = 0$ if the i^{th} variance PSU is not selected in variance stratum h. Define the bootstrap item weights as:

$$w^*_{hik} = w_{hik}\left[\left\{1 - \sqrt{\frac{m_h}{n_h - 1}}\right\} + \sqrt{\frac{m_h}{n_h - 1}\left(\frac{n_h}{m_h}\right)m^*_{hi}}\right]$$

where w_{hik} is the original sample weight for the item k in variance PSU i of variance stratum h. If $m_h = n_h - 1$, then $w^*_{hik} = w_{hik}\left[\left(\frac{n_h}{n_h - 1}\right)m^*_{hi}\right]$. The number of bootstrap replicates formed will be 150.

4.13 Bootstrap variance formula

We formed B (B=150) bootstrap replicates and computed the bootstrap variance estimator, v_{BT} , in the following way: $V_{BT} = \frac{1}{B}\sum_{b=1}^{B}\left(\hat{\theta}_b - \hat{\theta}_{full}\right)^2$ where $\hat{\theta}_b$ is a bootstrap estimator of replicate b and $\hat{\theta}_{full}$ is the estimator from the original sample.

4.2 Balanced Repeated Replication (BRR) and Fay's Method of BRR
4.21 Forming BRR and Fay'sBRR variance strata
Variance strata for the BRR and Fay's BRR methods were formed in the following way: Each certainty establishment was a separate variance stratum. For these certainty establishments, the PSUs were items. The number of certainty establishments was n_{cert}.
Each establishment selected with probability was paired with another one in a variance stratum. The probability establishments were the PSUs. If there were an odd number of PSUs, one stratum contained three PSUs. The number of probability establishments was n_{prob}. The number of variance strata, L was determined as follows. If n_{prob} was even, then $L = n_{cert} + \frac{n_{prob}}{2}$. If n_{prob} was odd, then $L = n_{cert} + \frac{n_{prob}-1}{2}$.
.

4.22 Replicate formation, number of replicates, and item weight adjustments
Within each variance stratum two variance groups were formed. Each variance group contained one or more PSUs. The items of each certainty establishment formed two variance groups within a variance stratum. Two adjacent probability establishments were combined to form one variance stratum, beginning with those with the largest measure of size. The number of BRR and Fay's BRR replicates formed is an integral multiple of 4 that is greater than the number of variance strata, L. For each replicate formed a corresponding complement was formed using the variance groups not included in the replicate. In the standard BRR method, the item weights of PSUs included in a replicate are multiplied by 2 to account for the PSUs excluded from a replicate. In the Fay's method the item weights of PSUs in a replicate are multiplied by a factor of $2-K$, where $0 \le K < 1$, and the remaining PSUs' item weights are multiplied by the factor K. All of the PSUs of a sample are used in a Fay's method replicate estimate.

4.23 BRR variance formulas
The following BRR variance formulas were used. The subscripts identifying these standard BRR variance formulas are BRR1_k0, BRR2_k0, BRR3_k0, and BRR4_k0. According to Wolter (1985), if the estimator $\hat{\theta}$ is linear, then these variance estimatorsare identical. In our case, $\hat{\theta}$ is nonlinear and the variance estimators are unequal. $\hat{V}_{BRR1}\left(\hat{\theta}\right)$, $\hat{V}_{BRR2}\left(\hat{\theta}\right)$, and $\hat{V}_{BRR3}\left(\hat{\theta}\right)$ are sometimes regarded as estimators of the mean squared error MSE $\left\{\hat{\theta}\right\}$, while $\hat{V}_{BRR4}\left(\hat{\theta}\right)$ is regarded as an estimator of variance Var $\left\{\hat{\theta}\right\}$. Wolter shows that $\hat{V}_{BRR3}\left(\hat{\theta}\right)$ will be larger than $\hat{V}_{BRR4}\left(\hat{\theta}\right)$. The k0 in our designation refers to the Fay's K value that is equal to 0 for the standard BRR method.

$$\hat{V}_{BRR1\_k0}\left(\hat{\theta}\right) = \frac{1}{G}\sum_{g=1}^{G}\left(\hat{\theta}_{(g,r)} - \hat{\theta}\right)^2,$$

where $\hat{\theta}_{(g,r)}$ is the gth replicate estimate of θ based on the items included in the gth replicate, $\hat{\theta}$ is the estimate of θ based on the full sample, G is the total number of replicates formed, and $\hat{V}_{BRR1_k0}\left(\hat{\theta}\right)$is the estimated BRR1_k0 variance of $\hat{\theta}$.

$$\hat{V}_{BRR2\_k0}\left(\hat{\theta}\right) = \frac{1}{G}\sum_{g=1}^{G}\left(\hat{\theta}_{(g,c)} - \hat{\theta}\right)^2,$$

where $\hat{\theta}_{(g,c)}$ is the gth complement estimate of θ based on the items included in the gth complement, $\hat{\theta}$ is the estimate of θ based on the full sample, G is the total number of complements formed, and $\hat{V}_{BRR2\_k0}(\hat{\theta})$ is the estimated BRR2_k0 variance of $\hat{\theta}$.

$\hat{V}_{BRR3\_k0}(\hat{\theta}) = \frac{1}{2}[\hat{V}_{BRR1\_k0}(\hat{\theta}) + \hat{V}_{BRR2\_k0}(\hat{\theta})]$,

where $\hat{V}_{BRR1\_k0}(\hat{\theta})$ is the estimated BRR1_k0 variance of $\hat{\theta}$, $\hat{V}_{BRR2\_k0}(\hat{\theta})$ is the estimated BRR2_k0 variance of $\hat{\theta}$, and $\hat{V}_{BRR3\_k0}(\hat{\theta})$ is the estimated BRR3_k0 variance of $\hat{\theta}$.

$\hat{V}_{BRR4\_k0}(\hat{\theta}) = \frac{1}{4G}\sum_{g=1}^{G}\left(\hat{\theta}_{(g,r)} - \hat{\theta}_{(g,c)}\right)^2$,

where $\hat{\theta}_{(g,r)}$ is the gth replicate estimate of θ based on the items included in the gth replicate, $\hat{\theta}_{(g,c)}$ is the gth complement estimate of θ based on the items included in the gth complement, G is the total number of pairs of replicates and complements, and $\hat{V}_{BRR4\_k0}(\hat{\theta})$ is the estimated BRR4_k0 variance of $\hat{\theta}$.

4.24 Fay's BRR variance formulas

The Fay's method variance formulas are similar to the standard BRR formulas above and the K or $2-K, (0 \le K < 1)$ factor is applied to the PSU item weights. Two K factors, 0.1 and 0.5, were tested. The Fay's BRR variance formulas are designated BRR1_k10, BRR1_k50, BRR2_k10, BRR2_k50, BRR3_k10, BRR3_k50, BRR4_k10, and BRR4_k50. The k10 refers to $K = 0.1$ and k50 refers to $K = 0.5$. Only the BRR1_k50 formula will be shown below as an example of the Fay's method formulas.

$$\hat{V}_{BRR1\_k50}(\hat{\theta}) = \frac{1}{G(1-K)^2}\sum_{g=1}^{G}\left(\hat{\theta}_{(g)} - \hat{\theta}\right)^2$$

where $\hat{\theta}$ is the estimate of θ based on the full sample, $\hat{\theta}_{(g)}$ is the gth replicate estimate of θ based on the full sample of items, where the weights of half of the items are multiplied by K and the other half of the items are multiplied by $(2-K)$, G is the total number of replicates formed, K is a factor used to modify the sample weights, and $\hat{V}_{BRR1\_k50}(\hat{\theta})$ is the estimated Fay's method variance of $\hat{\theta}$, which corresponds to the BRR1_k0 method with $K = 0.5$.

4.3 Stratified jackknife method

Forming stratified jackknife variance strata

Variance strata and PSUs for the stratified jackknife (JK) method are set up exactly as in the BRR and Fay's BRR methods. Each variance stratum has the same two variance groups formed for the BRR methods, with each variance group containing one or more variance PSUs.

4.31 Replicate formation, number of replicates, and item weight adjustments

The number of replicates is equal to two times the number of variance strata. For each variance stratum, one replicate is formed by dropping one variance group from the variance stratum and doubling the item weights of the remaining variance group in the stratum. The variance groups of all other variance strata are retained in the replicate without item weight adjustment. The second replicate in a variance stratum is formed by dropping the other variance group from the stratum, doubling the item weights of the remaining variance group and retaining the variance groups in all other variance strata without item weight adjustment.

4.32 JK variance formulas

Two stratified jackknife formulas are used to calculate the variance of $\hat{\theta}$, designated JK1 and JK2.

$$\hat{V}_{JK1}(\hat{\theta}) = \sum_{h=1}^{L} \frac{1}{2}\left(\left(\hat{\theta}_{(h1)} - \hat{\theta}\right)^2 + \left(\hat{\theta}_{(h2)} - \hat{\theta}\right)^2\right)$$

where $\hat{\theta}$ is the estimate of θ based on the full sample, $\hat{\theta}_{(hi)}$ is the replicate estimate of θ calculated when variance group i is dropped from stratum h, L is the total number of variance strata formed (recall that each stratum has 2 groups), $\hat{V}_{JK1}(\hat{\theta})$ is the estimated JK1 variance of $\hat{\theta}$.

$$\hat{V}_{JK2}(\hat{\theta}) = \sum_{h=1}^{L} \frac{1}{4}\left(\hat{\theta}_{(h1)} - \hat{\theta}_{(h2)}\right)^2$$

where $\hat{\theta}_{(hi)}$ is the replicate estimate of θ calculated when variance group i is dropped from stratum h, L is the total number of variance strata formed and $\hat{V}_{JK2}(\hat{\theta})$ is the estimated JK2 variance of $\hat{\theta}$.

5. Impact of missing price imputation

For missing prices, the PPI index estimation system imputes a price based on the percentage change of the reported prices in the cell. For any replicate, imputed prices will vary depending on which reporting prices are included in the replicate. In the simulation study we used the full imputation procedure. In the full imputation procedure, values from the PPI research database were imputed by the PPI index estimation system are treated as missing values, imputed separately for each replicate and complement, and imputed values are carried forward and used to calculate estimates for the following month.

6. Simulation Study

6.1 Creation of Simulation Frame and Drawing of Simulation Samples

6.12 Number of certainty and probability establishments

We used the original sample and pricing data for the two-year study period of each industry as the base for the sampling frames for the simulation study. We made some alterations to the original samples so that they could be used as simulation frames. The certainty establishments in a sample represent only themselves and were carried directly from the sample to the created simulation sampling frames. The probability establishments in the original sample each represent several probability establishments in the original frame. The probability establishments in the original sample were duplicated, with the number of duplicates determined as follows:

1) For each probability establishment in the sample, take the decimal part of the sampling factor, *d*
2) Round the sampling factor up to an integer with probability *d* or down to an integer with probability 1- *d*.
3) The resulting rounded sample factor *D* was the number of establishments that were placed in the created simulation sampling frame for the specified probability establishment. Certain *D* values were judged unusually high and were lowered on a case-by-case basis.

Having determined the number of establishments to be used in the created sampling frame based on the probability establishments, we next determined the values of those establishments. Using exact duplicates of the probability establishments may have resulted in the created simulation sampling frame being far more homogenous than the actual sampling frame, and hence, underestimating variances. We attempted to avoid this problem by adding noise to the duplicated establishments. Each duplicated establishment had the same item composition as the original establishment from which it was duplicated. However, noise was added to the initial prices of those items and to the manner in which they change over time.

6.13 Drawing of Simulation Samples

A simulation sample for an industry was formed by including from the simulation frame all of the certainty establishments with their items and a sample of the duplicated probability establishments with their items. The sample of probability establishments was drawn systematically with probability proportional to size (PPS). The number of probability establishments selected in the simulation sample was equal to the number of probability establishments selected in the original sample.

7. Comparison Statistics

Population values are first computed for 1-month percentage change and 12-month percentage change for each lowest level cell and aggregate level cell for each of the eleven industries in the study. These population values for index percentage change θ_{ct}, are computed using all of the items in the expanded frames for all cells *c* for all of the time periods *t*. There are twenty-three 1-month time periods from February 2004 to December 2005 and twelve 12-month time periods from January 2005 to December 2005. From the 300 samples that were drawn from the expanded frames the following statistics are computed for all cells *c*, which include the lowest level cells and aggregate level cells for each industry and each time period:

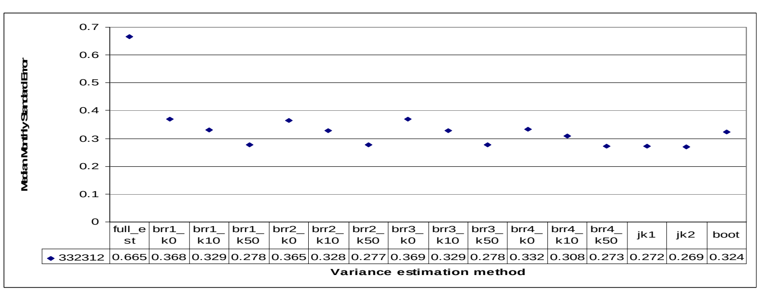
- Estimates of $\theta_{ct}\left(\hat{\theta}_{cts}\right)$, where $s = 1, \ldots, 300$ for cell c and period t.
- Simulation empirical values for variance, $V\left(\hat{\theta}_{ct}\right)$, and standard error, $\sigma\left(\hat{\theta}_{ct}\right)$, of 1-month and 12-month percentage change for cell *c* and time period *t*: $V\left(\hat{\theta}_{ct}\right) = \frac{1}{299}\sum_{s=1}^{300}\left(\hat{\theta}_{cts} - \hat{\bar{\theta}}_{ct}\right)^2$ where $\hat{\bar{\theta}}_{ct} = \frac{1}{300}\sum_{s=1}^{300}\hat{\theta}_{cts}$, $\sigma\left(\hat{\theta}_{ct}\right) = \sqrt{V\left(\hat{\theta}_{ct}\right)}$. ($\sigma\left(\hat{\theta}_{ct}\right)$ is called the Simulation Empirical Standard Error).
- Estimates of variance, $\hat{V}_m\left(\hat{\theta}_{cts}\right)$, and standard error, $\hat{\sigma}_m\left(\hat{\theta}_{cts}\right) = \sqrt{\hat{V}_m\left(\hat{\theta}_{cts}\right)}$, of 1-month and 12-month percentage change for each of the variance estimation methods m, for each cell c, time period t, and sample s.
- Relative bias of variance estimation method m for cell c in time period t:
  $$bias_{mct} = \frac{\left[\frac{1}{300}\sum_{s=1}^{300}\hat{\sigma}_m\left(\hat{\theta}_{cts}\right) - \sigma\left(\hat{\theta}_{ct}\right)\right]}{\sigma\left(\hat{\theta}_{ct}\right)}$$
- Relative stability (relative standard error of the variance) of variance estimation method m for cell *c* in time period *t*:
  $$Rel\ Stability_{mct} = \frac{\sqrt{\frac{1}{299}\sum_{s=1}^{300}\left[\hat{V}_m\left(\hat{\theta}_{cts}\right) - \hat{\bar{V}}_m\left(\hat{\theta}_{ct}\right)\right]^2}}{\hat{\bar{V}}_m\left(\hat{\theta}_{ct}\right)},$$
  where $\hat{\bar{V}}_m\left(\hat{\theta}_{ct}\right) = \frac{1}{300}\sum_{s=1}^{300}\hat{V}_m\left(\hat{\theta}_{cts}\right)$.

- The mean width of confidence interval for each variance estimation method: $\bar{l}_{mct} = \frac{1}{300}\sum_{s=1}^{300}\left(\hat{\theta}_{ctsU} - \hat{\theta}_{ctsL}\right)$, where $\hat{\theta}_{ctsL} = \hat{\theta}_{cts} - t_{df,\alpha/2}\sqrt{\hat{V}_m\left(\hat{\theta}_{cts}\right)}$; $\hat{\theta}_{ctsU} = \hat{\theta}_{cts} + t_{df,\alpha/2}\sqrt{\hat{V}_m\left(\hat{\theta}_{cts}\right)}$, $df1 = \frac{2\left[\bar{\hat{V}}_m\left(\hat{\theta}_{ct}\right)\right]^2}{\frac{1}{299}\sum_{s=1}^{300}\left[\hat{V}_m\left(\hat{\theta}_{cts}\right)-\bar{\hat{V}}_m\left(\hat{\theta}_{ct}\right)\right]^2}$, $df2 = (\#Variance\ Strata) - (\#\ PSUs)$.
- A bias-adjusted variance $\hat{V}_m'\left(\hat{\theta}_{cts}\right)$, , was calculated as shown below and was used in place of $\hat{V}_m\left(\hat{\theta}_{cts}\right)$ in an alternative method of calculating confidence intervals. Bias-adjusted variance = $\hat{V}_m'\left(\hat{\theta}_{cts}\right)$ = $\hat{V}_m\left(\hat{\theta}_{cts}\right) - \left\{\left[\frac{1}{300}\sum_{s=1}^{300}\hat{V}_m\left(\hat{\theta}_{cts}\right)\right] - V\left(\hat{\theta}_{ct}\right)\right\}$, if $\hat{V}_m\left(\hat{\theta}_{cts} \geq \left\{\left[\frac{1}{300}\sum_{s=1}^{300}\hat{V}_m\left(\hat{\theta}_{cts}\right)\right] - V\left(\hat{\theta}_{ct}\right)\right\}\right)$, otherwise Bias-adjusted variance = $\hat{V}_m'\left(\hat{\theta}_{cts}\right) = 0$, where $\hat{V}_m\left(\hat{\theta}_{cts}\right)$ is the original estimate of variance of 1-month or 12-month percentage change for each of the variance estimation methods *m*, for each cell *c*, time period *t*, and sample *s*. $\left\{\left[\frac{1}{300}\sum_{s=1}^{300}\hat{V}_m\left(\hat{\theta}_{cts}\right)\right] - V\left(\hat{\theta}_{ct}\right)\right\}$ is the bias of variance estimation method *m*, for cell *c*, in time period *t*.
- Confidence interval coverage rates of the variance estimation methods computed as the proportion of intervals that contain the population value of θ_{ct} . We could also use the simulation empirical value of $\hat{\hat{\theta}}_{ct}$ in place of θ_{ct} . The interval will be calculated using $Coverage\ rate_{mct} = \frac{1}{300}\sum_{s=1}^{300} I\left\{\theta_{ct}\epsilon\left(\hat{\theta}_{ctsL}, \hat{\theta}_{ctsU}\right)\right\}$

8. Results of simulation study

8.1 Median Monthly and Mean Yearly Percentage Change and Standard Error Estimates

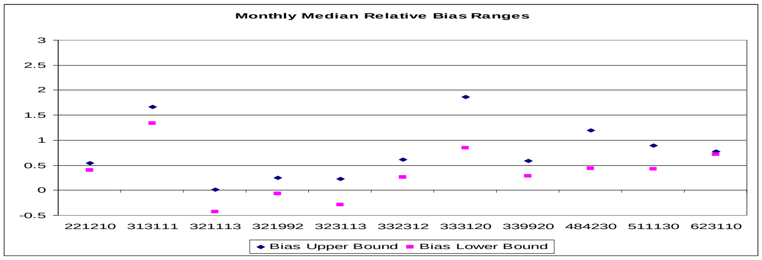
The median values of the 23 monthly standard error values for the variance estimators are shown in the graph 8.1 below. The median of the full sample percentage change estimates (full_est) is also shown on the same scale. This graph shows the typical pattern in the magnitude of the BRR standard error estimates with the standard BRR (k0) estimates larger than the Fay's method estimates, Fay's method k10 estimates greater than the k50 estimates, and BRR4 estimates smaller than BRR1, BRR2, and BRR3 estimates of same *k* value.



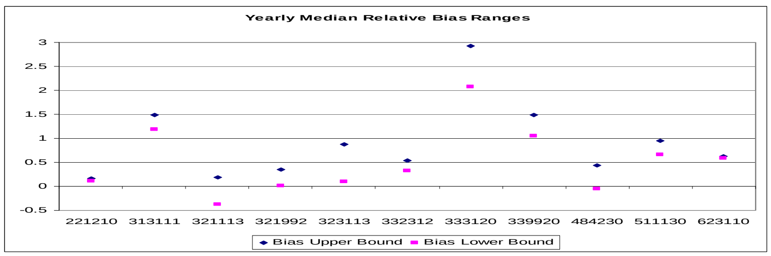
Graph 8.1 Median monthly sample pct change and standard error estimates

8.2 Relative Bias of Variance Estimators: Median Relative Bias Plots and Data for 6-digit level cells

The relative bias of each variance estimator was calculated for each time period. The median relative bias of all the time periods for each variance estimator is used to compare the variance estimators within a cell and to compare the relative bias values across the industries. The median relative bias range of the variance estimators for each industry for monthly and yearly percentage change estimates is shown on Graphs 8.2A and 8.2B below. The actual values are shown on Table 8.2A. All discussion refers to the 6-digit level cells unless stated otherwise.



Graph 8.2A Monthly median relative bias range



Graph 8.2B Yearly median relative bias ranges

	Monthly Yearly			
	Relative bias upper bound	**Relative bias lower bound**	**Relative bias upper bound**	**Relative bias lower bound**
NAICS				
221210	**boot brr4_k50**		**brr2_k0**	**brr4_k50**
	0.53762	0.40076	0.15858	0.11735
313111	**brr3_k0 boot**		**brr1_k0 brr4_k50**	
	1.66739	1.34318	1.48927	1.19022
321113	**brr1_k0 brr4_k50**		**brr2_k0 brr4_k50**	
	0.00932	-0.42945	0.19308	-0.37981
321992	**brr3_k0 brr4_k50**		**brr2_k0 brr4_k50**	
	0.25161	-0.06531	0.34639	0.01356
323113	**brr1_k0 brr4_k50**		**brr2_k0 brr4_k50**	
	0.23025	-0.28488	0.87126	0.10486
332312	**brr3_k0 jk2**		**brr1_k0 brr4_k50**	
	0.61116	0.25591	0.53392	0.32874
333120	**brr3_k0 brr4_k50**		**brr1_k0 jk2**	
	1.85934	0.8482	2.93095	2.06958
339920	**brr2_k0 brr4_k50**		**brr1_k0 brr4_k50**	
	0.59115	0.28991	1.49095	1.04499
484230	**brr3_k0 brr4_k50**		**brr3_k0 brr4_k50**	
	1.2009	0.44167	0.43342	-0.04527
511130	**brr3_k0 brr4_k50**		**brr2_k0 brr4_k50**	
	0.89827	0.41943	0.94931	0.6573
623110	**boot jk2**		**brr3_k0**	**jk2**
	0.77382	0.71492	0.62788	0.58401

Table 8.2A Median Relative Bias Ranges of Variance Estimators

The two graphs above show the median relative bias for the 1-month and 12-month percentage change estimates for all of the industries. For each industry, the range of median relative bias values of the variance estimators is indicated on the graphs by the bias upper bound and bias lower bound plotted values. In all of the industries, except for three of the monthly estimates and two of the yearly estimates, the median relative bias of the variance estimators is greater than zero, indicating that the sample percentage change standard errors are overestimating the "true" percentage change standard error represented by our simulation empirical standard error. The amount of overestimation varies by industry. For the monthly percentage change estimates, the midpoint of the ranges varies from a low of 0.43 (43% larger than the simulation empirical standard error) for 332312 to a high of 1.51 (151% larger than the simulation empirical standard error) for 313111. For the yearly percentage change estimates, the midpoint of the ranges varies from a low of 0.14 (14% larger than the simulation empirical standard error) for 221210 to a high of 2.50 (250% larger than the simulation empirical standard error) for 333120. For the industries with at least one of the variance estimators showing negative bias, for the monthly estimates, the median relative bias range midpoint falls between -0.21 and 0.09. For the yearly estimates, the median relative bias range midpoint falls between -0.09 and 0.19.

The variance estimators tend to be ordered with respect to median relative bias. The variance estimators with the largest and smallest values tend to be consistent across the industries. One of the standard BRR estimators, either BRR1_k0, BRR2_k0, or BRR3_k0, has the largest median relative bias in 9 of 11 industries for the monthly estimates and all 11 industries for the yearly estimates. One of these estimators will have the largest positive bias of all the variance estimators in a cell or the smallest negative bias in a cell if all the variance estimators are negatively biased. The BRR4_k50 estimator tends to have the smallest value of the median relative bias. In 8 of 11 industries for the monthly estimates and 9 of 11 industries for the yearly estimates, the BRR4_k50 estimator has the smallest positive or largest negative median relative bias. Between the k0 and BRR4_k50 estimators, the other estimators may fall in any order, except that the B, k10, and JK1 estimators tend to fall closer to the k0 estimators more often and the JK2 and BRR3_k50 estimators tend to fall closer to BRR4_k50 more often. The BRR bias values tend to be grouped by *k* value with BRR4 having the smallest value for a particular *k* value.

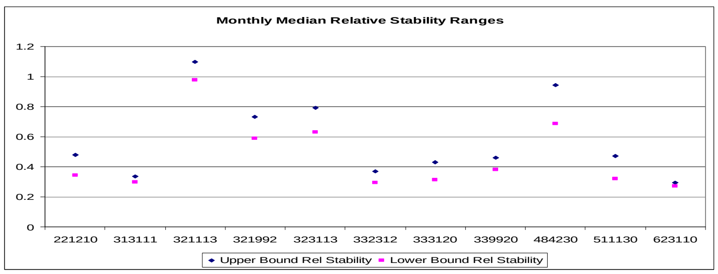
8.3 Relative Stability of the Variance Estimators

8.31 Median Relative Stability Plots and Data for 6-digit level cells

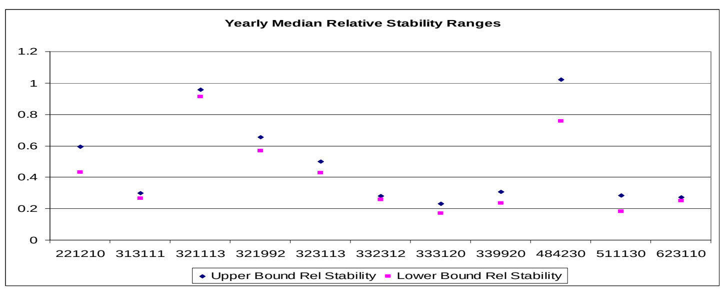
The relative stability, or relative standard error of the variance, of each variance estimator was calculated for each time period. The median relative stability of all the time periods for each variance estimator is used to compare the variance estimators within a cell and to compare the relative stability values across the industries and is shown in Graphs 8.31A and 8.31B below. All discussion refers to the 6-digit level cells unless stated otherwise.

The two graphs below show the median relative stability for the 1-month and 12-month percentage change variance estimates for all of the industries. For each industry, the range of median relative stability values of the variance estimators is indicated on the graphs by the Upper Bound Rel Stability and Lower Bound Rel Stability plotted values. The upper and lower bound values are shown in Table 8.31 below the graphs.

For the monthly median relative stability values, for 7 of 11 of the industries, the midpoint of the relative stability range of the variance estimators is around 0.3-0.4. The values of relative stability of the 4 industries with larger relative stability range midpoints are near 0.7-0.8 for 3 industries with one industry with midpoint over 1.0. Three of the 4 industries with larger relative stability values are the industries that had negative lower bound values for median relative bias. The yearly median relative stability values are similar to the monthly values. Six of 11 of the industries have the midpoint of the range of median relative stability values of the variance estimators from 0.2-0.3. The midpoints of 3 of the industries are from 0.4-0.6. The 2 remaining industries have the largest median relative stability midpoint values of around 0.9. With the yearly estimates, the 2 industries with the largest median relative stability values of the variance estimators, 321113 and 484230, had some variance estimators with negative median relative bias values.



Graph 8.31A Monthly median relative stability ranges



Graph 8.31B Yearly median relative stability ranges

	Monthly Yearly			
	Upper Bound Rel Stability	**Lower Bound Rel Stability**	**Upper Bound Rel Stability**	**Lower Bound Rel Stability**
NAICS				
221210	**jk2 boot**		**brr4_k50**	**boot**
	0.48041	0.34196	0.59547	0.43135
313111	**brr2_k0 boot**		**brr1_k0**	**brr3_k0**
	0.33505	0.29774	0.29818	0.26501
321113	**jk1 brr2_k50**		**brr2_k10**	**brr3_k0**
	1.0972	0.97899	0.95622	0.91053
321992	**jk2 boot**		**jk2**	**brr3_k0**
	0.73089	0.58877	0.65341	0.56865
323113	**brr2_k0 brr2_k50**		**jk1**	**brr4_k10**
	0.79164	0.62968	0.49819	0.42899
332312	**brr2_k0 brr3_k0**		**jk2**	**boot**
	0.36965	0.29545	0.27876	0.25618

333120	**brr2_k0 brr3_k50**		**brr2_k0**	**brr3_k10**
	0.42911	0.31441	0.23211	0.17112
339920	**jk1 brr3_k0**		**boot**	**brr3_k0**
	0.45867	0.38107	0.30713	0.23567
484230	**jk1 boot**		**jk1**	**boot**
	0.94175	0.68859	1.02299	0.75851
511130	**jk1 brr3_k10**		**jk1**	**boot**
	0.47123	0.32079	0.28481	0.18269
623110	**jk1 boot**		**boot**	**brr3_k0**
	0.29396	0.27258	0.27375	0.2504

Table 8.31 Median Relative Stability Ranges of Variance Estimators

8.32 Median Relative Stability Comparison Measures

A rating method similar to the one used for the median relative bias values was used to determine the variance estimator that most frequently has the smallest median relative stability, or, in other words, the smallest median relative variance of the variance. Considering the results of all of the industries combined, the bootstrap method has the smallest median relative stability most frequently for the monthly estimates, with BRR3_k0 and BRR3_k50 occurring next most frequently. However, for the yearly estimates, two estimators appear with about equal frequency in the sets of cells considered to have the smallest median relative stability – bootstrap and BRR3_k0.

8.4 Confidence Interval Coverage Rates of the Variance Estimators

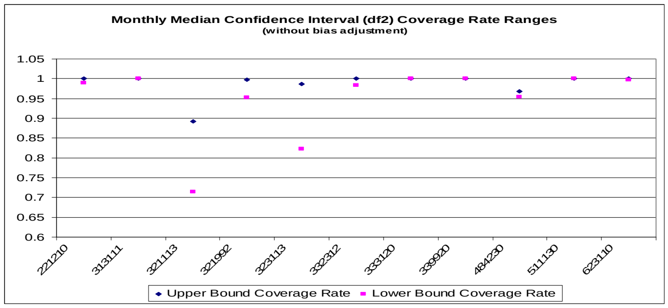
8.41 Median Confidence Interval Coverage Rates for 6-digit level cells

The confidence interval coverage rate of each variance estimator was calculated for each time period. The median coverage rate of all the time periods for each variance estimator is used to compare the variance estimators within a cell and to compare the coverage rates across the industries and is shown in Graphs 8.41A and 8.41B below. All discussion refers to the 6-digit level cells unless stated otherwise.

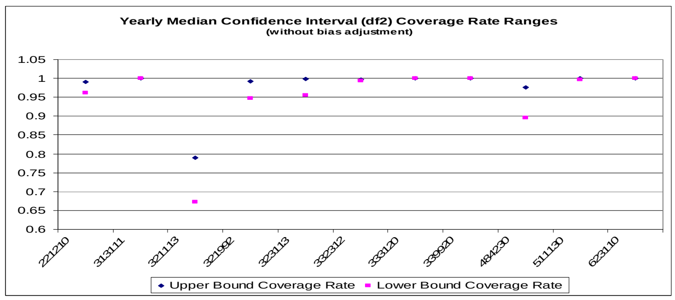
The two graphs below show the median coverage rates for the 1-month and 12-month percentage change estimates for all of the industries. For each industry, the range of median coverage rates of the variance estimators is indicated on the graphs by the Upper Bound Coverage Rate and Lower Bound Coverage Rate plotted values. For most of the industries the median coverage rates for all of the variance estimators are high – greater than 0.95 or 95%. For both the monthly and yearly estimates, 9 of 11 industries have coverage rates of greater than 95% for all of the variance estimators. For the monthly estimates, the midpoint of the range of coverage rates for the variance estimators for 321113 is 80%, and for 323113 is 90%.

The rest of the industries have midpoints of 96% or higher. For 4 of these industries all of the variance estimators have a coverage rate of 100%. The coverage rates for the yearly estimates are similar. Two of the industries, 321113 and 484230, have midpoints of coverage rate range of variance estimators of 73% and 94%, respectively.

Industries with the lowest coverage rates for the variance estimators are also the ones with the most unstable variance estimators. These industries also have variance estimators with negative median relative bias. For the monthly estimates, these industries are 321113, 321992, and 323113. For the yearly estimates, these industries are 321113 and 484230.



Graph 8.41A Monthly median CI coverage rate ranges (calculated with degrees of freedom df2 and without bias adjustment)



Graph 8.41B Yearly median CI coverage rate ranges (calculated with degrees of freedom df2 and without bias adjustment)

8.5 Median Confidence Interval Coverage Rates (using df2=#PSUs - #VS and without bias adjustment)

8.51 Issue of Magnitude of Coverage Rates

Coverage rates are expected to be around 95%. Many of ours were greater than 95%, which was felt to be too large and may be due to the large positive bias of the variance estimators in several of the industries. We recalculated the coverage rates using a bias-adjusted variance to form the confidence intervals. This led to coverage rates that were thought to be too low; however, the range of coverage rates of the variance estimators was larger, and we were able to compare the coverage rates of the variance estimators.

8.52 Median Coverage Rate Comparison Measures

The BRR4_k50 estimator has the largest coverage rate most frequently in both the monthly and yearly estimates. The JK2 and bootstrap methods have the next most frequent appearances of the largest coverage rate in a cell, with JK2 more frequent than bootstrap in the monthly estimates and bootstrap more frequent than JK2 in the yearly estimates. These comparisons of the variance estimators are made with the coverage rates based on confidence intervals calculated using the bias-adjusted variance estimates and degrees of freedom df2.

9. Recommendation of Variance Estimation Method

Because no single variance estimator was superior to all the others based on the comparative measures of relative bias, relative stability, and confidence interval coverage rates, we are recommending the bootstrap method.

10. Suggestions for Further Study

The methodology for calculating percentage change variances at levels above the 6-digit industry level should be tested. Also, calculating bootstrap variance estimates for commodity aggregation structures, and other alternative aggregations should be investigated. Finally, additional research may be required to investigate the cause of the positive bias observed in the results. One potential reason for the overestimation observed in the study is the inconsistency between simulation-sample generation and replicates formation. More specifically, there was no sampling variability involved for certainty establishments in the simulation sample. However, in forming replicates from the simulation sample, secondary sampling units were sampled for certainty establishments and, therefore, sampling variability was added. This inconsistency could have led to overestimation of variance, and overestimation could have been severe if proportions of certainty establishments were high in given industries. Tests may need to be conducted on a diverse range of industries to determine the effects of differing approaches to generating simulation samples and forming replicates in the study. In particular, the handling of certainty establishments in the simulation samples should be revisited.

References

Bobbitt, Patrick A., Chen, Te-Ching., M.J. Cho, Ernst, Lawrence R., Himelein Jr, James A., and Paben, Steven P., *International Price Program (IPP) Variance Project*. Rep. April 18, 2008. Print.

Chipperfield, James, and John Preston. *Efficient bootstrap for business surveys*. Publication. No.2 ed. Vol. 33. Statistics Canada, 2007. Print.

Girard, Claude. *How to avoid getting all tied up bootstrapping a survey: A walk-through featuring the National Longitudinal Survey of Children and Youth*. Working paper. 2007th ed. Vol. HSMD. Household Survey Methods Division, Statistics Canada, 2007. Print. 001E.

Judkins, David R. (1990). *Fay's Method for Variance Estimation.* Journal of Official Statistics, Vol. 6, No. 3, 223 – 239.

Lahiri, P. (2003). *On the Impact of Bootstrap in Survey Sampling and Small-Area eatimation.* Statistical Science, Vol.83, No.401, 231 – 241.

Rao, J.N.K. and Wu, C.F.J. (1988). *Resampling Inference with Complex Survey Data*. Journal of the American Statistical Association, Vol. 83, No. 401, 231 – 241.

Sloane, N.J.A. "Hadamard Matrices." AT&T Labs Research. Web. 10 Sept. 2009. <http://research.att.com/~njas/hadamard>.

U.S.A. Department of Labor. Bureau of Labor Statistics. *BLS Handbook of Methods*. By BLS. 2004. BLS. Web. <http://www.bls.gov/opub/hom/>.

WesVar. *WesVar Complex Samples 3.0 User's Guide*. WesVar, 1998. Print.

Wolter, Kirk. *Introduction to Variance Estimation (Statistics for Social and Behavioral Sciences)*. 2nd ed. New York: Springer, 2007. Print.