# Variance Estimates for Price Changes in the Import and Export Price Indexes January-December 2013

Helen McCulley and David Mead

Each month, the International Price Program (IPP) at the Bureau of Labor Statistics collects a sample of approximately 16,000 import and 10,000 export price quotes from roughly 5,000 establishments. The number of prices from imported items is larger because of the larger volume of imports versus exports. The sample of import and export transaction prices is drawn from the complete universe of all items traded during a calendar year as measured by the Census Bureau. Sampling variability results whenever a sample is used rather than the complete universe.

## Variance Results

The percent changes for overall import and overall export prices along with the standard error for those estimates was calculated for each of the 12 months running from January 2013 to December 2013. Tables 1 and 2 display the estimated percent change for each index, as well as the standard error of that estimate, defined as the square root of the variance. The values of the estimated percent changes show the average percent changes on a monthly, quarterly, and annual basis.

**Table 1. Percent change and standard error for the Import Price Indexes, for one, three, and 12-month intervals, January-December 2013**

| Month | All Commodities Import Price Index | | | | | |
|---|---|---|---|---|---|---|
| | One-month | | Three-month | | 12-month | |
| | Percent Change | Standard Error | Percent Change | Standard Error | Percent Change | Standard Error |
| January | 0.5 | 0.05 | -0.8 | 0.06 | -1.5 | 0.09 |
| February | 0.9 | 0.03 | 0.8 | 0.05 | -0.6 | 0.05 |
| March | -0.1 | 0.04 | 1.3 | 0.07 | -2.1 | 0.05 |
| April | -0.7 | 0.04 | 0.1 | 0.08 | -2.7 | 0.07 |
| May | -0.6 | 0.04 | -1.3 | 0.05 | -1.8 | 0.06 |
| June | -0.4 | 0.03 | -1.7 | 0.05 | 0.1 | 0.04 |
| July | 0.1 | 0.05 | -0.9 | 0.06 | 0.9 | 0.09 |
| August | 0.4 | 0.03 | 0.0 | 0.05 | 0.0 | 0.07 |
| September | 0.3 | 0.03 | 0.7 | 0.04 | -0.7 | 0.06 |
| October | -0.6 | 0.04 | 0.0 | 0.07 | -1.6 | 0.07 |
| November | -0.9 | 0.03 | -1.2 | 0.05 | -1.8 | 0.06 |
| December | 0.1 | 0.07 | -1.4 | 0.09 | -1.1 | 0.11 |

**Table 2. Percent change and standard error for the Export Price Indexes, for one, three, and 12-month intervals, January-December 2013**

| Month | All Commodities Export Price Index | | | | | |
|---|---|---|---|---|---|---|
| | One-month | | Three-month | | 12-month | |
| | Percent Change | Standard Error | Percent Change | Standard Error | Percent Change | Standard Error |
| January | 0.4 | 0.05 | -0.4 | 0.08 | 1.2 | 0.12 |
| February | 0.7 | 0.04 | 1.0 | 0.09 | 1.5 | 0.06 |
| March | -0.5 | 0.04 | 0.6 | 0.05 | 0.2 | 0.05 |
| April | -0.6 | 0.03 | -0.4 | 0.07 | -0.8 | 0.05 |
| May | -0.5 | 0.03 | -1.6 | 0.05 | -0.8 | 0.05 |
| June | -0.1 | 0.03 | -1.2 | 0.05 | 0.8 | 0.08 |
| July | -0.2 | 0.07 | -0.7 | 0.09 | 0.3 | 0.09 |
| August | -0.5 | 0.03 | -0.8 | 0.04 | -1.1 | 0.06 |
| September | 0.4 | 0.13 | -0.3 | 0.13 | -1.6 | 0.13 |
| October | -0.6 | 0.04 | -0.8 | 0.11 | -2.2 | 0.08 |
| November | 0.2 | 0.02 | -0.1 | 0.05 | -1.5 | 0.09 |
| December | 0.4 | 0.03 | -0.1 | 0.15 | -1.0 | 0.06 |

The standard error, the square root of the estimated variance, is a common measure used to derive confidence intervals for percent changes in the import and export price indexes. Confidence intervals can be used to determine if an index change is significantly different than zero.

Take as an example the 1-month percent change of 0.7 percent for export prices in February, as seen in Table 2. The standard error for exports in February was 0.04. So, deriving a confidence interval plus or minus two standard errors from the point estimate, assuming a normal distribution, the true change would most likely be 0.7 percent plus or minus 0.08, or between 0.62 percent and 0.78 percent.

**Analyzing the Data**

One result from the data is that the *relative* standard error, defined as standard error divided by the percent change, generally decreases the longer the period in question. For example, looking at the index for overall import prices as seen in Table 1, for the 1-month percent change in March, the standard error divided by the percent change (0.04/-0.1) is -0.40. For the 3-month and 12-month percent changes, the relative standard errors are 0.05 and -0.02, respectively, meaning that relative to shorter intervals, the precision of the estimated percent change improves as the calculation interval lengthens.

**Sources of Error**

There are different types of errors that are introduced when calculating the estimates of average price changes for imports and exports published by the Bureau of Labor Statistics. One way to look at measurements of error is the difference between sampling error and nonsampling error. Sampling error, what is reported on in the tables in this report, is the error resulting from drawing a *sample* of imported and exported items to and from the United States, rather than using the entire universe of trade.

Nonsampling error can take a number of different forms. One form is misspecification error, which takes place if the universe of data from which the sample is being drawn does not correctly measure the actual population. For import and export prices indexes, this type of error could result if there are mistakes in the trade dollar value statistics measured by the Census Bureau. A second type of nonsampling error is nonresponse error. Each month, a subset of the items sampled by the Bureau of Labor Statistics does not have prices reported. This type of error results if the respondents and nonrespondents do not represent a similar cross section of the total universe. Another form of nonsampling error can be introduced from misreported prices, and is possible regardless of whether deriving the indexes from a sample or from the complete universe of data.

One issue when deriving an estimate from a sample is the potential trade-off between variance and bias. Variance is a measure of how much the estimates derived from numerous samples differ from the true value of the estimate. Bias results if the expected value of the estimate is either higher or lower than the true value of what is being estimated. An estimate could have a high sampling variance and still be unbiased if the expected value of the estimate is equal to the true value. Likewise, an estimate may have small variation over numerous samples, yet be biased if the expected value of the estimate deviates from the true estimate.

The Bureau of Labor Statistics strives to minimize both sampling and nonsampling error as much as possible. Sampling error is reduced by maintaining as many prices as possible to support an index given resource and company burden constraints. Nonsampling error is reduced by subjecting the data to careful review using automated checks and a staff of professional economists, as well as by employing methods to estimate missing observations.

**Sampling in the International Price Program**

Trade into and out of the United States is highly regulated and therefore highly documented. This allows the IPP to sample from a fairly complete and detailed frame.

U.S. Customs and Border Protection provides the sampling frame for import merchandise, while the export merchandise frame is created from a combination of data collected by the Canada Border Services Agency, for exports to Canada, and from the Census Bureau for exports to the rest of the world.

The import and export merchandise universes are divided into two panels each, with each product-based panel representing approximately half of the trade dollar value for its respective universe. One import and one export panel is sampled each year, resulting in a fully sampled universe every 2 years. Each panel is sampled using a 3-stage sample design. The first stage independently selects establishments within product-based sampling strata using systematic probability proportional to size, where the measure of size is the total trade dollar value for establishments within the sampling stratum. The second stage selects more highly detailed product categories, known as classification groups, within each establishment stratum combination selected during the first stage using the technique of systematic probability proportional to size with replacement. The final stage of sampling, which results in a unique item to price, occurs in the field and is a random selection technique with the probability of selection proportionate to field collected trade estimates.

**Index Calculations**

IPP calculates its indexes using a modified Laspeyres formula. Rather than calculate indexes relative to a base period, the IPP indexes are calculated relative to the previous period and are theoretically chained to the reweight period. For this reason, IPP's indexes are often referred to as chained Laspeyres indexes. Explicitly, the index formula for the modified chained Laspeyres is derived from the classic formula as follows:

$$LTR_t = \left(\frac{\sum_i p_{i,t} q_{i,0}}{\sum_i p_{i,0} q_{i,0}}\right)(100) = \left(\frac{\sum_i \frac{p_{i,t}}{p_{i,0}} p_{i,0} q_{i,0}}{\sum_i p_{i,0} q_{i,0}}\right)(100) = \left(\frac{\sum_i r_{i,t} w_{i,0}}{\sum_i w_{i,0}}\right)(100) =$$

$$\left(\frac{\sum_i r_{i,t} w_{i,0}}{\sum_i r_{i,t-1} w_{i,0}}\right)\left(\frac{\sum_i r_{i,t-1} w_{i,0}}{\sum_i w_{i,0}}\right)(100) = \left(\frac{\sum_i r_{i,t} w_{i,0}}{\sum_i r_{i,t-1} w_{i,0}}\right)(LTR_{t-1}) = (STR_t)(LTR_{t-1}),$$

where

$LTR_t$ = the long term relative of a collection of items at time $t$;

$p_{i,t}$ = price of item $i$ at time $t$;

$q_{i,0}$ = quantity of item $i$ in base period 0;

$w_{i,0}$ = $p_{i,0}\, q_{i,0}$, or the total revenue generated by item $i$ in base period 0;

$r_{i,t}$ = $p_{i,t} / p_{i,0}$, or the long term relative of item $i$ in period $t$; and

$STR_t = \frac{\sum_i r_{i,t} w_{i,0}}{\sum_i r_{i,t-1} w_{i,0}}$, or the short term relative of a collection of items at time $t$.

Depending on the level of aggregation, the weights used during index aggregation are either trade dollar based or probability based. At the lowest level of aggregation, items are weighted by probability-based weights calculated monthly corresponding to detailed categories within establishments.

These weighted item price relatives are combined across establishments and aggregated to the lowest level stratum indexes as

$$P_{h,t} = \frac{\sum_k \sum_j \sum_i w_{k,t} w_{j,t} w_{i,t} \left(\frac{p_{i,t}}{p_{i,0}}\right)}{\sum_k \sum_j \sum_i w_{k,t} w_{j,t} w_{i,t}},$$

where

$P_{h,t}$ = the price index for lowest level stratum $h$, at time $t$;

$w_{k,t}$ = the weight of detailed product category $k$, within stratum $h$;

$w_{j,t}$ = the weight of establishment $j$, within detailed product category $k$;

$w_{i,t}$ = the weight of item $i$, within establishment $j$ and detailed product category $k$; and

$\frac{p_{i,t}}{p_{i,0}}$ = the price relative of item $i$, from period t, to base period 0.

The weights used for these lowest level stratum indexes are derived from sampling frame trade dollar values, divided by the corresponding probabilities of selection determined by the sample design.

At the next level of aggregation, child strata level indexes are aggregated to their corresponding parent stratum level indexes. A child stratum index is simply one level of aggregation less than its parent stratum index. The weights used for this aggregation are based on Census Bureau trade dollar values for the base period. The aggregation formula for these upper index levels is

$$P_{H,t} = \frac{\sum_h w_{h,t} P_{h,t}}{\sum_h w_{h,t}},$$

where

$P_{H,t}$ = the price index at period *t*, for upper level index *H*;

$w_{h,t}$ = the weight at period *t*, for child index *h*; and

$P_{h,t}$ = the price index at period *t*, for child index *h*.

**Replication and Variance Estimation**

A modified bootstrap method, applying rescaled sampling weights, is used to produce 150 replicate index set estimates from 150 simulated item set samples. Item set replicates are constructed according to IPP's 3-stage sample design. At both of the first two stages of sampling, it is possible for a selection to be either a certainty selection (i.e. the probability of selection is greater than the iteratively calculated sampling interval) or a probability selection. The replicate resampling method takes this into consideration by first partitioning the selected items within each sampling stratum *m* into those items that resulted from certainty establishment selections and those items resulting from probability establishment selections. The item set resulting from establishment certainty selections is further partitioned into two item sets: sampling classification group certainty selections and sampling classification group probability selections. Thus, the set of all sampled items *S* is the union of these three partitions over all sampling stratum *m*;

$$S = \bigcup_{m=1}^{N} S_m = \bigcup_{m=1}^{N} \left( \bigcup_{p=1}^{3} S_{m_p} \right)$$

where *N* is the number of sampling strata, $p \in \{1,2,3\}$, with *p*=1 for items selected from probability establishments, *p*=2 for items selected from probability sampling classification groups within certainty establishments, and *p*=3 for items selected from certainty sampling classification groups within certainty establishments.

Each bootstrap sampling, *b*, selects $n_{m_p}^b$ units within each partition of each sampling stratum as follows:

$$n_{m_p}^b = \begin{Bmatrix} n_{m_p} - 1 & n_{m_p} > 1 \\ 1 & n_{m_p} = 1 \end{Bmatrix},$$

where $n_{m_p}$ is the number of units originally sampled in partition *p* of sampling stratum *m*.

Bootstrap item weights are then calculated as

$$w_{m_p,j,i}^b = \begin{Bmatrix} w_{m_p,j,i} \left( \frac{n_{m_p}^b + 1}{n_{m_p}^b} \right) d_{m_p,j}^b & \text{for } n_{m_p} > 1 \\ w_{m_p,j,i} & \text{for } n_{m_p} = 1 \end{Bmatrix}$$

where

> $w_{m_p,j,i}^b$ = the $b^{\text{th}}$ replicate item weight for item *i*, within establishment *j* and sampling stratum partition $m_p$;
>
> $w_{m_p,j,i}$ = the standard item weight for item *i*, within establishment *j* and sampling stratum partition $m_p$; and
>
> $d_{m_p,j}^b$ = the number of times establishment *j*, within partition *p* of sampling stratum *m*, is selected in bootstrap sample *b*.

In the rare instances that $n_{m_p} = 1$, a simple random sample of items within that establishment is selected. If only one item exists under this establishment singleton, that item is chosen with certainty.

For each of the 150 bootstrap samples, chained indexes of the desired length are calculated at all levels of aggregation using these modified item weights, original probabilities of selection, trade dollar values, and collected price data. For variance estimates, the variance is calculated across replicate percent change values for all published indexes as $\nu_B = \frac{1}{150}\sum_{b=1}^{150}\left(\hat{\theta}_b - \hat{\theta}\right)^2$ where $\hat{\theta}$ is the full sample estimate.