Using Administrative Data to Calculate Export Price Indexes

Authors:
Don Fast (fast.don@bls.gov) and Susan Fleck, Ph.D. (Fleck.Susan@bls.gov)
International Price Program, Office of Prices and Living Conditions
U.S. Bureau of Labor Statistics
2 Massachusetts Ave., NE, Rm. 3955
Washington, DC 20212

Discussant:
Katheryn Russ, Ph.D. (knruss@ucdavis.edu)
University of California, Davis
Session: Improving Economic Price Statistics through the Use of Alternative Data
Session sponsored by AEA Committee on Economics Statistics
Presented at the ASSA Annual Meetings, San Diego, CA, January 5, 2020

Abstract: The BLS currently uses the administrative dataset of international trade transactions of merchandise goods as the sample frame to select representative products and companies to collect a basket of approximately 22,000 representative goods that form the basis of published Import and Export Price Indexes, Principal Federal Economic Indicators. Previous research using the administrative trade data established a prototype unit value index for homogeneous products. This research project expands the analysis to create historical monthly time series for 2012-2017 of all 127 5-digit BEA End Use monthly export unit value indexes, based on 20+ million trade records. The unit value indexes are evaluated for homogeneity and comparability to official price indexes. Among the 127 indexes, 24 unit value indexes covering 23% of export trade are deemed high quality. The impact of the historical data on the real value of exports from 2012 to 2017 is estimated. If used as deflators, real value of goods exports comprising these 24 product areas would have declined .75% more per year, and the impact on real value of all-goods exports would have been -.15 % yearly.

JEL codes: E3, E31, F0, F1, F3
Keywords: price indexes, exports, trade, measurement

We recognize and thank these BLS staff for their research and data support for this paper: Michael Havlin, Laurence Lang, Christina Qiu, Daryl Slusher, and Dominic Smith. Additional support was received from Antonio Caraballo, Brett Edgerton, Ara Khatchadourian, Aric Schneider, and Praveenkumar Yerramareddy.

Use of the export trade data are subject to Agreement No. 2067-2018-001, Memorandum of Understanding (MOU) between the U.S. Census Bureau and the Bureau of Labor Statistics (BLS). The BLS has received prior approval from the U.S. Census Bureau, which affirms that the research results do not present disclosure risks and approve the publication of the results.

Calculations are subject to revision.

Introduction

The US Bureau of Labor Statistics (BLS) International Price Program (IPP) mission is to measure price changes for US imports and exports with the official Import and Export Price Indexes (MXPI), which is a goods-only index. Presently IPP uses an establishment survey to track unique-item price changes over time to construct monthly price indexes. Work by academic economists[1,2] provides evidence that big data sources such as scanner data and the Billion Prices Project provide opportunities to track average prices of groups of similar goods and calculate price indexes. This paper analyzes whether unit value indexes – based on average prices extracted from US administrative trade data – can be used on a large scale to track price change; if this research provides promising results, it could bolster the number and improve the quality of BLS official price indexes at the detailed and, by extension, all-goods levels. This paper analyzes export trade data for 2012 to 2017 and compares unit value indexes created from the trade data with official export price indexes for 5-digit BEA End Use product categories.

Official price indexes are the standard against which other price indexes are measured, and they must withstand greater scrutiny to maintain confidentiality and provide timely, accurate, reliable, and transparent data. With this public obligation in mind, the IPP has been carrying out research and sharing results to gain feedback on the potential of using the complete dataset of U.S. trade transactions to calculate the goods' MXPI. The administrative trade transaction dataset is the universe from which the MXPI sample is selected on a rolling two year basis. Lack of timely receipt and unit value bias have been the two limiting factors to using the administrative trade data directly in price indexes. However, the wealth of data and the ability to maintain confidentiality have been the impetus for developing alternative approaches to directly calculate price indexes with the US Customs data[3]. The MXPI survey program faces resource constraints and falling response rates, and the research results described here - using monthly US Customs transaction data from 2012 to 2017 – will be used to evaluate whether this dataset is a viable alternative to expand the number of prices and items available for calculating current price indexes.

Most academic research using the US Customs administrative trade data implicitly make the strong assumption that unit value indexes calculated at the detailed product category level are not biased. However, this assumption conflicts with the consensus view that unit values are biased measures of heterogeneous or differentiated product prices. The results of this paper provide strong evidence that there are real and measurable differences in the homogeneity of the trade data, and that the existence of unit value bias requires product categories to be homogeneous if they are to track market price

[1] Feenstra, Robert C., and Matthew D. Shapiro, eds. Scanner data and price indexes. Vol. 64. University of Chicago Press, 2007.

[2] Harchaoui, Tarek M., and Robert V. Janssen. "How can big data enhance the timeliness of official statistics?: The case of the US consumer price index." International Journal of Forecasting 34, no. 2 (2018): 225-234.

[3] For all country destinations but Canada, U.S. export merchandise trade transactions are registered by the Shipper's Export Declaration of each harmonized system (HS) category per shipment and recorded by the Census Bureau through the Automated Export System (AES) web portal; records of U.S. exports to Canada are recorded through Canadian import records transmitted to the U.S. Census. The dataset accurately records total trade dollar value information for each shipment, and each record describes the shipping characteristics all goods at the 10-digit HS product level for each company in each shipment. U.S. Census uses the data for international trade measures and subsequently provides it to BLS. To construct a price from the transaction trade value, this number must be divided by the provided quantity. Some quantity data are missing and others are imputed by the U.S. Census.

movements. The need to establish rigor to define homogeneity and to validate a way to test unit value indexes for unit value bias and goodness of fit are the drivers of the research presented here. BLS research into unit values and unit value indexes is intended to evaluate their usability as inputs into official import and export price indexes. In Fast and Fleck[4] we develop a proof of concept that evaluates homogeneity and establishes the least biased approach for calculating unit values and short term ratios (STRs) at the 10-digit HS product classification strata, and then aggregate price changes using HS-BEA End Use concordance for two detailed export product categories – dairy and vegetables – for 2015-16.

In this paper we expand this work and evaluate all 127 export 5-digit BEA end-use indexes (the detailed indexes used to deflate net trade in the GDP) for 2012 to 2017, using more than 200 million transaction records to implement the proof of concept described in Fast and Fleck. The characteristics of this research database, the methodology used to calculate unit value indexes, and the evaluation of index quality compared to official price indexes are described in this paper. We address three issues that sort the price indexes into "good", "undecided" and "poor" bins. The first is the known unit value bias in trade statistics. The second is the statistical comparison of the long term and monthly differences between the unit value indexes with their corresponding official BLS XPI. The third is the potential impact of these changes on trends and on deflation if they were to replace the official price indexes. There are also operational issues of calculating timely measures which are not addressed in this paper.

Judging the quality of the unit value index from these three criteria, we identify 24 "Good" indexes that are homogeneous, exhibit little unit value bias and have similarities with their comparable official XPI. We then estimate the impact of replacing the "Good" administrative trade price indexes in the top level price index. These 24 unit value indexes cover 23 percent of export trade and together resulted in .75 percent higher prices annually from 2012 to 2017, when compared to the price change tracked by the same 24 official indexes. When weighted with the official all-goods XPI, the 24 product categories contributed to .14 percent higher export prices annually.

From Proof of Concept to Road Map

Fast and Fleck 2019 established a proof of concept after analyzing 36 options to 1)select the records needed to calculate unit values for entry level items (ELI), 2)group the records for maximum homogeneity, 3)address outliers, 4)establish imputation rules for months an ELI is not traded, 5)select index method and weights to aggregate unit values to the lower level strata – the HS classification groups. These options were evaluated for their homogeneity and goodness of fit against official BLS price indexes and the approach that was deemed the best fit was adopted for the analysis of the 127 5-digit BEA End Use price indexes.

To recap the choices for the proof of concept, the issues are described below.

Unit values and grouping records. The average transaction price of a shipment record is a blunt measure of prices. All shipment records are based on the HS product classification and the employer ID number. Other shipping information is available. Various groupings of shared characteristics – which we call item keys - were evaluated and the most detailed grouping of characteristics that track the price survey was found to have the least price variation, at least initially. Given our assumption that substitutability is maximized when price variation is minimized, we use the most detailed grouping to begin the analysis.

[4] Fast, Don, and Susan Fleck. "Unit Values for Import and Export Price Indexes – A Proof of Concept." NBER, 17 Oct. 2019, www.nber.org/papers/w26373.

Addressing outliers and establishing imputation. The sheer number of records results in price outliers that need to be effectively limited on a systemic basis. Outliers may be a result of data entry error, compositional abnormalities or one-time shipments, and one outlier can greatly skew both one-month and long term movements in the index. A trimmed mean (0.5%) of the entry level item prices eliminates outliers but does not dampen actual market price variation. Imputation of ELI unit values using the cell mean are calculated for up to three months to provide continuity to the item, in case it is not traded for a couple of months. Furthermore, the ELI is calculated as a weighted geomean average price to mitigate the impact of outliers on the ELI unit value.

Select index method and weights. Total trade value for each transaction record provides the opportunity to use the Tornquist superlative index formula to aggregate ELIs to the lower level strata, the HS classification group. Not only does the Tornquist index address formula bias, it corrects for substitution bias and new goods bias by accounting for changes in both price and weight, or share of exports.

Starting with this proof of concept, we set out to create a road map to evaluate the quality of 127 5-digit BEA End Use export unit value indexes based on more than 5,000 HS classification groups for more than 200 million transaction records from 2012 to 2017. First, geomean unit values are calculated from detailed item keys to create ELIs for each month, then ELIs' month-to-month change is used in a Tornquist index formula to calculate the price change for the 10-digit HS product classification groups, or lower level strata. From the lower level strata, the HS-to-BEA End Use concordance is applied to match the HS classification groups to calculate the BEA End Use unit value indexes using the official MXPI methodology for the upper level calculation - a modified Laspeyres index that uses two year lagged weights. Any newly introduced HS classification is backward-linked to its previous classification in order to assure that the lagged weights do not undercount new items.

For exports, these 127 detailed unit value indexes are split into homogeneous and heterogeneous groups using a test for price variability. The motivation for this test is described below. This test provides strong evidence that price indexes trend differently depending on their homogeneity.

Unit Value Bias, Price Variability, and the Coefficient of Variation

Whether unit values can be used in price indexes – and if so, which ones – is the first and arguably most important conceptual concern when considering blending the values into official price statistics. The BLS price indexes are based on the matched model concept, which tracks transaction price movements of unique items. The MXPI concepts and methods were established in direct response to public criticism of official unit value indexes in the 1960s. Because a unit value is an average price based on shipping transaction records, the price is not likely to represent one unique item, thus conflating changes in product mix with changes in market price. Unit value bias is more prevalent for heterogeneous goods, and there is general consensus that homogeneous product categories can produce viable unit value indexes (see Silver[5]).

Interestingly, many researchers use the trade transaction records to create price indexes but they do not differentiate between heterogeneous and homogeneous goods. This lack of concern for unit value bias is curious given the criticism that official price index measures receive for insufficient quality

[5] Silver, Mick. "The wrongs and rights of unit value indices." Review of Income and Wealth 56 (2010): S206-S223.

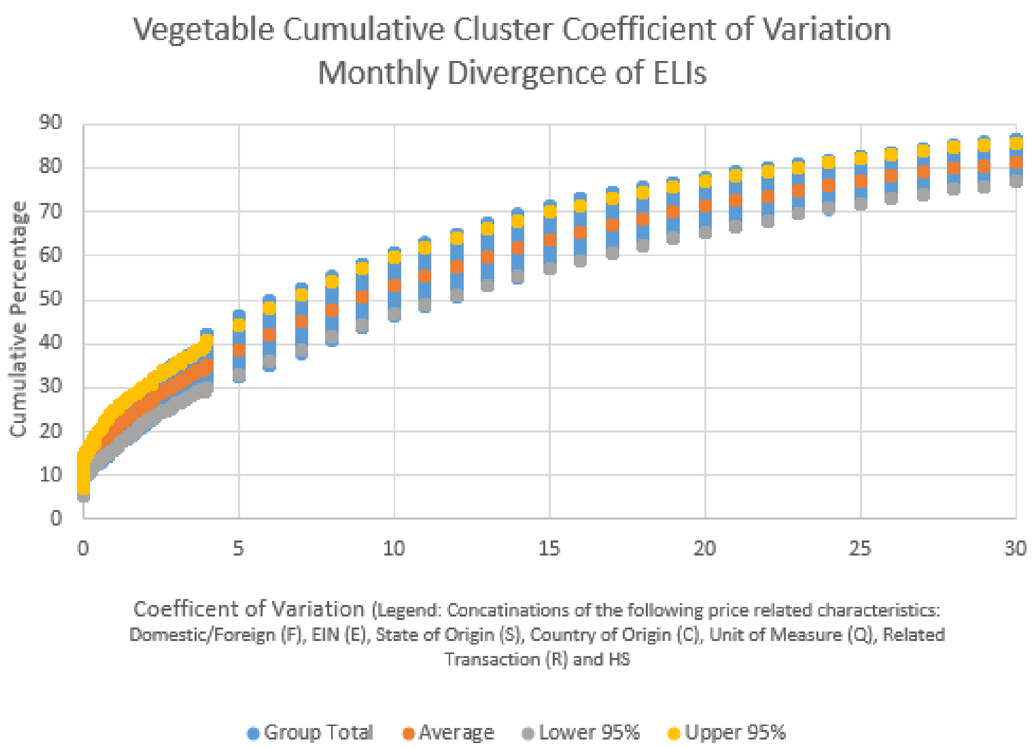
adjustment of heterogeneous goods, particularly of advanced technology products. Unit value indexes of heterogeneous goods product categories are likely to experience upward drift, given that price increases are not adjusted for changes on product attributes. For official statistical purposes, however, we must consider the potential impact of unit value bias if we are to use the trade transaction records.

Because there is no agreed-upon measurable concept of homogeneity, we develop one. We base our general concept in price-setting theory. We assume that homogeneous unit values track close substitutes over time, and thus there should be little variability of prices of similar goods. At the same time, it is known that some homogeneous goods face fluctuations in prices, possibly due to seasonality and global demand and supply factors (Gopinath and Itskhoki (2010) and Gopinath and Rigobon (2008)). We group transactions in such a way as to minimize compositional effects and to maximize substitutability. Matching transactions to a greater level of specificity than the 10-digit HS product categories takes into account price and non-price trade characteristics that separate goods into unique bins of substitutable items. This approach aligns with work by Rauch (2001) and Clausing (2003) that identifies the importance of non-price characteristics on price-setting in international trade. Given the high frequency of transactions in trade data, each grouping of like characteristics is likely to have more than one transaction – which could add to price variability; but price dispersion between months of similar transactions would be dampened the more similar the groupings. The focus on what we call intra-item substitutability results in our enforcing stricter category inclusion criteria.

The necessary but not sufficient condition for BLS to integrate unit values into price indexes is to gauge unit value bias. Fast and Fleck (2019) identified the coefficient of variation test as the most relevant test for evaluating price variability and unit value bias. The presence of price variability does not prove unit value bias, but significant price variability of entry level items (which are constructed average prices of transaction records that share select characteristics) may be the result of unit value bias. In Fast and Fleck, we chose the threshold of price dispersion as the average price dispersion of the ELIs of the Vegetable BEA End Use Export unit value index. Why vegetables? We take a reasonable judgmental approach – vegetable products are both considered to be relatively homogeneous and yet subject to price variation due to seasonality. There are 161 HS product classifications that comprise the Vegetable BEA End Use Export product category. Each classification is internally homogeneous although in aggregate there is much variability. Vegetable prices also vary significantly throughout the year, subject to domestic and world seasonal and weather conditions. Thus the Vegetable unit value index, like its comparable official price index, is on the high end of price variability among homogeneous product classifications.

By using a coefficient of variation, and by maintaining a constant, if judgmental, limit in price dispersion, all the price dispersion trends of all other indexes can be compared to it. Those unit value indexes that show less dispersion can be considered more homogeneous than the vegetable unit value index. The threshold is established with the results of the coefficient of variation test. The cumulative coefficient of variation tracks the degree of price dispersion of the short term ratios (STRs), or monthly price changes. A convex trend as seen below shows that the majority of prices are grouped at the lower bound of price change. The x axis tracks the value of the coefficient and the y axis counts the cumulative share of monthly observations at each incremental change in the coefficient of variation. The majority of price dispersion is clustered at the lower bound, with half of all observations holding a value less than 9. (Given the number of item prices that had smaller coefficient of variations (CV), more clusters were created for these lesser CV values. For the larger CV numbers calculations were performed on larger cluster ranges given the dispersion of the data.)

Graph 1



In graph 1 below, all month-to-month changes for the unique detailed ELIs[6] comprising the vegetable unit value index were used to calculate the coefficient of variation simultaneously across 6 years, 2012-17. A benchmark mean and lower bound were defined based on 24 months (2015-16) of clustered Vegetable entry level items' mean and second standard deviation (from Fast and Fleck's previous research). This lower bound forms the basis of comparison; the mean value calculated for the other 126 export 5-digit BEA End Use unit value indexes are compared to the lower bound of the Vegetable coefficient of variation trend line. Similar to the comparison in Fast and Fleck based on two years of prices, 51 of the 126 indexes showed even less price dispersion than Vegetables. We categorize these 52 unit value indexes (including vegetables) into the homogeneous category, i.e. not exhibiting extreme price variability and thus less prone to unit value bias.

Given the judgmental nature of the cutoff point that defines a homogeneous good, there is a possibility that some of the BEA End Use product categories near the mean may not meet the cutoff for homogeneity. Inversely, it is possible that an index may appear homogeneous because the cumulative percentage of price variability is above the Vegetable mean, but it may not accurately track the price index. These impacts at the margin result in ambiguities that we take into consideration when we compare the indexes with their comparable official price index.

---

[6] The entry level items are grouped by matched records of all of the following characteristics: harmonized number, domestic or foreign content, employer ID, state of origin, country of destination, unit of measure, and related company transaction.

Some homogeneous and heterogeneous unit value indexes are clearly delineated. We use the C.V. test as the first but not sole decision point to evaluate unit value bias and goodness of fit. Using the mean of the vegetable C.V. results, there are 52 homogeneous and 74 heterogeneous unit value indexes. The statistical comparisons described in the next section provide additional information to further disaggregate the indexes by goodness of fit measures (which we discuss later in this paper); thus we separate the 127 indexes into three bins, or quality groups – "Good", "Undecided", "Poor". All "Good" indexes are homogeneous and all "Poor" indexes are heterogeneous, but the "Undecided" bin of indexes includes both homogeneous and heterogeneous indexes. These indexes' coefficients of variation range close to the mean of the vegetable coefficient of variation threshold and exhibit some comparability to official price indexes (details described in Benchmark Comparisons). These indexes deserve further investigation.

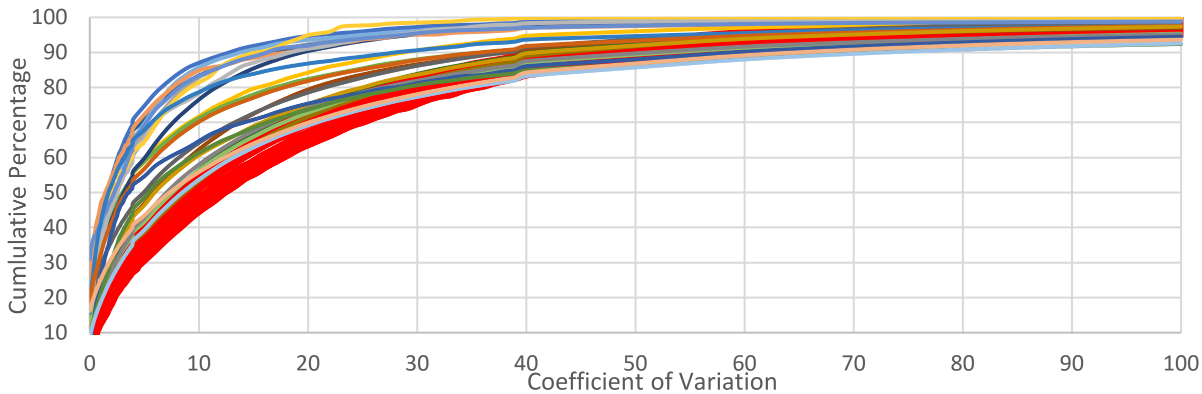
At a glance, one can see the different trends in price variability between these three quality groups in Graphs 2, 3, and 4, in which the mean and lower bound of price dispersion are noted by a thick red line, the mean cumulative percent of the coefficient of variation for Vegetables . For "Good" indexes (graph 2) more than seventy percent of variability as defined by the coefficient of variation is less than 20 and in all cases above both the lower bound and mean of the vegetable C.V. For "Undecided" indexes (graph 3) 11 of the 30 are heterogeneous, and most of them range around the mean. They have significantly greater price variability, such that seventy percent of the coefficients of variation observations range between 25 and 55. These "Undecided" unit value indexes primarily constitute homogeneous "Other NES" and "Manufacturing" heterogeneous product categories. Except for a handful of unit value indexes (which show data anomalies we have not yet corrected), the "Poor" indexes (graph 4) report coefficients of variation for which seventy percent of monthly changes are far below the vegetable lower bound cutoff with C.V.s ranging between 25 and 100.

At a minimum, this approach, if implemented, would allow 23 percent of goods' export prices to be represented by unit value indexes (i.e. "Good" indexes), valued at $328 billion dollars of 2015 trade. If the "Undecided" unit value indexes were to have a better 'fit with a different grouping of like-characteristics, an additional 20% of export trade could use unit value indexes to measure price change. These results also show that the 9 homogeneous and 64 heterogeneous indexes in the "Poor" bin have extreme price variability and unit value bias and/or a poor fit with the official price indexes.

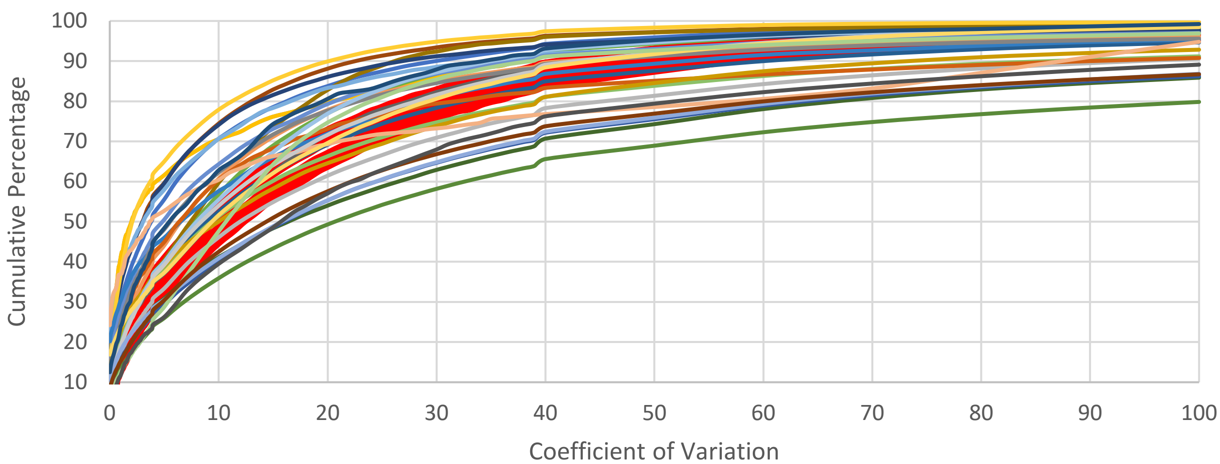
Table 1. Characteristics of 5-digit BEA End Use Export Price Indexes

Homogeneous/Heterogeneous	Index Quality	Number of 5-digit BEA End Use U.V. Indexes	Trade Dollar Value, 2015, In millions	Percent of Indexes	Percent Trade Weight
Homogeneous	Good	24	$328,869	18.9	22.5
Homogeneous	Undecided	19	$150,099	15.0	10.3
Homogeneous	Poor	9	$68,781	7.1	4.7
Heterogeneous	Undecided	11	$136,100	8.7	9.3
Heterogeneous	Poor	64	$777,116	50.4	53.2
	All	127	$1,460,964	100.0	100.0

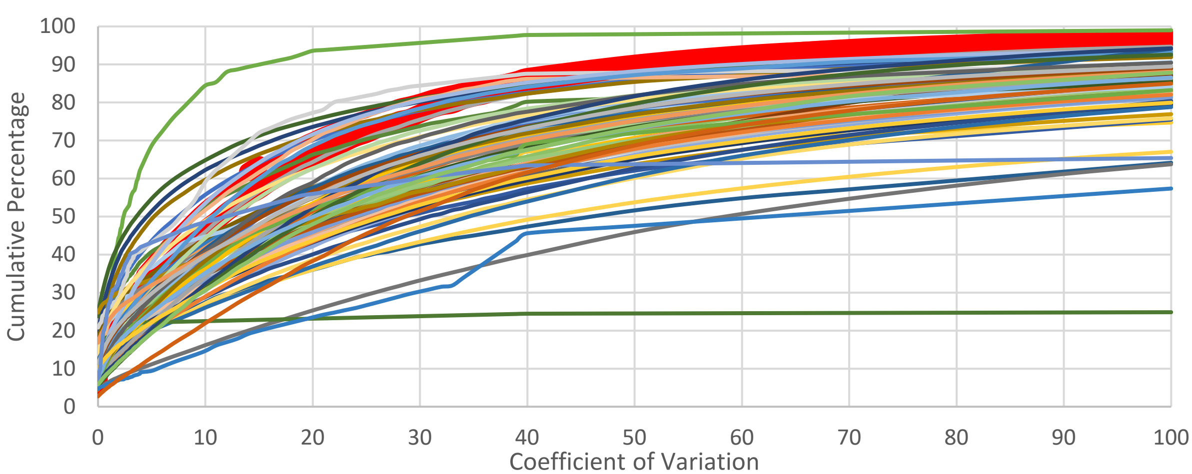
Graph 2. Cumulative Percentage of the Coefficient of Variation for "Good U.V. Indexes, compared to the Vegetable C.V. Mean and Lower Bound (n=24)



Graph 3. Cumulative Percentage of the Coefficient of Variation for "Undecided" U.V. Indexes, compared to the Vegetable C.V. Mean and Lower Bound (n=27)



Graph 4. Cumulative Percentage of the Coefficient of Variation for "Poor" U.V. Indexes, compared to the Vegetable C.V. Mean and Lower Bound (n = 76)



The ability to create ELIs that are similar enough to eliminate or mitigate unit value bias by limiting price dispersion is a new approach that consistently results in clear differences between homogeneous and heterogeneous goods. Note that each index has been calculated with the variables that result in the least intra-item variability, and thus the greatest homogeneity, using an approach in which we select the best combinations of six variables, or data fields, from the complete dataset of administrative trade transactions that are similar to the price determining characteristics of the directly collected survey data. The six variables selected are the same for each index but not necessarily the same across indexes. Not all combinations of variables have been evaluated for their intra-item variability - this research is ongoing.

Benchmark Comparison of Unit Value Indexes with Official BLS Price Indexes

As described above, the determination of unit value bias is not sufficient information to evaluate the impact and the quality on the MXPI. There are four statistical tests that we conduct to evaluate both the monthly change and the trend over time to evaluate data quality, comparing them to the benchmark of official price indexes. The benchmark comparisons are carried out for the short term ratios (STRs) and long term relatives (LTRs). These tests support the condition of intra-item substitutability for each period and over the whole period. The more similar the monthly change (STRs) the more likely the price variability reflects trade of similar items. The more similar the base-period to end-period change (LTRs), the more likely the price trends over time reflect similar goods and a similar mix of goods.

We evaluate all 127 indexes using their short term ratio and long term relative trends in comparison to their corresponding official price indexes (published and unpublished) for the 6-year period. This comparison is based on the reasonable assumption that the items and prices collected for the official price index are representative of trade, and that unit value indexes with little unit value bias are reasonably similar to the official price index. The benchmark price indexes are the official 5-digit BEA End Use export price indexes, of which about half are published. The other half are not published because, although data exist to support the higher level indexes, BLS has concerns about the quality or protection of confidentiality of the data. Thus, not all details for unpublished price indexes are included in this paper.

With the results from the statistical tests, we sort indexes into “Good”, “Undecided”, and “Poor” bins, or quality groups. We find along the way that no index performs well on all tests, and so we establish one additional condition to evaluate the index quality. To compliment these analyses of short term ratios and long term relatives we visually inspect the indexes and evaluate the gap of the indexes at the December 2017 endpoint and throughout the time series. This additional judgmental evaluation is both a check on the basic assumption that the benchmark correctly captures price change, and the evaluation of the unit value time series as a deflator for revisions to the net export measure of Gross Domestic Product. By considering the percentage point difference in the December 2017 index value between unit value and official price indexes, we can whether the two indexes act alike over 6 years. Even though statistical tests may show similarities between indexes, large one-time differences and small but cumulative differences can create a large gap between the unit value and price indexes along the trend line as well as at the end point. Additionally, lags in price changes of one index that parallel the other index are not captured by any of the statistical tests, but may result on a small gap between indexes in December 2017. The indexes that tested well but had large gaps between LTRs in December 2017, and the indexes that did not test well with current-month comparisons but that had small gaps between LTRs were the grey areas that constitute the "Undecided" unit value indexes.

Results of statistical comparisons and quality groups

The comparison of the unit value indexes against a benchmark of official price indexes is the second step in evaluating the quality of unit value indexes. We carry out statistical comparison using four tests: 1)Mean and Standard Deviation of long term relatives; 2) the Root Means Squared Error of short term ratios, 3) the Correlation Coefficient of short term ratios, and 4) Cointegration test of long term relatives.

Each statistical test is briefly described below.

- *Mean and Standard Deviation.* Calculations are made of long term relatives cast backwards and forwards from each month as a point of reference for all months and for both unit value and official price indexes. These figures provide a population from which the mean and standard deviation are calculated and compared for each 5-digit BEA End Use category. Then the mean difference between the unit value and benchmark index is evaluated, to evaluate whether the clustering around the mean is similar between the two indexes.
- *Root Mean Squared Error Comparison*. Root mean squared error measures the accuracy of the indexes based on the population – in this case the unit value indexes – and the sample – the benchmark price indexes. The smaller the first degree difference between like indexes for the same month, across the six years, the lower the RMSE value will be, and the more predictable is the price change for the period covered. However, because the first degree differences are squared, the magnitude may be the same while the direction differs; such differences over time may result in divergent indexes and thus a low error value may still result in a long term trend that is not comparable with the existing official price index. Values less than 4 are used as the cut-off for a 'pass', to contain price variability.
- *Correlation coefficient comparison*. Correlation coefficients assess the degree of predictability of the first degree differences of the unit value index on benchmark price indexes, for each month. The correlation coefficient ranges between 1 and -1, with 1 describing positive correlation and -1 describing negative correlation. Values over .5 are used as the cut-off for a 'pass', i.e. the two indexes' STRs being correlated.
- *Cointegration Comparison.* This test is applied differently from most cointegration tests. Rather that identifying stationarity and correcting for it, the use of the cointegration test in this context is to determine whether there is stationarity, and if there is, we assume that the unit value indexes and official indexes have stable, long-run relationships.[7] Lower P-values suggest that the unit value and official indexes are cointegrated, and therefore follow similar trends over time. This test is not suitable for the heterogeneous indexes, and may describe an opposite effect. That is, truly heterogeneous unit value indexes may be consistently divergent from the benchmark index, and thus the stationarity may show a stable relationship that provides no predictive power of comparability. The cut-off for quality is p-value < 0.05.

[7] To test for cointegration, the official price and unit value long term relatives were regressed in an ordinary least squares model; an Augmented Dickey-Fuller (ADF) test was applied to estimate whether or not the residuals were stationary; residuals will pass an ADF test if their mean and variance are stable over time.

The aggregate results of the statistical tests are listed on Table 2, and the detailed results for RMSE, Correlation Coefficient and Cointegration are in Appendix Table A1. The "Good" product categories are primarily agricultural, energy, and industrial commodities. The 30 "Undecided" product categories are primarily secondary inputs into textiles or industry, basic food, drinks and smokes, and a full one third are the grab bag product categories of "Other" or "Not elsewhere classified", which by their nature will have a wider variety of products. Half of. Not all of these "Other" product categories met the vegetable homogeneity test, but the fact that these "Other" product categories comprise primary or secondary materials gives these products the potential to be homogeneous. Finally, the quality group with the comparably "Poor" unit value indexes primarily measure industrial, final consumer, and advanced technology products.

The official publication quality review is a regular step in monthly production; the share of price indexes that are officially published versus suppressed is an important consideration in this research on the feasibility of the unit value indexes as price indexes. We find that the share of official price indexes that are published directly corresponds to the quality groups to which the unit value indexes were assigned. Of all 5-digit BEA End Use export price indexes, only 9 of the 24 product categories are not currently published for the "Good" quality group of indexes. Half of the "Undecided" quality group are not currently published. And a full two thirds of the official price indexes that have "Poor" fitting unit value indexes are suppressed for publication.

The first statistical test of unit value quality was to estimate the mean[8] and standard deviation[9] of the unit value indexes versus the official price indexes, and then also to evaluate the mean difference between the unit value and price indexes. Thirty-two 5-digit BEA product categories passed the test of similar means and proximate standard deviations; and the mean difference between unit value and price indexes fell roughly within one standard deviation of the official price index[10]. Of the 32, two were heterogeneous product categories, and of the homogeneous product categories thirty passed all three tests. Of the 30 homogeneous product categories, 24 indexes passed all 3 tests of mean and standard deviation as well as at least one other statistical test. These 24 unit value indexes are judged to be "Good" indexes at this point, showing strong intra-item substitutability and consistent price variation between the unit value and benchmark indexes.

---

[8] $|(IPP\ Index_{Mean} - Unit\ Value\ Index_{Mean})| < 10$

[9] $Unit\ Value\ Index_{Stdv} < \ IPP\ Index_{Stdv} + 3$

[10] $Rounded\ IPP\ Index_{Stdv} - abs(IPP\ Index_{Mean} - Unit\ Value\ Index_{Mean}) > 0$

Table 2. Results of Statistical Comparison of Unit Value Indexes to Benchmark Official Price Indexes

		Homogeneous Products						Heterogeneous Products					
	Quality group	Good, N=24		Undecided, N=19		Poor, N=9		Good, N=0		Undecided, N=11		Poor, N=64	
Statistical test	Short or Long Term	Pass	Fail	Pass	Fail	Pass	Fail	Pass	Fail	Pass	Fail	Pass	Fail
Standard Deviation and Mean	LTR	24	0	6	13	0	9	-	-	2	9	0	64
Root Mean Square Error	STR	14	10	9	10	1	8	-	-	8	3	13	51
Correlation Coefficient	STR	12	12	1	18	0	9	-	-	0	11	0	64
Cointegration	LTR	16	8	8	11	5	4	-	-	4	7	31	33

Furthermore, the “Poor” unit value indexes are able to be defined by their failure to meet all three standard deviation and mean thresholds. While a number of “Undecided” unit value indexes also failed to meet at least one of the three comparative tests, the “Undecided” indexes successfully met the threshold for one of the other statistical tests, or for the level difference from January 2012 to December 2017. The values of the root mean square error, correlation coefficient, and cointegration tests for individual unit value indexes are shown graphically in the appendix (graphs A1-A4).

In Table 3, for the official price and unit value indexes from January 2012 to December 2017, the average mean of all unit value index of “Good” quality are almost identical to that of the official price index, and the average standard deviations for both are slightly over 10. The difference between means and standard deviations is larger for the “Undecided” and “Poor” quality indexes. For the 30 “Undecided” unit value indexes the unit value and official price index means are 106.0 and 98.2 respectively. The divergent values reflect an upward drift in the unit value indexes, a flag for potential unit value bias. The unit value indexes’ average standard deviation is 13.9, higher than the 8.4 of the official price indexes. The variability around the mean for the unit value index implies that intra-item substitutability is not maintained, and thus unit value bias is a concern. For the “Poor” quality, the unit value indexes have a mean index value twice as much as the official price index. This is to be expected as the heterogeneous unit value indexes comprise most of these product groups and are not expected to follow intra-item substitutability because products are dissimilar or changing in nature. The standard deviation for all “Poor” unit value indexes is extremely high, 158.9, compared to the official price index standard deviation of 4.4.

Table 3. Average Index Means and Standard Deviations

Type	Mean – Unit Value Index	STDV – Unit Value Index	Mean – Official Price Index	STDV – Official Price Index
Good	90.9	11.2	89.6	13.4
Undecided	106.0	13.9	98.2	8.6
Poor	220.6	158.9	100.7	4.4

To recap, an index that falls in the "Good" category passed at least one of the three additional statistical tests: correlation coefficient, root mean squared error, or the cointegration (BIC), in addition to the three mean/standard deviation tests. The "Undecided" and "Poor" categories both include indexes that failed all three tests. However, some may have validly passed some of these statistical tests. There are some "Undecided" indexes that passed the comparative tests on the mean and standard deviations and even may have passed one or more of the three additional statistical tests (e.g. correlation coefficient, root mean squared error, or the cointegration test), but the trend line and the percentage point difference in the December 2017 index value were not in sync with the test results, and we judged them to be not as high quality. For example, the highest correlation coefficient for any index of "Undecided" quality was 0.5019 and the highest of "Poor" quality was 0.2771. The differences in the ways that the unit value indexes compare to the official indexes is distinctly different for the indexes in the "Undecided" and "Poor" quality groups. Interestingly, not only are the price changes not correlated, they do not show as much of a range of price change. An analysis of the average index ranges from the base period to December 2017 for "Good" quality indexes was 188.1, and much less for "Undecided" and "Poor" quality indexes, 141.3 and 121.0 respectively. As mentioned previously, homogeneous products have greater price variability, but the degree of price variability is bounded by intra-item substitutability.

Index methods and bias

The statistical tests provide more information to improve the evaluation of the quality of unit value indexes. The homogeneity condition that is first established with the coefficient of variation is further refined with the statistical tests of unit values against a benchmark price index. We assume that bias is less likely to exist when the time series unit value index based on the complete dataset have similar trends and levels compared to the official price index. It is important to consider the differences between unit value and official price indexes that occur as a result of the different index methods used to calculate each.

First, we do not address criticisms regarding quality adjustment and official price indexes, given that quality adjustment is relevant to heterogeneous product categories and we narrow our scope to homogeneous product categories. Second, we consider the potential of nonresponse bias and fixed weight bias in the official price index, which potentially could mean that the benchmark is flawed or the comparisons are not precise. Third, we explain the methods in unit value indexes that correct for existing bias in official indexes, while also reflecting on the potential of introducing new bias in the measurement of prices. We think this discussion of bias will contribute to an improved understanding of the differences between unit value and official price indexes in the "Undecided" quality category.

The official MXPI are commonly criticized for not adequately accounting for timely changes in products imported and exported, not adequately capturing the CES-driven substitution effect, which should result in importers to switch to lower priced imports or exporters to respond to competitive pressures by lowering their prices, and not addressing sourcing substitution bias from domestic inputs to foreign imported inputs (Reinsdorf and Yuskavage 2014[11]). The operational exigencies of sampling and direct

[11] Offshoring, Sourcing Substitution Bias and the Measurement of US
Import Prices, GDP and Productivity, Marshall Reinsdorf and Robert Yuskavage, 2014. Bureau of economic analysis. Working Paper 2014-5 URL: https://www.bea.gov/system/files/papers/WP2014-5.pdf.

data collection align with the matched model concept; this results in logistical lags to incorporate new goods and replacement of substitutes. Less commonly discussed but quite relevant for index quality is the impact of nonresponse bias on index quality. MXPI are published only when the value of trade of a product category meets the trade dollar cut-off that is set in the sampling process to select sufficient coverage and items to be representative of trade. Even if the trade dollar cut-off is met to publish an index, the product area requires sufficient coverage of companies and items to maintain publishability; lack of participation by companies reduces the number of detailed product MXPI that can be published.[12]

These issues in the official MXPI are addressed with the data and methods proposed for calculating unit value indexes. First, new and exiting goods are accurately counted because the complete dataset accounts for all current trade. Furthermore, the CES-driven substitution effect can be addressed because of the availability of current weights, ie trade values. The Tornquist index is used to aggregate average prices to lower level strata. And while sourcing substitution bias is an import phenomenon, the proposed method described here for exports is flexible enough to account for changes in import quantities and prices when sourcing shifts from one foreign country to another, if not from one supplier to another; however, these methods do not evaluate sourcing changes from domestic to foreign markets. Nonresponse and lack of participation by businesses does not affect the administrative trade dataset, because reporting trade is a legal requirement. An additional benefit is that outliers are mitigated by applying a geomean calculation to the ELIs and using current weights to aggregate ELIs to the lower level strata with a Tornquist index calculation.

However, not all proposed methodology changes are clear improvements. We discovered, for example, that Tornquist indexes result in price trends *dampening* in both directions for most indexes. This is a non-intuitive result. When evaluating results for the CES-drive substitution, we expected that the unit value index would be lower than the official price index, as the Tornquist use of current weights accounts for changes in demand in the face of lower prices. This result holds with the unit value indexes that rise. However, when unit value and price indexes declined, this theoretical consideration does not bear out; with declines in the official price index, declines in the comparable unit value index were less steep, or flatter. This practical result does not hold with theory. We determined that this dampening is a mathematical characteristic of the changing monthly weights, because in numerous artificial data scenarios with fluctuating and stable prices and weights, we found that the only situation in which the Tornquist consistently trended lower (rather than flatter) than the Laspeyres was in the case of unchanging q, ie unchanging weights.

This potential new type of bias – due to weight variability of monthly trade in the complete dataset and the addition of current weights in the Tornquist calculation – will be the subject of further research and possibly different calculations; it may be outweighed with the other improvements that present themselves in the alternative data source and methods. Nonetheless, understanding the issue is important; weight variability is more common with industry statistics and especially MXPI, because international trade is lumpy. Business purchases and sales may not be market driven, as shipping costs, customs regulations, and production processes will affect how much and when something is purchased or sold. Logistical aspects of credit and shipping also can contribute to uneven trade. The data exhibit these characteristics of trade, showing wide variability in trade values across months, which impacts

[12] Note some product prices are collected from other government surveys (ie some agriculture commodity exports) or as a spot price (some import and export metals).

measurement. These issues are considered as we continue to parse out the reasons for differences in the “Undecided” quality unit value indexes which represent 20 percent of export goods trade.

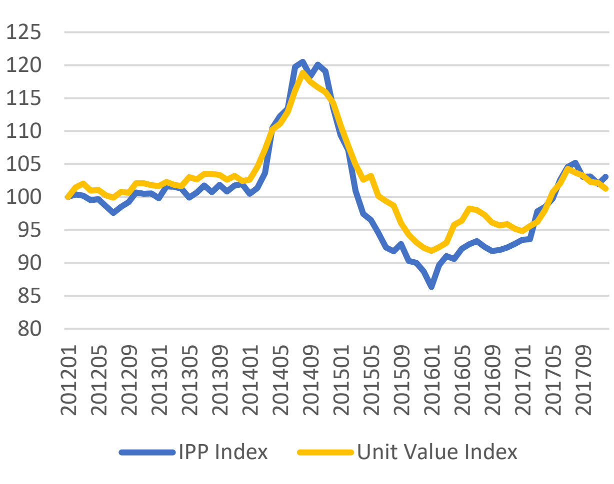
Snapshots of Index Comparisons – and the time series deflator “sniff” test

We present here examples of “Good”, “Undecided, and “Poor” indexes and the statistics that underpin the evaluation of quality. Table 4 shows a snapshot of the statistics that describe the differences between a meat product category, a commodity product category, and computer parts, a heterogeneous product category. The homogeneous meat and computer parts product categories trend as expected, as can be seen in Graphs 1 and 2. The meat products unit value index closely matches the official index, has a strong correlation with it, and closely tracks the long term relative. In contrast, the computer parts unit value index is widely divergent, as expected, and the statistical analysis highlights the differences with the official price index. The unit value index of the homogeneous category of rubber is not as clear cut because, while it appears that the mean and standard deviation are neither proximate nor distant, the gap in the end index value is large at 20 index points different. (See Graph 3.) The STRs are not closely correlated, but there exists some consistent relationship between the XPI and the unit value index. These ambiguities could be related to errors or mismeasurement in either the unit value index or its corresponding official price index.

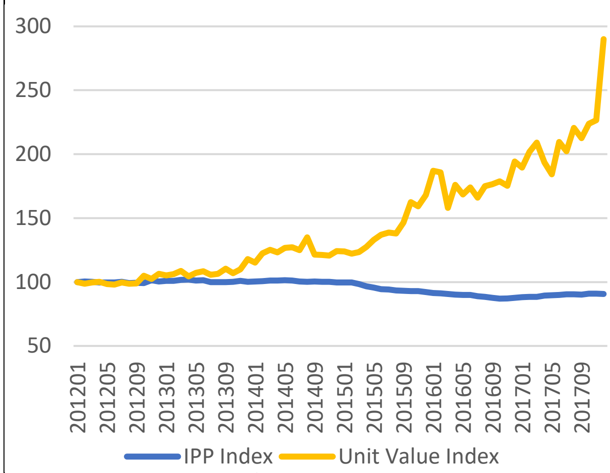
Table 4. Statistical Results Comparing Three Different Quality Unit Value Indexes with Corresponding XPI

Product Category	Meat, Poultry, and Other Edible Animal Products (00300) "Good"		Synthetic Rubber-Primary (12700) "Undecided"		Computer Peripherals, Accessories and Parts (21301) "Poor"	
Statistical Test and Results by Index Type	Unit Value	Official	Unit Value	Official	Unit Value	Official
Mean	102.39	100.24	87.19	75.26	143.75	95.93
Standard Deviation	5.89	7.79	10.26	15.26	41.92	5.10
Dec 2017 (Jan 2012=100)	101.25	103.07	59.01	79.18	290.02	90.73
Correlation coefficient	0.67	-	0.34	-	-0.09	-
RMSE	1.50	-	3.49	-	6.18	-
Cointegration	0.0002	-	0.0672	-	0.9912	-

Graph 1. Index Comparison - Meat, Poultry, and Other Edible Animal Products

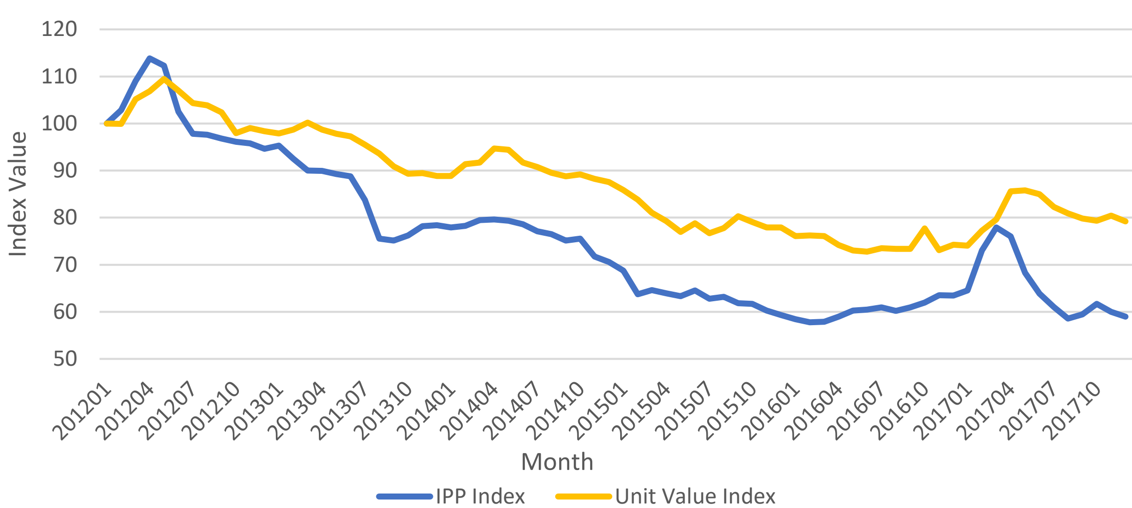


Graph 2.Index Comparison – Computer peripherals, accessories and parts



Some of the statistical results seen in Table 4 would put the homogeneous rubber index into the "Good" category. But the gap between the final indexes is not conducive to use as a substitute for the existing deflator for this index. On further inspection, the two indexes appear to run in parallel, trending in similar directions. The correlation coefficient showed poor predictive power, but we see that there is a correlation in the lagged STR by one to two months. This lag is not captured in any of the statistical tests that we have applied. The flatter trend of the unit value index was discussed previously and is characteristic of the Tornquist formula bias, but this bias does not present in the "Good" quality indexes. It would appear that an alternative calculation, a better item key or a lag could improve the fit of these two indexes and potentially close the gap in the end value.

Graph 3. Index Comparison - Synthetic Rubber-Primary



Impact of Unit Values on GDP Measurement

MXPI at a detailed level are used as deflators for net exports by both the US Census and the Bureau of Economic Analysis. Net exports are a component of Gross Domestic Product. If we reach our goal of providing more and better quality statistics with the administrative trade data, we will also contribute to improving the measure of GDP.

To simulate the impact of replacing the historic BEA End Use 5-digit export price indexes for 2012-17, we aggregate price indexes for each subgrouping of the “Good”, “Undecided”, and “Poor” price indexes and evaluate the difference between the trade-weighted unit value index endpoint of December 2017 of each quality group with that of the corresponding official export price indexes - first for the partial aggregation and then to the top-level index, holding all else equal. As can be seen in Table 5, the 24 5-digit BEA End Use “Good” unit value indexes show a December 2017 index value of 83.06, calculated from the lower level strata using the official method of 2-year lagged fixed dollar weights and the Laspeyres index formula. This compares to the estimate of the corresponding official XPIs of 80.07. The total dollar value of these indexes is 23 percent of export trade. We calculate the impact on all goods’ export prices by assuming that the other 77 percent of trade value does not change prices. Applying official MXPI methodology, the 5-digit indexes are aggregated to an all-goods measure, which results in a top-level aggregate unit value index of 96.08 and the official XPI of 95.38. The price change gap in the “Good” price indexes is 3.7% over the time period, or .75% per year. For the top-level index comparison, the price gap between the “Good” unit value indexes versus the historic official price indexes is .73% percent from 2012 to 2017, or a .15% annual difference. That is, if the “Good” unit value indexes were to be incorporated into the all-goods XPI, export prices would rise .15% more per year. And when applied as deflators, real export prices would fall by .15 % more per year. The same thought experiment comparing the “Undecided” unit value indexes with their comparable 5-digit official price indexes results in a smaller gap and smaller impact of .26% annual price gap for the 30 unit value indexes and a .05% price gap for the top-level index, but in the opposite direction. Despite the relatively poorer quality of indexes, the 30 unit value indexes of “Undecided” quality, when aggregated and compared to their comparable official price indexes at both the detailed and top-level all-goods indexes, had less of an impact. The “Poor” unit value indexes show an extreme annual price gap of 41.8% for the 73 unit value indexes in this category, and 24.2% for the top-level indexes.

Table 5. Aggregate Impact of Goods Indexes
December 2017

Type	Aggregation of Official Indexes	Aggregation of Unit Value Indexes	Annual % difference	Top level % change of Official Indexes	Top level % change of UV Index	Annual % difference	Percent of Total Trade $ Value
	Jan 2012=100	Jan 2012=100	Jan 2012-Dec 2017	Jan 2012=100	Jan 2012=100	Jan 2012-Dec 2017	Base year 2015
Good (N=24)	80.07	83.06	0.75%	95.38	96.08	0.15%	23%
Undecided (N=30)	99.75	98.47	-0.26%	99.95	99.69	-0.05%	20%
Poor (N=73)	99.47	307.47	41.82%	99.69	220.57	24.25%	57%

This range of price gaps, and the extreme variability for the “Poor” quality grouping validates the basic intra-item substitutability tenet that guides our research and qualitatively affirms that we have selected reasonable quality groupings.

The effort to measure the impact on net trade and thus GDP will depend on the results from the import comparison. Hypothetically, if there were a commensurate adjustment upward to the import index prices, and thus downward impact on real goods imports, the impact on net trade may be small, and thus have a minimal impact on GDP. The direction of the impact is sensitive to the choice and number of indexes determined to be “Good”. The different composition of imports, with a larger share of heterogeneous products, must be calculated before the possible impact on the real trade balance and GDP can be measured. The opportunity to identify goods by ‘manufacturer ID’ in the locality of origin of the import could potentially benefit the fit of MPI to the published price indexes.

Conclusion and Future Steps

Our research shows that not all unit value indexes are the same, and that unit value indexes used in the official MXPI measures must be homogeneous and is expected to closely align with the comparable official price index. The test run of unit value indexes for 127 detailed product categories over six years shows that defining homogeneity matters, and that there are statistical tests to evaluate differences and make judgments on the consistency, reliability, and comparability of unit value indexes relative to the official export price indexes. The maximum range of coverage of unit value indexes up to 43 percent depends on a better understanding of indexes of middling quality – “Undecided” – and we have identified some alternative approaches to improve the quality of these unit value indexes, by evalaating the degree of intra-item substitutability, the bias introduced by the new methodology, and/or the nonresponse bias in the official price index.

We continue to refine our efforts to secure the maximum coverage with the best quality even as we get closer to a decision to operationalize the measures. For example, we will estimate hedonic linear regressions on the complete dataset to develop a systematic method of identifying the ideal item key combination for each strata. Also, we will consider whether to explore time-dummy hedonic models to reduce the specification constraints of grouping data variables and calculating ELIs. Hopefully not only can we improve the quality of the “Good” set of indexes with this refinement, but additionally “Undecided” indexes possibly can move to the “Good” category. The Tornqvist index formula helps with substitution bias, but does not fully alleviate the drift found in the chained Laspeyres index. Because frequent chaining has been determined to exacerbate chain drift[13][14], we will evaluate the possibility and resources needed to measure chain drift in the Unit Value indexes and investigate alternative aggregation methods, such as the base construction strategy described by Statistics Finland[15].

[13] Ivancic, L., W.E. Diewert and K.J. Fox (2009), “Scanner Data, Time Aggregation and the Construction of Price Indexes”, Discussion Paper no. 09-09, Department of Economics, University of British Columbia, Vancouver, Canada.
[14] de Haan, J. and H.A. van der Grient (2011), “Eliminating Chain Drift in Price Indexes Based on Scanner Data”, Journal of Econometrics 161, 36-46.
[15] Nieminen, Kristiina and Montonen, Satu (2018), “The foundation of index calculation”, Statistics Finland, Helsinki. (URL: http://www.stat.fi/static/media/uploads/meta_en/menetelmakehitystyo/foundation_of_index_calculation_nieminenmontonen_final.pdf)

In addition, work will be done to analyze and compare import unit value indexes with official import price indexes and to calculate partial-month measures with low variance compared to full-month data. From our work less than two years ago beginning to consider whether unit value indexes are viable alternatives to price indexes, we have come a long way. We are now driving on the road, but we have not yet arrived at our destination.

Appendix

Table A.1 Statistical Test Results Unit Value Index compared to Official Price Index Benchmarks

BEA Code	Index Description	Results	Homogeneous Heterogeneous	Correlation Coefficient > .5	RMSE < 4	Cointe-gration (BIC) < .05
00000	Wheat	Good	HM	0.5035	5.4007	0.0060
00010	Rice other food grains	Good	HM	0.3991	2.4711	0.2435
00100	Soybeans and soybean by-products, prior to the extraction of oil	Good	HM	0.5483	4.8812	0.0000
00110	Oilseeds other than soybeans, and food oils	Good	HM	0.2324	3.0893	0.0187
00200	Corn	Good	HM	0.5914	4.3994	0.0000
00220	Other animal feeds, n.e.s.	Good	HM	0.5659	2.6767	0.0000
00300	Meat, poultry other edible animal products	Good	HM	0.6685	1.5001	0.0002
00330	Vegetables and vegetable preparations and juices	Good	HM	0.3193	7.8922	0.0207
00340	Nuts preparations	Good	HM	0.3021	3.9120	0.0632
00350	Bakery confectionery products	Good	HM	0.0848	1.2338	0.3965
10000	Cotton incl linters-raw	Good	HM	0.3462	3.6008	0.0631
10100	Tobacco, unmanufactured	Good	HM	-0.0661	4.6763	0.0000
11010	Metallurgical grade coal	Good	HM	0.6222	4.3259	0.0477
11100	Crude	Good	HM	0.5930	6.0992	0.0944
11110	Fuel oil	Good	HM	0.8172	4.9671	0.0060
11120	Other petroleum products	Good	HM	0.7575	3.0244	0.0284
11130	Natural gas liquids and mfd gas	Good	HM	0.6205	7.9260	0.0854
12000	Steelmaking ferroalloying mater	Good	HM	0.6581	3.7291	0.0003
12200	Aluminum alumina	Good	HM	0.3983	2.0630	0.0000
12210	Copper	Good	HM	0.7184	2.2597	0.0000
12260	Nonmonetary gold	Good	HM	0.4739	3.6209	0.0558
12270	Other precious metals	Good	HM	0.2692	8.2255	0.0003
12430	Linerboard, newsprint, and other paper/paperboard	Good	HM	0.2009	1.2398	0.8604
12500	Plastic materials	Good	HM	0.5845	1.2116	0.0853
00210	Other feedgrains	Undecided	HM	0.2312	7.4431	0.0449
00310	Dairy products eggs	Undecided	HM	0.5019	5.8990	0.0707
00320	Fruit and fruit preparations, including fruit juices	Undecided	HM	0.0799	6.2916	0.0000
00360	Other foods and food preparations	Undecided	HM	0.1325	1.4546	0.0024
01000	Fish and shellfish	Undecided	HM	-0.0006	3.9712	0.3807
01010	Distilled alcoholic beverages	Undecided	HM	0.0920	5.4131	0.0230
10120	Hides, skins furskins-raw	Undecided	HM	0.3026	5.0731	0.0382
10130	Other agricultural materials	Undecided	HM	0.2386	3.7681	0.0333

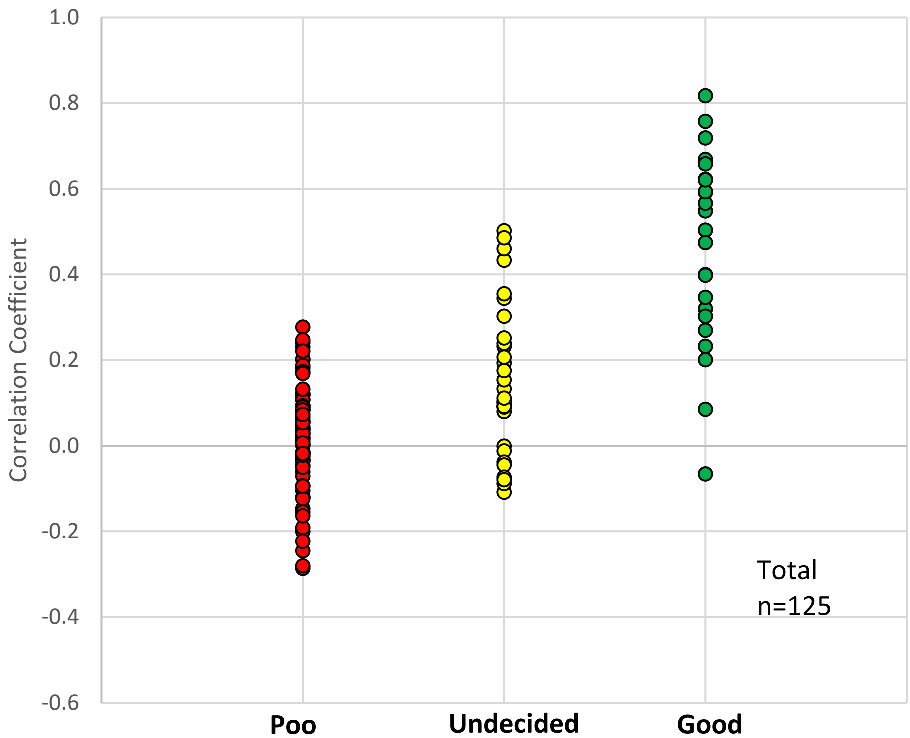
10150	Other agricultural materials-manufactured	Undecided	HT	-0.0124	2.7133	0.7118
11020	Other coal related fuels	Undecided	HM	0.4331	11.4035	0.0000
12100	Iron steel mill products	Undecided	HM	0.4598	1.6488	0.8333
12290	Other nonferrous metals	Undecided	HT	0.1931	3.6110	0.9963
12420	Woodpulp and recovered paper	Undecided	HM	0.4856	3.3320	0.7595
12600	Cotton other natural clothing	Undecided	HM	0.1000	2.6335	0.0101
12620	Manmade cloth and thread cordage	Undecided	HT	-0.0383	1.9114	0.6517
12630	Oth materials	Undecided	HM	-0.0451	4.5173	0.2502
12650	Unmanufactured leather and fur	Undecided	HM	0.0903	7.3395	0.1831
12700	Synthetic rubber-primary	Undecided	HM	0.3443	3.4938	0.0672
12720	Nonmetallic minerals, nes	Undecided	HT	0.1036	4.2136	0.1625
12760	Mineral supplies-manufactured	Undecided	HT	0.0918	2.2544	0.6851
12770	Other goods manufactured and unmanufactured	Undecided	HT	-0.1086	2.9617	0.0015
13100	Logs, lumber, plywood and veneers	Undecided	HM	0.3546	1.4349	0.7951
22000	Civilian aircraft	Undecided	HT	0.2518	2.4538	0.0000
30000	New complete assembled automobiles	Undecided	HM	0.1533	41.1392	0.0000
30100	Trucks, buses, spec-purpose veh	Undecided	HM	-0.0886	1.8557	0.7742
30200	Engines and engine parts for automotive vehicles	Undecided	HT	0.1751	1.4124	0.4041
30220	Automotive tires tubes	Undecided	HT	-0.0738	2.6888	0.0737
40130	Cigars, cigarettes, and other tobacco	Undecided	HM	0.1111	11.1467	0.1849
40140	Other products	Undecided	HT	-0.0796	139.4971	0.0000
42000	Nursery stock, cut flowers	Undecided	HT	0.2065	20.6520	0.0084

Table A.2 Comparison of Good Indexes between the Official and Unit Value Indexes in December 2017

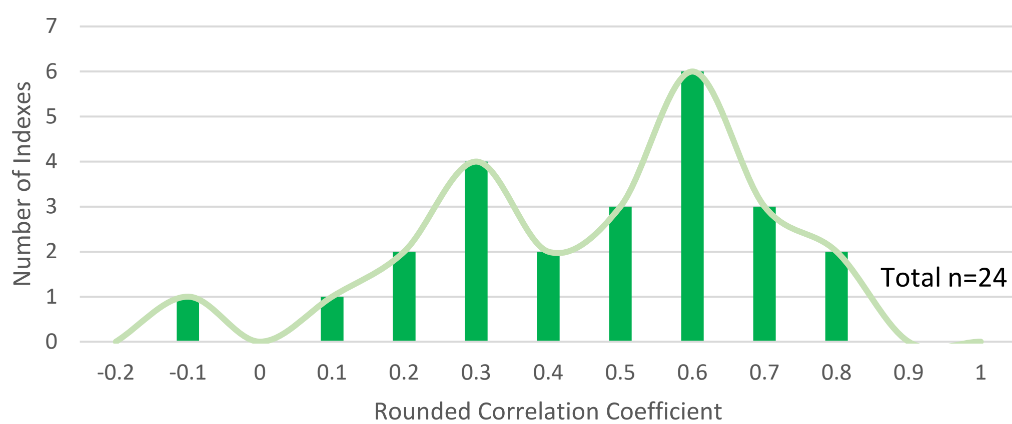
Index	Pub-lished	Official 2015 Weight	Official Dec. 2017 LTR	Unit Value Dec. 2017 LTR	Official-UV Differ-ence	Official-UV 5-year Percent Gap
Wheat	Y	6,223,098,400	75.08	84.13	9.05	12.06
Rice other food grains	N			89.75		
Soybeans and soybean by-products, prior to the extraction of oil	Y	22,224,705,489	81.32	78.97	-2.35	-2.89
Oilseeds other than soybeans, and food oils	N			83.34		
Corn	Y	10,151,665,553	56.02	64.41	8.40	14.99
Other animal feeds, n.e.s.	Y	8,237,438,320	102.65	104.71	2.05	2.00
Meat, poultry other edible animal products	Y	18,837,591,560	103.07	102.81	-0.26	-0.25
Vegetables and vegetable preparations and juices	N			106.54		
Nuts preparations	N			120.20		
Bakery confectionery products	N			106.08		
Cotton incl linters-raw	N			72.59		
Tobacco, unmanufactured	N			102.21		
Metallurgical grade coal	N			59.52		
Crude	N			45.50		
Fuel oil	Y	38,100,763,240	62.19	60.78	-1.41	-2.26
Other petroleum products	Y	49,081,922,540	69.85	82.51	12.65	18.12
Natural gas liquids and mfd gas	N			88.64		
Steelmaking ferroalloying mater	N			75.00		
Aluminum alumina	Y	7,820,182,985	92.55	94.83	2.28	2.46
Copper	Y	6,172,567,283	92.29	93.67	1.38	1.50
Nonmonetary gold	Y	21,021,114,128	77.02	77.33	0.31	0.40
Other precious metals	Y	6,631,330,913	74.54	71.58	-2.96	-3.97
Linerboard, newsprint, and other paper/paperboard	Y	13,274,733,248	106.94	111.53	4.59	4.29
Plastic materials	Y	34,715,404,167	95.96	101.10	5.14	5.35

Index difference from 201201 to 201207 for "Good" Quality Indexes

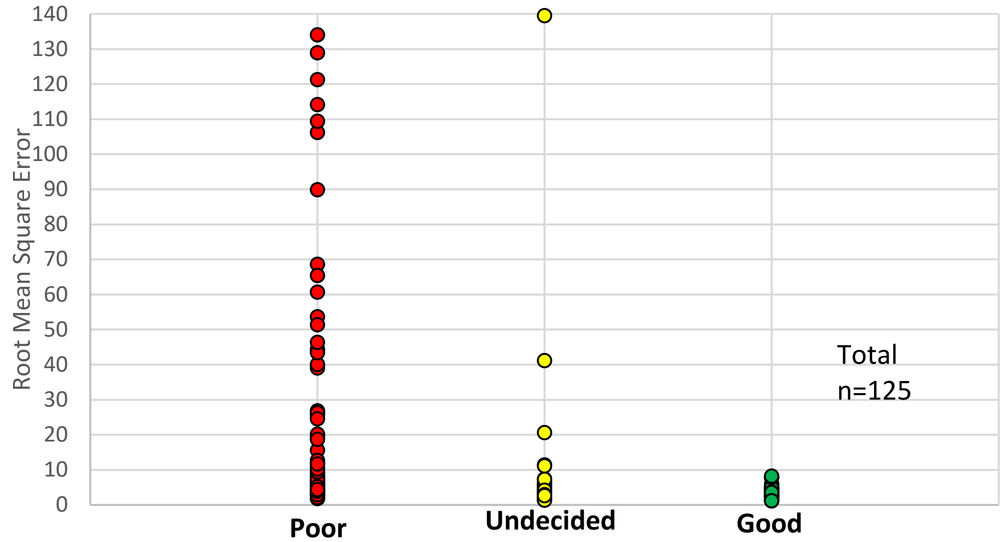
Graph A1. Distribution of Correlation Coefficients by Quality Categorizations



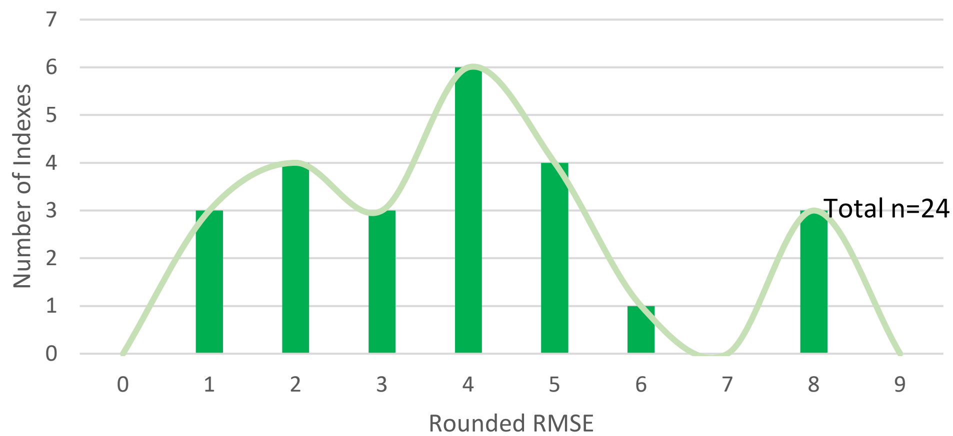
Graph A2. Frequency Distribution of Rounded Correlation Coefficient Categories for "Good" Indexes



Graph A3. Distribution of Root Mean Square Errors by Quality Categories



Graph A4. Frequency Distribution of Rounded Root Mean Square Error for "Good" Indexes



Note for graphs A1-A4: Data displayed are truncated to exclude extreme values.