

Measuring Export Price Movements With Administrative Trade Data

by

Don A. Fast and Susan E. Fleck

February 2019

Prepared for the CRIW Conference Big Data for 21st Century Economic Statistics

Abstract

The International Price Program (IPP) surveys establishments to collect price data of merchandise trade and calculates import and export price indexes (MXPI). In an effort to expand the quantity and quality of MXPI, we research the potential to augment the number of price indexes by calculating thousands, and potentially millions, of prices directly from export administrative trade transaction data maintained by the Census Bureau. This pilot research requires reconsideration of the long-held view that unit value price indexes are biased because product mix changes account for a large share of price movement. The research addresses this methodological concern and identifies others by analyzing two semi-homogeneous product categories among the 129 5-digit BEA End Use export categories. The results provide a road map of a consistent and testable approach that aligns with the concepts used in existing MXPI measures, maximizes the use of high volume data, and mitigates the risk of unit value bias. Preliminary analysis of all 129 5-digit BEA End Use categories for exports shows potential for calculating export price indexes for 50 of the 5-digit classification categories, of which 21 are currently not published.

Don A. Fast is a Senior Economist in the International Price Program. Susan E. Fleck is the Assistant Commissioner of the International Price Program at the U.S. Bureau of Labor Statistics (BLS). Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the view of the U.S. Bureau of Labor Statistics. We thank Daryl Slusher, Laurence Lang, Michael Havlin, Tamar Schmidt, Helen McCulley, Praveenkumar Yerramareddy, Antonio Caraballo, Sudha Polumatla, Jeff Blaha, Aric Schneider, Robert Martin, Jenny Fitzgerald, Ara Khatchadorian, David Friedman, Steven Paben of BLS, and Brian Moyer, Ana Aizorbe, and Jon Samuels of BEA. Author emails: fast.don@bls.gov, fleck.susan@bls.gov.

The BLS has received prior approval from the U.S. Census Bureau, which affirms that the paper does not present any disclosure risks and approves this document to be shared with the appropriate parties related to the CRIW Big Data for 21st Century Economic Statistics. Use of the 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).

Introduction

BLS Import and Export Price Indexes (MXPI) track price changes in internationally traded merchandise goods. The indexes are used to adjust U.S. net exports and trade balances from current to constant dollars. In addition, the indexes and the data that underlie them are used to analyze cross-country price competitiveness, exchange rate pass through effects on domestic inflation, price transmission in the global value chain, and productivity. For example, Gopinath (2010) uses currency and price data for imports to evaluate exchange rate pass-through of goods; Clausing (2003) examines taxation impacts on intrafirm trade prices; and Houseman, Kurz, Lengermann, and Mandel (2011) discuss how U.S. productivity growth can be overestimated due to offshoring.

The International Price Program's mission is to maintain and expand the level of detail and quality of import and export price indexes, but this mission is hindered by the high cost of direct data collection. The limited amount of product and industry detail impacts the quality of real Gross Domestic Product (GDP) and international trade balances. Where detailed import and export product indexes are not published, the Bureau of Economic Analysis (BEA) and the Census Bureau use deflators from parent-level indexes as proxies for missing detail or turn to domestic Producer Price Indexes (PPI), as a best alternative.

In an effort to expand the coverage and quality of MXPI, BLS is piloting research to use administrative export trade data provided by the U.S. Census Bureau to calculate detailed price indexes that could supplement directly collected data. The administrative trade records represent the universe of nearly all export transactions. The 2.9 million monthly records of export trade dwarfs the 20,000 items in the directly collected international price survey. Quantity, however, does not necessarily mean quality. A major drawback to using this data source for price indexes is that prices calculated from administrative data will be, by definition, unit value prices - that represent the average price of a group of products and not a specific product. Unit value indexes reflect price changes of both like-product market factors and changes in product mix; the methodology is commonly known to distort and bias price movements, especially for heterogeneous product areas (Alterman 1991). It was precisely this quality problem that led the Bureau of Labor Statistics to initiate the international price survey to directly collect data using a matched-item model approach. The sampled approach replaced and improved upon the Census Bureau's unit value import and export price index series.

Maintaining the standard for Principal Federal Economic Indicators requires thoughtful and thorough review of new methods. Such is the intent of this research. The research develops and evaluates new methods to test the usability of administrative trade data and the quality of resulting price index measurements. This paper describes the pilot research and the systematic approach to create and evaluate detailed transaction prices and to calculate and evaluate monthly price indexes. The selection of two semi-homogeneous BEA End Use export product categories was based on their high degree of homogeneity of the 10-digit product detail, and also because the two 5-digit indexes had experienced difficulties with respondent cooperation. The first assumption is shown to be a valid premise for selection, but the second created difficulties in validating the consistency and quality of the pilot measures.

The paper describes how the approach for selecting items and index methods is motivated by the concepts and methods of the official MXPI. The numerous permutations of item prices and indexes are analyzed and evaluated for the quality of the results using all 2015-16 export transaction records for two semi-homogeneous product categories, BEA End Use export product categories 00310 (Dairy Products and Eggs) and 00330 (Vegetables and Vegetable Preparations and Juices). First, item-specific concepts and methods are described and options for calculating item prices are put forth, analyzing variable choice, consistency and outliers and their impact on indexes. Second, prices are calculated for the detailed proxy items, and the prices and price changes are tested for unit value bias within and across months to identify the proxy item that

demonstrates the least bias. Third, the numerous variations of BEA End Use indexes are statistically compared to select published benchmarks, and the panoply of pilot indexes is narrowed down to those that best fit the benchmark.

Background

The official MXPI cover nearly all trade in merchandise goods.ⁱ The matched-item model is the conceptual basis of the import and export price indexes. Approximately 20,000 items are priced monthly to calculate both import and export price indexes. The majority of item prices are directly collected. Some homogeneous product categories (e.g. export grains uses data from the U.S. Department of Agriculture and import petroleum uses data from the U.S. Department of Energy) are priced using administrative trade data.

Annually approximately 2,000 establishments across half of all merchandise product categories are sampled proportional to their weight in trade to provide continuous representative coverage in the baskets of import and export goods (the full sample covers two years and no more than 4,000 establishments for imports and exports combined). Cooperative establishments provide prices for precisely-defined items and report the price of each item over the life cycle of the sample. Up to four panels of item prices are included in the MXPI. A Laspeyres index formula is used to aggregate item and detailed product groups (also referred to as classification group) into product and industry strata to the topline import and export goods price index measures. At the item level a probability-based establishment sample weight is used to calculate the product group index. Product group to the topline index weights are calculated with 2-year lagged annual trade dollar values based on product/industry classifications, such that 2019 price indexes are calculated using 2017 trade weights and classifications.ⁱⁱ

A total of 60 out of 129 5-digit BEA End Use export indexes are presently published. Other indexes do not meet the publication criteria due to insufficient coverage, potential confidentiality concerns, or dollar value of trade. Trade in a product or industry category must meet a trade dollar value threshold for an index to be published. This minimum quality threshold is dependent on the size of the sample that is fielded. The establishments and items must be sufficient to ensure representativeness while also protecting respondent identifiable information to support the publication of an index. In the face of finite data collection resources, an increase in the value of trade or a reduction in data collection resources results in fewer establishments and items in the survey and negatively impacts index quality. Both published and unpublished product categories are sampled to assure that the higher levels of aggregation are properly representative of the trade.

The magnitude of observations in the administrative trade data far outnumbers those in the directly collected data. The volume of usable records would virtually eliminate the risk of disclosure and assure confidentiality, while also eliminating the bias that accompanies nonresponse. On average, 8,000 prices are collected per month to support the publication of all export indexes. The average number of monthly observations over the

ⁱ Some 10-digit product areas are excluded from MXPI for merchandise goods (e.g. returned goods, used goods, military goods and firearms). BLS also publishes import and export service price indexes for air passenger and air freight. All subsequent references to indexes refers to merchandise trade price indexes.

ⁱⁱ MXPI are published for three classification systems (HS, BEA, and NAICS). A mapping, or “tree” structure is used for aggregating data and/or weights annually. A concordance is created by taking the relevant year of HTSUSA (imports) and Schedule B (exports) classification and mapping these classifications to IPP ten-digit classification groups. These IPP ten-digit classification groups are then mapped to four-digit HS, five-digit BEA, and six-digit NAICS categories and vertically mapped into higher level indexes.

two-year period considered in this paper nears the export item official count for Dairy and far exceeds it for Vegetables with 6,839 and 32,430 monthly records respectively. The potential for administrative data to expand the quantity and improve the quality of price measures, however, faces two hurdles. The first – whether there is a method to calculate indexes accurately – is the focus of this paper. The second – whether there is a way to integrate the lagged administrative data into official monthly production – will not be addressed here, but is not insignificant.

Research Approach

The goal of this research is to determine if the transaction data can be used to calculate accurate indexes that can be used as a substitute for semi-homogeneous import and export price indexes. Pilot research will focus on two semi-homogeneous export product categories (stratum) - the BEA End Use export product categories of “Dairy Products and Eggs” (BEA classification 00310) and “Vegetables and Vegetable Preparations and Juices” (BEA classification 00330) for 2015 and 2016. Within these product categories in 2016, hundreds of companies traded close to 3.9 and 7 billion dollars, respectively.

In order to use administrative data for more homogeneous product categories and to blend these indexes with directly collected data/indexes for more heterogeneous product categories, the selection of data characteristics should adhere as closely as possible to the concepts used in the official MXPI measures.

The administrative trade data that is the subject of IPP’s pilot research is currently used as the sample frame for IPP’s sample selection of exports.ⁱⁱⁱ The data are maintained by the U.S. Census Bureau. Regulation requires exporters to enter information on their shipments on a timely basis. Individual records must be submitted for each product type in each shipment and include Employer Identification Numbers (EINs) of the importing and exporting establishments, HTSUSA or Schedule B categories, customs dollar value, quantity, shipping logistics and other information about the shipment. All goods are classified by ten-digit codes based upon the Harmonized System (HS), which is based on the six-digit World Customs Organization (WCO) international Harmonized classification.

This following sections describe item keys used to group trade records by unique characteristics, the method of calculating prices, weights, and indexes, and the types of concerns that arise when trade data are not consistent, missing, or have outliers. The objective of creating a unique item is to attempt to replicate a matched item model.

Alternative Products and Prices

The administrative trade data consists of transaction records of goods’ shipments across borders that list many details of the price, quantity, provenance, and destination of the shipment, for each product category (10-digit harmonized), and party. A number of the data fields in the record format align with existing concepts used in defining price characteristics in the official MXPI. We select the variables that most closely align with price determining characteristics in the international price survey.

For each export establishment, the type of product they trade and the sales price depend on the firm’s production capacity, the preferences in the market where they are selling the good, and the relationship between the seller and the buyer. These characteristics of a product and its price are integrated into the international price survey processing. First, the sample is drawn based upon an establishment’s trade within a

ⁱⁱⁱ Import administrative data are maintained by the U.S. Customs and Border Patrol agency.

detailed published index level, or stratum, and further selection is done using the 10-digit HS category/grouping. Business enterprises are identified by EIN (Employer ID Number) or, for Canada, establishment name. Because one company may have more than one establishment (physical plant), many establishments may have the same EIN, so survey units are broken out by zip code and state; this information is used to initiate contact with an establishment. Administrative data fields– E, D, S, Z, and H – correspond with these sampling characteristics. Second, fixing price characteristics for a product are an important step in defining the price that the respondent will provide monthly. Different goods are shipped to different countries because consumer and business preferences differ by country. The official MXPI publish price indexes by country of destination. The seller-purchaser relation also determines the price. Related trade, e.g. between subsidiaries or related companies, often has different prices than arms-length trade, ceteris paribus. The unit of measure, e.g. gross, piece, ton, is also a price determining characteristic that significantly affects the unit price. The data fields C, S, Q, and R correspond to these price characteristics.

The table below provides data field descriptions from Census’s “2016 Export Net Record Field Names and Descriptions” along with the abbreviation used in this paper:

Table 1 Export Shipment Characteristics

| Abbreviation | Variable | Variable description |
|--------------|---------------------------|---|
| E | EIN | 11-digit (USPPI Number) employer identification number (non-Canadian exports). For Canadian exports where no EIN is provided, the IPP uses the USPPI Name concatenated with the five-digit ZIP code to create a proxy EIN. |
| S | Stateori | 2-digit state of origin code |
| Z | Zipcode | 5-digit ZIP code |
| C | Country | 4-digit country of destination code (Schedule C) |
| D | Dist_Exp + Port_Exp | 2-digit Customs district of exportation (Schedule D) + 2-digit Customs port of exportation (Schedule D) |
| Q | Qty1_dsg | 2-digit designation code of QTY1 reported (Refer to ‘Attachment A for a list of designator codes) In layman’s terms this is the unit of measure for the transaction. |
| R | Related | 1-digit related/non-related code 'Y' = Parties to the transaction are related 'N' = Parties are not related |
| H | HS | 10-digit numeric Harmonized Commodity Classification code (Schedule B) Note: All keys include HS. |

Using these data fields, data are grouped to create item keys. These item keys result in different subgroupings of the administrative data into unique products, or proxy items. All item keys include the Schedule B code (H) as part of the unique identifier.

Table 2 Item Keys

| Item Keys | Description |
|-----------|--|
| ESZCDQRH | EIN, State of Origin, Zip Code, Country of Destination, U.S. Port, Unit of measure, Related, 10-digit HS |
| ESZQRH | EIN, State of Origin, Zip Code, Unit of measure, Related, 10-digit HS |
| ESQRH | EIN, State of Origin, Unit of measure, Related, 10-digit HS |
| EQRH | EIN, Unit of measure, Related, 10-digit HS |
| QRH | Unit of measure, Related, 10-digit HS |
| H | 10-digit HS |

Looking at BEA classifications 00310 (Dairy products & eggs) and 00330 (Vegetables and vegetable preparations and juices) one can see how much data exists to support these alternative products and prices.

Table 3 Volume of trade and trade transactions for 2016

| Export Indexes, BEA End Use | No. of annual trade transactions | No. of EINs with > 50 trades & > \$100k traded | No. 10-digit product categories | Annual trade \$ value |
|---|----------------------------------|--|---------------------------------|-----------------------|
| 00310 - Dairy products & eggs | 82,100 | 220 | 41 | \$3.8 billion |
| 00330 -Vegetables and vegetable preparations and juices | 392,000 | 950 | 161 | \$6.9 billion |

The average price for each proxy item defined by an item key was calculated with a weighted geometric mean. Calculating a geometric mean instead of an arithmetic mean mitigates the impact of outliers, an important consideration given the volume of data. The records with a missing or imputed quantity, are excluded from price calculation.

Weighted Geometric Mean

$$P = \left(\prod_{i=1}^n p_i^{w_i} \right)^{\frac{1}{\sum_{i=1}^n w_i}}$$

Where:

P = Weighted Geomean Price

i = Individual trade transaction

p = Average unit price within a transaction

w = Total dollar value of the unique transaction used as a weight

The prices and price change for each of these item keys are used in the unit value bias analysis later in this paper. Indexes at both the product category and BEA End Use level are calculated from the proxy items that comprise an item key. However, in order to compare indexes to the benchmarks – for all pilot indexes – additional dimensions of index quality are considered below.

Alternative Index Concepts and Methods

Administrative data record patterns of trade by company and product over time. Proxy items are the anchor of monthly price changes. Tracking the proxy item price change over time, however, is subject to variation and variability. This section considers the characteristics of trade to identify potential issues of bias in index methodology and develops a range of options to consider when calculating indexes. These options address choice of index methodology and quality measurement methodologies related to consistent trade, outlier management, and missing price imputation. In the subsequent section on Benchmark comparisons, the variety of indexes based on these different concepts and methods along with the dimension of item keys will be compared to the Benchmarks for Dairy and Vegetables.

Index Formulas. The fact that the administrative trade data provides current dollar value of trade begs the question of how we could use this data for index calculation. The official MXPI uses a Laspeyres index formula, because of the data constraint that current trade weights are not available. The availability of current period weights in the administrative data expands our options to use a Tornqvist index formula to aggregate the price data from detailed proxy item to the 10-digit classification group index. In “Economic Theory and BEA’s release of Alternative Quantity and Price Indexes”, Triplett discusses the benefit of superlative indexes, which Diewert had previously demonstrated in 1976. Specifically, superlative indexes (Fisher Ideal or Tornqvist) have an advantage by accommodating for substitution (Triplett 1992). The application of these ideas for this pilot research is that 10-digit classification group indexes are representative of current trade and account for substitution.

For this study, two levels of indexes are calculated – the IPP 10-digit classification groups and stratum level indexes – 5-digit BEA End Use. Two index methods – Laspeyres and Tornqvist – are used to calculate classification group indexes. Stratum level indexes are calculated using a Laspeyres index formula. Index values are then compared against benchmarks. Consistency of trade and outlier treatment, as described below, will only be performed on the Tornqvist indexes based indexes for this study. The two index formulas are below.

Laspeyres Index

$$I_{t,0}^L = \left(\frac{\sum_i P_{i,t} Q_{i,0}}{\sum_i P_{i,0} Q_{i,0}} \right) * 100$$

Tornqvist Index

$$I_{t,0}^T = \left\{ \prod_i \left(\frac{P_{i,t}}{P_{i,0}} \right)^{\frac{(W_{i,0} + W_{i,t})}{2}} \right\} * 100$$

Where:

$$W_{i,t} = \left(\frac{P_{i,t} Q_{i,t}}{\sum_i P_{i,t} Q_{i,t}} \right)$$

$$W_{i,0} = \left(\frac{P_{i,0} Q_{i,0}}{\sum_i P_{i,0} Q_{i,0}} \right)$$

Where:

I^L = Laspeyres Index

I^T = Tornqvist Index
i = trade transaction
t = Current time period
0 = Base period
P = Price
Q = Quantity
W = Weight (share of trade dollar value)

For any type of index (e.g. Laspeyres, Tornqvist), trade weights are needed to calculate the index values. The summed dollar value of all observations for each proxy item in the administrative trade data is the proxy item weight. Using the characteristics that define a proxy item and the IPP's Schedule B-to-IPP classification group mappings for the selected period, item weights are calculated using the transaction data. It should be noted that for index calculation, proxy items map to classification groups. For Laspeyres indexes, 2013 and 2014 annual weights will be calculated at the item level in order to match with the IPP's current Laspeyres trade dollar value based index methodology. For Tornqvist indexes, monthly current and previous weights will be calculated using the total dollar value of the proxy items for the current and previous months.

Proxy item monthly percent changes are calculated as index values using the following formula for Geometric mean prices:

Index Value

$$I_{t,0} = \left(\frac{P_t}{P_0} \right) * 100$$

Where:

I = Index value
t = current time period
0 = base time period
P = Price

These proxy item values are then aggregated using the proxy item index values and item weights, for both the Laspeyres and Tornqvist index formulas, to the classification group level.

Index Refinement. The quality and quantity of prices for each proxy item is intertwined with the quality of the index. Index refinement methods are proposed in this pilot to address two data quality concerns – consistency of trade and outliers. In the official MXPI sampling process, establishment's probability of selection is greater when they historically have traded for more months and quarters. To test how consistent trade affects index quality, pilot indexes are run both with and without the condition that proxy items have at least six prices each year (2015 or 2016). If an item does not meet the minimum consistency requirement within the year, all prices for that proxy item during the year will be removed from the indexes when the condition is imposed.

The pilot indexes are evaluated for the impact of outliers. In the official MXPI price verification process, prices that are beyond set thresholds (typically 2.5 standard deviations from the historical mean) are identified as outliers. The price changes may reflect a real market condition, or a change in the product quality, or even a new item. Each of these situations is treated differently. In the case of outliers in the administrative trade data; proxy item prices may reflect real market changes or be a price for different good that shares the same proxy item characteristics. Whether the change is market driven or product mix driven can not be determined. To test the impact of outliers on index quality, pilot indexes with increasing restrictions on outliers are calculated. Proxy item price changes that fall outside a set number of weighted standard deviations (either 2 or 3) from the weighted mean are removed as outliers. The analysis of the index quality occurs only after the

outlier price is excluded from index calculation. To carry out outlier identification, the test is calculated using the following formulas for each classification group based upon the monthly item index value changes: weighted mean, weighted standard deviation, and the upper and lower price change bounds. The steps in the process are as follows:

Calculate the weighted mean for each classification group for each month:

$$Wt. Mean = \frac{\sum_{i=1}^N \left[W_{i,t} * \left(\frac{I_{i,t}}{I_{i,t-1}} \right) \right]}{\sum_{i=1}^N W_{i,t}}$$

Calculate the weighted standard deviation for the classification group for each month:

$$Wt. STDV = \sqrt{\frac{\sum_{i=1}^N \left\langle W_{i,t} * \left[\left(\frac{I_{i,t}}{I_{i,t-1}} \right) - Wt. Mean \right]^2 \right\rangle}{\left(\frac{N-1}{N} \right) * \sum_{i=1}^N W_{i,t}}}$$

Where (for items within a classification group):

- I = Item Index value
- i = Items with a relative
- t = Current time period
- t-1 = Previous time period
- W = Weight (trade dollar value)
- N = Total Number of Item Relatives

Calculate the upper and lower price change bounds for the classification group for each month:

$$\begin{aligned} Lower Price Change Bound &= Wt. Mean - (number of STDV * STDV) \\ Upper Price Change Bound &= Wt. Mean + (number of STDV * STDV) \end{aligned}$$

Once these standard deviations and limits have been calculated and set in place, indexes are recalculated excluding item prices for the periods that lie outside of the parent (higher-level) classification group's bounds. Thus, the items prices for these months will be imputed.

Imputation of missing data, new data, and nonexistent data. In any time period, when there is a missing price, cell mean imputation is applied to the item without a price quote. A missing price can occur when an item is not traded or simply not reported. Cell mean imputation applies the parent's (higher-level) index percentage change as the child's (lower level) proxy item's price change to continue any index series when no current price exists..New items' index values are also initialized based upon their parent's index values. These new items can only impact the indexes once they have weights, i.e. an actual value of trade. Indexes are calculated for classification groups (IPP-created ten-digit classifications closely related to 10-digit Schedule B codes) based upon the child proxy item data and for strata (BEA classifications 00310 and 00330) based upon the child classification group data. The structure of indexes is graphically as follows:

Figure 1 Tree Structure

Stratum (Classification Group’s Parent, Item’s grandparent)

Classification Group (Item’s Parent, Stratum’s Child)

Item (Classification Group’s Child, Stratum’s grandchild)

An analysis of monthly trade by EIN and proxy item shows that establishments regularly export, but the goods they export are irregularly traded. With irregular, or volatile, trade, records for proxy items are absent across months. The choice to either impute a price – which over time could lead to index bias – or drop a proxy item – which could significantly reduce coverage of an index – is a judgmental decision. The analysis below sets out the approach taken in this research to limit cell mean imputation to under 4 months.

As can be seen in the following table company level EINs vary monthly for 2015 and 2016. Cross-month EIN is a compilation of comparisons counting the instances of an EIN appearing for the BEA strata in the current and previous month (e.g. 201501 and 201502), thus being an average of 23 data points. At first glance, one may think trade is not that volatile since approximately 82% and 78% percent of the companies, respectively, have trade on a monthly basis.

Table 4 EIN Volatility

| BEA Strata | Cross-Month EINs | Monthly EINs | Percentage |
|------------|------------------|--------------|------------|
| 00310 | 1584 | 1943 | 82% |
| 00330 | 3268 | 4165 | 78% |

However, trade transactions are irregular at a more granular level, as demonstrated by the monthly average number of proxy items used in index calculation for BEA classification 00310 (see table 5). For the more detailed proxy items, there are two to three times the number of imputed prices for items than there are actual items. extensive amount of cell mean imputation by a factor of six between new items and estimated this data, it will be shown that index values exhibit bias due to imputation adjustments. To minimize this bias, logic was added to price calculation to limit imputation of like-proxy items to three months. This decision reduces the number of prices used in indexes, because prices in two separate months are needed to start an index series. Given the total number of prices used in the indexes, the overall impact of having to restart some item price index series, may not be significant. This balance between imputation bias and the loss of prices in index calculation is necessary to address the known bias. Volatility still remains in trade/substitution of goods as demonstrated by the counts, but by applying constraints to cell means, the imputation bias has been reduced. The descriptive results of this are shown below, and a stand-alone comparison of continuous imputation vs. constrained imputation is highlighted for Dairy in the benchmark comparison.

Table 5 Item Volatility of Trade for BEA Classification 00310 (Dairy Products and Eggs)

| Item Data | ESZCDQRH | ESZQRH | ESQRH | EQRH | QRH | H |
|---------------------------|----------|--------|-------|------|-----|----|
| New Items | 886 | 473 | 391 | 283 | 4 | 0 |
| Imputed Items | 4025 | 2231 | 1893 | 1418 | 24 | 3 |
| Items Used in Calculation | 1717 | 1248 | 1191 | 1057 | 85 | 37 |

Analyzing the BEA classification 00330 Vegetables with the constraints on imputation logic applied, one can see that the BEA classification 00330 has significantly more trade, and hence more price data.

Table 6 Item Volatility of Trade for BEA Classification 00330 (Vegetables and Vegetable Preparations and Juices)

| Item Data | ESZCDQRH | ESZQRH | ESQRH | EQRH | QRH | H |
|---------------------------|----------|--------|-------|------|-----|-----|
| New Items | 3256 | 1832 | 1367 | 1065 | 14 | 1 |
| Imputed Items | 14564 | 8349 | 6323 | 5047 | 96 | 13 |
| Items Used in Calculation | 5971 | 4278 | 3727 | 3402 | 346 | 148 |

As demonstrated by the data, both the BEA classification 00310 and 00330 exhibit significant evidence of trade volatility/substitution. Limiting imputation to three months provides a balance between potential imputation bias and loss of item price index series.

As with items, if a classification group does not exist, its index value will be initialized based upon its parent's index value. Cell mean imputation will be used to continue any index series when data does not exist. At the classification group level, separate sets of indexes were constructed using both the Laspeyres and Tornqvist index formulas. At the stratum level (BEA five-digit classification) only the Laspeyres index formula was used to construct indexes; the strata calculations were limited to the Laspeyres index formula to replicate existing calculation for IPP's index methodology. The Laspeyres index formula at the stratum level used for this test, however, only used 2014 annual fixed weights, as 2013 data were not available. There is an issue that the IPP will need to address, if this methodology is incorporated into IPP's production indexes in the future. The data provided by the Census Bureau contains the Schedule B code that is appropriate for the month where the trade occurred. However the IPP's mapping structure is lagged, thus not all Schedule B codes will exist in the IPP's mapping structure. For this study the Schedule B code changes all fell in minor revision years, where commonly less than 30 export Schedule B codes of over 8000 are replaced annually. However, in a revision year hundreds of Schedule B codes can be replaced. Given this study is performed using non-revision year data, the loss of price and weight data due to unmapped new Schedule B classifications (and the loss of additional HS data due to related IPP classif group mapping changes) across all indexes should be significantly reduced. BEA classification 00310 is not impacted by any Schedule B changes, but BEA classification 00330 had five Schedule B codes that were replaced of over 100 in 2015 with new Schedule B codes.

Unit Value Bias

The primary issue limiting the use of administrative trade data for price indexes is that the lack of detail results in an average price of dissimilar goods, and that the price per unit, or unit value, will vary widely month to month because of the types of goods that are traded. Unit value bias is a valid concern because changes in the composition of goods in a product area can drive the average price of an aggregate good in a direction that does not accurately reflect the underlying price trend for that product area. Alterman describes the concerns and Silver discusses the potential use of unit value indexes with homogeneous goods (Alterman 1991; Silver 2010). The question facing the IPP is whether creating product groupings from detailed characteristics in the administrative trade data could produce unique enough products that prices and indexes track market changes in price that can proxy the matched-item model. By choosing semi-homogeneous strata with homogeneous goods, we expect like-characteristic products to be similar and have similar prices. Unique

items with like-characteristics can be proxied and tracked over time, imitating the matched-item model with less detailed information but significantly more observations.

In this section we analyze unit value bias in the Dairy and Vegetable trade records data from 201501 to 201612. We use the prices of products created from the item keys described in table 2. Three tests will be applied to the data including a within month item price dispersion test, a cross-month item percentage change test, and a price clustering test. For the price dispersion and cross-month tests, the individual monthly results will be aggregated by category, then the individual monthly categorical numbers (receiving equal weight) will be combined.

Price Dispersion

Testing the dispersion of prices within a month for each proxy item and month is an important consideration of price quality. It is measured by the coefficient of variation. The relationship between the mean and the standard deviation is measured. While variability does not prove unit value bias – because there are examples such as when prices change due to market conditions between the first and the last of the month - variability can be an indicator that unit value bias exists. For example if the mean were 100 and the standard deviation were 80, one would note significant variability in the data. That being said, this specific pricing condition is not happening across the board on a monthly basis for all goods.

The first test that will be applied to the price data within an item key for a month. If more than one price exists for a month, the coefficient of variation (C_v) will be calculated for the item key and month for the price data.

$$C_v = 100 * \frac{Wt. STDV}{Wt. Mean}$$

$$Wt. Mean = \frac{\sum_i P_{i,t} Q_{i,t}}{\sum_i Q_{i,t}}$$

$$Wt. STDV = \sqrt{\frac{\sum_{i=1}^N \langle W_{i,t} * [P_{i,t} - Wt. Mean]^2 \rangle}{\left(\frac{N-1}{N}\right) * \sum_{i=1}^N W_{i,t}}}$$

Where (for prices for an item for a month):

i = Transaction for a proxy item

t = Current time period

P = Transaction Price

W = Transaction Weight (trade dollar value = Price * Quantity)

N = Total Number of Transactions

For analysis purposes the compiled coefficient of variations will be grouped in five-point range categories. For each item key there are two columns. The first column is the percentage for that category and the second column is the cumulative percentage, summing the previous percentage with the new category. Looking first at the BEA classification 00310 coefficient of variation categories, one notices that the more detailed key shows the least variability and the least detailed key shows the greatest amount; that is, the more detailed keys have the highest cumulative percentage (i.e. lower variability) at the lower number categories of the coefficient of variation. This is naturally expected because more detailed products are likely to represent more specific types of trade, but the percentage divergence is what is of greater significance. The data shows that

ESZCDQRH through EQRH are all showing significantly less variability than QRH through H. The cumulative coefficient of variation category of 47.5-52.49 is only 63.6 and 48.6 for QRH and H, respectively, and at a category of 97.5-102.49 is respectively at 85.2 and 79.7. Item keys ESZCDQRH, ESZQRH, ESQRH and EQRH are respectively at 87.6, 82.2, 80.6 and 76.8 with a cumulative coefficient of variation category of 17.5-22.49. Given these statistics, one can surmise from the data that item keys QRH and H for the BEA classification 00310 index are exhibiting significant potential unit value bias.

Table 7 Coefficients of Variation of Transaction Prices for Item Keys within BEA Classification 00310 (Dairy Products and Eggs)

| C. of V. Category | ESZCDQRH | | ESZQRH | | ESQRH | | EQRH | | QRH | | H | |
|-------------------|----------|--------------------|--------|-----------|-------|-----------|------|-----------|------|-----------|------|-----------|
| | Per-cent | Cumulative Percent | Pct. | Cum. Pct. | Pct. | Cum. Pct. | Pct. | Cum. Pct. | Pct. | Cum. Pct. | Pct. | Cum. Pct. |
| 0-2.49 | 45.7 | 45.7 | 35.6 | 35.6 | 33.3 | 33.3 | 29.7 | 29.7 | 5.6 | 5.6 | 0.9 | 0.9 |
| 2.5-7.49 | 22.1 | 67.8 | 21.7 | 57.3 | 21.3 | 54.6 | 20.0 | 49.7 | 4.2 | 9.8 | 0.7 | 1.6 |
| 7.5-12.49 | 9.8 | 77.6 | 12.2 | 69.5 | 12.5 | 67.1 | 12.7 | 62.4 | 3.1 | 13.0 | 1.2 | 2.8 |
| 12.5-17.49 | 5.9 | 83.5 | 7.7 | 77.2 | 8.0 | 75.1 | 8.3 | 70.7 | 5.0 | 18.0 | 2.0 | 4.8 |
| 17.5-22.49 | 4.0 | 87.6 | 5.0 | 82.2 | 5.5 | 80.6 | 6.1 | 76.8 | 5.9 | 23.8 | 4.5 | 9.3 |
| 22.5-27.49 | 2.7 | 90.3 | 3.7 | 85.9 | 4.0 | 84.6 | 4.7 | 81.5 | 5.4 | 29.3 | 6.3 | 15.6 |
| 27.5-32.49 | 1.9 | 92.2 | 2.7 | 88.6 | 2.9 | 87.5 | 3.8 | 85.3 | 7.0 | 36.2 | 5.5 | 21.1 |
| 32.5-37.49 | 1.6 | 93.8 | 2.2 | 90.8 | 2.3 | 89.8 | 2.8 | 88.0 | 8.0 | 44.2 | 7.1 | 28.2 |
| 37.5-42.49 | 1.3 | 95.2 | 1.7 | 92.5 | 1.9 | 91.7 | 2.2 | 90.3 | 7.5 | 51.7 | 7.0 | 35.2 |
| 42.5-47.49 | 0.9 | 96.1 | 1.5 | 94.0 | 1.6 | 93.4 | 1.8 | 92.1 | 5.9 | 57.6 | 5.8 | 41.0 |
| 47.5-52.49 | 0.7 | 96.8 | 1.0 | 95.0 | 1.1 | 94.4 | 1.3 | 93.4 | 6.0 | 63.6 | 7.5 | 48.6 |
| 72.5-77.49 | 0.2 | 98.7 | 0.4 | 97.6 | 0.4 | 97.2 | 0.5 | 96.7 | 1.6 | 79.8 | 1.5 | 71.2 |
| 97.5-102.49 | 0.1 | 99.1 | 0.2 | 98.4 | 0.2 | 98.3 | 0.2 | 97.9 | 0.8 | 85.2 | 0.7 | 79.7 |
| > =102.5 | 0.9 | 100.0 | 1.6 | 100.0 | 1.7 | 100.0 | 2.1 | 100.0 | 14.8 | 100.0 | 20.3 | 100.0 |

Looking at the BEA classification 00330, one sees similar results. Once again item keys ESZCDQRH through EQRH are showing significantly less variability than item keys QRH through H. However, comparing the BEA classification 00330 to BEA classification 00310, there is more variation for item keys ESZCDQRH through EQRH, about the same amount for item key QRH and less variability for item key H for the BEA classification 00330. The cumulative coefficient of variation category of 47.5-52.49 is 63.6 and 50.9 for item keys QRH and H, respectively, and at a category of 97.5-102.49 is respectively at 85.8 and 82.0. Item keys ESZCDQRH, ESZQRH, ESQRH and EQRH are respectively at 76.9, 72.2, 70.5 and 68.7 at the coefficient of variation category of 17.5-22.49. Once again, the variation in the price data for item keys QRH and H, flags a valid concern for unit value bias in these areas. Technically, even the more detailed keys could be demonstrating some unit value bias, but given the limited variation (larger percentages at the lower number categories for the coefficient of variation) shown in the data and the known variability in commodity prices the unit value concern is significantly mitigated.

Table 8 Coefficients of Variation of Transaction Prices for Item Keys within BEA Classification 00330 (Vegetables and Vegetable Preparations and Juices)

| C. of V. Category | ESZCDQRH | | ESZQRH | | ESQRH | | EQRH | | QRH | | H | |
|-------------------|----------|--------------------|--------|-----------|-------|-----------|------|-----------|------|-----------|------|-----------|
| | Per-cent | Cumulative Percent | Pct. | Cum. Pct. | Pct. | Cum. Pct. | Pct. | Cum. Pct. | Pct. | Cum. Pct. | Pct. | Cum. Pct. |
| 0-2.49 | 30.7 | 30.7 | 24.9 | 24.9 | 23.6 | 23.6 | 22.0 | 22.0 | 7.9 | 7.9 | 1.0 | 1.0 |
| 2.5-7.49 | 18.4 | 49.1 | 17.5 | 42.4 | 16.9 | 40.5 | 16.5 | 38.5 | 7.8 | 15.7 | 1.7 | 2.7 |
| 7.5-12.49 | 12.3 | 61.4 | 12.8 | 55.1 | 12.7 | 53.2 | 12.6 | 51.0 | 5.4 | 21.1 | 2.3 | 5.0 |
| 12.5-17.49 | 8.9 | 70.3 | 9.7 | 64.8 | 9.7 | 62.9 | 9.9 | 60.9 | 5.7 | 26.8 | 3.5 | 8.5 |
| 17.5-22.49 | 6.6 | 76.9 | 7.4 | 72.2 | 7.6 | 70.5 | 7.8 | 68.7 | 6.2 | 33.1 | 5.1 | 13.6 |
| 22.5-27.49 | 4.8 | 81.7 | 5.7 | 77.9 | 5.9 | 76.4 | 6.0 | 74.7 | 6.8 | 39.8 | 7.3 | 20.9 |
| 27.5-32.49 | 3.7 | 85.3 | 4.3 | 82.2 | 4.4 | 80.8 | 4.6 | 79.3 | 6.0 | 45.8 | 7.5 | 28.4 |
| 32.5-37.49 | 2.8 | 88.1 | 3.2 | 85.4 | 3.4 | 84.3 | 3.6 | 82.9 | 5.3 | 51.1 | 6.0 | 34.4 |
| 37.5-42.49 | 2.2 | 90.3 | 2.7 | 88.0 | 2.8 | 87.1 | 3.0 | 86.0 | 4.5 | 55.6 | 6.5 | 41.0 |
| 42.5-47.49 | 1.7 | 92.0 | 2.0 | 90.0 | 2.1 | 89.2 | 2.3 | 88.2 | 3.8 | 59.4 | 5.0 | 46.0 |
| 47.5-52.49 | 1.4 | 93.4 | 1.7 | 91.7 | 1.8 | 91.0 | 1.9 | 90.1 | 4.2 | 63.6 | 5.0 | 50.9 |
| 72.5-77.49 | 0.4 | 96.7 | 0.5 | 95.9 | 0.6 | 95.5 | 0.6 | 95.0 | 2.4 | 78.3 | 3.1 | 70.2 |
| 97.5-102.49 | 0.2 | 98.0 | 0.2 | 97.5 | 0.3 | 97.3 | 0.3 | 97.0 | 1.2 | 85.8 | 2.2 | 82.0 |
| > =102.5 | 2.0 | 100.0 | 2.5 | 100.0 | 2.7 | 100.0 | 3.0 | 100.0 | 14.2 | 100.0 | 18.0 | 100.0 |

Cross-Month Comparisons

Another study that was pursued on the data was a cross month item index value percentage change (Index Short Term Ratio) comparison. For this comparison only non-imputed index values for the current month were used for the data from 201502 through 201612. Percentage changes were counted by range category, then equally weighted by month, and these monthly counts were further compiled into aggregate percentage change range categories. As with the coefficient of variation data, this data will be presented in five point categories. If significant price changes are confirmed cross month for an item key, it could be a sign that non-matching items are being priced (product mix is shifting) from month to month. For example, if an item has a transaction price of \$1 this month and has a transaction price of \$1000 the next month, it most likely is not the same item.

For the calculated values that are based upon 23 months of item index percentage change comparisons for all the item keys for the BEA classification 00310, once again the most detailed item key shows the smallest month over month percentage change variability in price. As before, each item key receives two percentage columns, with the second being cumulative. For the 0-2.49 and 2.5-7.49 percentage change categories, ESZCDQRH through EQRH all out perform QRH through H with cumulative percentage changes for category 2.49-7.5 with values of 60.7, 56.5, 55.2, 52.5, 44.3 and 49.7, respectively. However, the differences are not that significant. When the cumulative percentage change category hits 17.5-22.49, H's cumulative value has surpassed all other values, excluding ESZCDQRH with a value starting from the detailed key of 85.1, 82.8, 81.9, 80.2, 76.7, and 84.0. Given the high values for the cross month percentage change data for the BEA classification 00310, there is no observable unit value bias of significance that is demonstrated by the data.

Table 9 Item Key Percentage Change Comparisons within BEA Classification 00310 (Dairy Products and Eggs)

| Percent Change Category | ESZCDQRH | | ESZQRH | | ESQRH | | EQRH | | QRH | | H | |
|-------------------------|----------|--------------------|--------|-----------|-------|-----------|------|-----------|------|-----------|------|-----------|
| | Per-cent | Cumulative Percent | Pct. | Cum. Pct. | Pct. | Cum. Pct. | Pct. | Cum. Pct. | Pct. | Cum. Pct. | Pct. | Cum. Pct. |
| 0-2.49 | 36.2 | 36.2 | 31.0 | 31.0 | 29.8 | 29.8 | 27.6 | 27.6 | 19.2 | 19.2 | 22.0 | 22.0 |
| 2.5-7.49 | 24.5 | 60.7 | 25.4 | 56.4 | 25.5 | 55.2 | 24.9 | 52.5 | 25.0 | 44.3 | 27.7 | 49.7 |
| 7.5-12.49 | 12.3 | 73.0 | 13.2 | 69.5 | 13.2 | 68.5 | 13.7 | 66.2 | 15.3 | 59.6 | 17.1 | 66.9 |
| 12.5-17.49 | 7.3 | 80.4 | 8.2 | 77.7 | 8.4 | 76.9 | 8.7 | 74.9 | 9.4 | 69.0 | 10.2 | 77.1 |
| 17.5-22.49 | 4.7 | 85.1 | 5.1 | 82.8 | 5.1 | 81.9 | 5.4 | 80.2 | 7.7 | 76.7 | 6.9 | 84.0 |
| 22.5-27.49 | 3.2 | 88.3 | 3.7 | 86.5 | 3.8 | 85.7 | 4.0 | 84.2 | 4.3 | 81.1 | 4.1 | 88.1 |
| 27.5-32.49 | 2.3 | 90.5 | 2.5 | 89.0 | 2.7 | 88.4 | 2.9 | 87.1 | 3.5 | 84.5 | 2.9 | 91.0 |
| 32.5-37.49 | 1.9 | 92.4 | 1.9 | 90.9 | 2.1 | 90.5 | 2.3 | 89.4 | 1.8 | 86.3 | 1.8 | 92.8 |
| 37.5-42.49 | 1.3 | 93.6 | 1.5 | 92.5 | 1.5 | 92.0 | 1.6 | 91.0 | 2.0 | 88.3 | 1.8 | 94.6 |
| 42.5-47.49 | 1.1 | 94.7 | 1.2 | 93.7 | 1.3 | 93.3 | 1.3 | 92.3 | 1.5 | 89.8 | 1.2 | 95.8 |
| 47.5-52.49 | 0.7 | 95.4 | 0.9 | 94.6 | 1.0 | 94.2 | 1.1 | 93.5 | 0.7 | 90.5 | 0.3 | 96.0 |
| >=52.5 | 4.6 | 100.0 | 5.4 | 100.0 | 5.8 | 100.0 | 6.5 | 100.0 | 9.5 | 100.0 | 4.0 | 100.0 |

Examining the other index area of interest, BEA classification 00330, one finds similar results. For the 0-2.49 percentage change category ESZCDQRH through EQRH, all outperform QRH and H with cumulative percentage changes with values of 28.8, 25.2, 25.1, 24.1, 18.3 and 19.3, respectively. With the 2.5-7.49 percentage change category, H is outperforming all other item keys excluding ESZCDQRH with values starting from the most detailed item key of 48.7, 45.6, 45.7, 44.6, 43.3 and 46.8, respectively. By the 7.5-12.49 percentage change category, H is outperforming all others. The values in the same item key order are as follows: 60.5, 58.0, 58.3, 57.5, 58.9, and 62.6. Given the data, even if one item key is outperforming another the values are all high and similar, with no item key presenting an obvious unit value bias of concern.

Table 10 Item Key Percentage Change Comparisons within BEA Classification 00330 (Vegetables and Vegetable Preparations and Juices)

| Percent Change Category | ESZCDQRH | | ESZQRH | | ESQRH | | EQRH | | QRH | | H | |
|-------------------------|----------|--------------------|--------|-----------|-------|-----------|------|-----------|------|-----------|------|-----------|
| | Per-cent | Cumulative Percent | Pct. | Cum. Pct. | Pct. | Cum. Pct. | Pct. | Cum. Pct. | Pct. | Cum. Pct. | Pct. | Cum. Pct. |
| 0-2.49 | 28.8 | 28.8 | 25.2 | 25.2 | 25.1 | 25.1 | 24.1 | 24.1 | 18.3 | 18.3 | 19.3 | 19.3 |
| 2.5-7.49 | 19.9 | 48.7 | 20.4 | 45.6 | 20.5 | 45.7 | 20.6 | 44.6 | 25.0 | 43.3 | 27.5 | 46.8 |
| 7.5-12.49 | 11.8 | 60.5 | 12.5 | 58.0 | 12.6 | 58.3 | 12.8 | 57.5 | 15.6 | 58.9 | 15.9 | 62.6 |
| 12.5-17.49 | 8.2 | 68.7 | 8.6 | 66.6 | 8.7 | 67.0 | 8.9 | 66.3 | 9.7 | 68.6 | 10.9 | 73.5 |
| 17.5-22.49 | 6.1 | 74.8 | 6.5 | 73.1 | 6.6 | 73.6 | 6.7 | 73.0 | 6.7 | 75.3 | 6.1 | 79.6 |
| 22.5-27.49 | 4.6 | 79.4 | 5.1 | 78.3 | 5.0 | 78.6 | 5.1 | 78.1 | 5.3 | 80.7 | 3.9 | 83.5 |
| 27.5-32.49 | 3.8 | 83.2 | 3.8 | 82.1 | 4.0 | 82.6 | 4.1 | 82.2 | 3.7 | 84.3 | 2.9 | 86.4 |
| 32.5-37.49 | 3.0 | 86.3 | 3.2 | 85.3 | 3.2 | 85.7 | 3.2 | 85.3 | 2.3 | 86.6 | 2.1 | 88.6 |
| 37.5-42.49 | 2.5 | 88.7 | 2.4 | 87.7 | 2.4 | 88.1 | 2.5 | 87.8 | 2.1 | 88.7 | 1.8 | 90.4 |
| 42.5-47.49 | 1.9 | 90.7 | 2.1 | 89.8 | 2.0 | 90.1 | 2.1 | 89.9 | 1.8 | 90.5 | 1.6 | 92.0 |
| 47.5-52.49 | 1.6 | 92.2 | 1.7 | 91.5 | 1.7 | 91.8 | 1.7 | 91.7 | 1.4 | 91.9 | 1.1 | 93.1 |
| > =52.5 | 7.8 | 100.0 | 8.5 | 100.0 | 8.2 | 100.0 | 8.3 | 100.0 | 8.1 | 100.0 | 6.9 | 100.0 |

Clustering

A third method was used to determine if unit value bias exists. Ward's Minimum-Variance Method was applied to any proxy item that had 100 or more transactions from 201501 to 201612. Using Ward's Minimum-Variance Method the data are clustered for each proxy item to determine how many clusters exist for each item based upon the transaction prices. The optimal number of clusters for each proxy item should be one if there is no unit value bias, as the item's price represents one good only. Valid reasons can exist where an item exhibits more than one price cluster. For example, a price shock for a good could create two valid clusters for a proxy item. However, in most conditions one would expect to see one cluster for one item.

The output for BEA classification 00310 is displayed by item key with the number of clusters, the total number of proxy items that have that number of clusters, and the percentage of items that fall in the number of clusters.

Table 11 Cluster data for BEA Classification 00310 (Dairy Products and Eggs)

| # of Clusters | ESZCDQRH | | ESZQRH | | ESQRH | | EQRH | | QRH | | H | |
|---------------|----------|-------|--------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | Total | Pct. | Total | Pct. | Total | Pct. | Total | Pct. | Total | Pct. | Total | Pct. |
| 1 | 191 | 82.33 | 225 | 77.05 | 232 | 78.38 | 224 | 80.00 | 48 | 82.76 | 31 | 88.57 |
| 2 | 29 | 12.50 | 48 | 16.44 | 46 | 15.54 | 39 | 13.93 | 5 | 8.62 | 1 | 2.86 |
| 3 | 8 | 3.45 | 12 | 4.11 | 13 | 4.39 | 10 | 3.57 | 3 | 5.17 | 1 | 2.86 |
| 4 | 2 | 0.86 | 2 | 0.68 | 2 | 0.68 | 5 | 1.79 | 1 | 1.72 | 1 | 2.86 |
| >4 | 2 | 0.86 | 5 | 1.71 | 3 | 1.01 | 2 | 0.71 | 1 | 1.72 | 1 | 2.86 |

Table 12 Cluster data for BEA Classification 00330 (Vegetables and Vegetable Preparations and Juices)

| # of Clusters | ESZCDQRH | | ESZQRH | | ESQRH | | EQRH | | QRH | | H | |
|---------------|----------|-------|--------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | Total | Pct. | Total | Pct. | Total | Pct. | Total | Pct. | Total | Pct. | Total | Pct. |
| 1 | 1003 | 76.74 | 1084 | 79.59 | 1044 | 78.09 | 995 | 78.16 | 150 | 77.32 | 89 | 74.79 |
| 2 | 200 | 15.30 | 191 | 14.02 | 204 | 15.26 | 186 | 14.61 | 33 | 17.01 | 19 | 15.97 |
| 3 | 71 | 5.43 | 64 | 4.70 | 62 | 4.64 | 61 | 4.79 | 9 | 4.64 | 9 | 7.56 |
| 4 | 22 | 1.68 | 18 | 1.32 | 17 | 1.27 | 25 | 1.96 | 1 | 0.52 | 2 | 1.68 |
| >4 | 11 | 0.84 | 5 | 0.37 | 10 | 0.75 | 6 | 0.47 | 1 | 0.52 | 0 | 0.00 |

Looking at the data one notices that most data for all item keys fall into one cluster. When this data and the price dispersion (Coefficient of Variation) data are reviewed, based upon the statistical tests for the least detailed keys (QRH and H) there is a large price dispersion that typically forms one cluster for most items. For all other keys, once again one typically finds one cluster, but significantly less price dispersion.

Given the three tests, one can state that item keys QRH and H for the BEA classification 00310 and 00330 are exhibiting significant potential for unit value bias based upon their price dispersions.

Benchmark Comparisons

For this study a range of indexes were calculated for the two dimensions – First on item keys, and then additionally with options to address potential bias in the measures, including consistency of trade, volatility of trade, existence of outliers, and the quality differences in Laspeyres and Tornqvist indexes. These options provide a wide range of measures against which the benchmark measures will be compared. The statistical tests carried out here are an effort to pin down the index method/item key combination that shows the least biased and best fit. Data results will be presented using the following index keys, and will be calculated for each item key to compare against benchmarks. With the benefit of hindsight, the two BEA End Use indexes that were selected for research based upon their index quality concerns, meant that there were no solid benchmarks of export price indexes against which to gauge the quality of the administrative price indexes. Nonetheless, the effort to carry out this work will be a key aspect to validate future research.

Table 13 Index Keys

| Index Keys | Description |
|------------|---|
| LL | Laspeyres index formula was used to calculate the stratum lower. Laspeyres index formula was used to calculate the classification group. |
| LL4 | Laspeyres index formula was used to calculate the stratum lower. Laspeyres index formula was used to calculate the classification group. If an item is Cell Mean Imputed for a fourth month, the index is no longer imputed. |
| LT4 | Laspeyres index formula was used to calculate the stratum lower. Tornqvist index formula was used to calculate the classification group. If an item is Cell Mean Imputed for a fourth month, the index is no longer imputed. |
| LT42 | Laspeyres index formula was used to calculate the stratum lower. Tornqvist index formula was used to calculate the classification group. If an item is Cell Mean Imputed for a fourth month, the index is no longer imputed. Outliers outside the 2nd standard deviation are removed. |
| LT43 | Laspeyres index formula was used to calculate the stratum lower. Tornqvist index formula was used to calculate the classification group. If an item is Cell Mean Imputed for a fourth month, the index is no longer imputed. Outliers outside the 3rd standard deviation are removed. |
| LT46 | Laspeyres index formula was used to calculate the stratum lower. Tornqvist index formula was used to calculate the classification group. If an item is Cell Mean Imputed for a fourth month, the index is no longer imputed. Only include items in 2015 with at least 6 usable prices in the 12 month period. Only include items in 2016 with least 6 usable prices in the 12 month period. |
| LT462 | Laspeyres index formula was used to calculate the stratum lower. Tornqvist index formula was used to calculate the classification group. If an item is Cell Mean Imputed for a fourth month, the index is no longer imputed. Outliers outside the 2nd standard deviation are removed. Only include items in 2015 with at least 6 usable prices in the 12 month period. Only include items in 2016 with least 6 usable prices in the 12 month period. |
| LT463 | Laspeyres index formula was used to calculate the stratum lower. Tornqvist index formula was used to calculate the classification group. If an item is Cell Mean Imputed for a fourth month, the index is no longer imputed. Outliers outside the 3rd standard deviation are removed. Only include items in 2015 with at least 6 usable prices in the 12 month period. Only include items in 2016 with least 6 usable prices in the 12 month period. |

Three statistical tests of each of these index keys were run against the benchmarks.

- Correlation Coefficients
- Root Mean Squared Errors
- Mean Absolute Error

Correlation coefficients (r) were calculated for the percentage changes (Short Term Ratios, or STRs) using the benchmarks' values for the specified year and month and the related pilot percentage change to determine how strong the relationship is between the data. For this test a value of one is the preferred value, with values closer to one showing a strong positive correlation.

$$r = \frac{\sum_{t=1}^n (Pilot\ Index\ STR_t - \overline{Pilot\ Index\ STR}) (Benchmark\ STR_t - \overline{Benchmark\ STR})}{\sqrt{\sum_{t=1}^n (Pilot\ Index\ STR_t - \overline{Pilot\ Index\ STR})^2 \sum_{t=1}^n (Benchmark\ STR_t - \overline{Benchmark\ STR})^2}}$$

Root Mean Squared Errors were calculated comparing the pilot indexes' percentage changes to the benchmark indexes percentage changes. This test is impacted more by outliers given that the numerator is squared. To minimize the impact of outliers the Mean Absolute Errors can be used. For this test a zero is the preferred value, with values closer to zero showing less prediction error.

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (Pilot\ STR_t - Benchmark\ STR_t)^2}{n}}$$

Mean Absolute Errors were also calculated comparing the pilot indexes monthly percentage changes to the benchmark indexes percentage changes. For this test a zero is also preferred value, with values closer to zero showing less prediction error.

$$MAE = \frac{\sum_{t=1}^n |Pilot\ STR_t - Benchmark\ STR_t|}{n}$$

Official indexes

Spot prices, Consumer Price Indexes (CPI) (U.S. city average, all urban consumers, not seasonally adjusted), Producer Price Indexes (PPI), and Import/Export Price indexes (XPI for Export Price Indexes) were all considered as possible comparative measures. It was determined that the XPI and CPI indexes were the benchmark indexes of choice for BEA classifications 00310 and 00330 respectively. Given that the Export Prices indexes are sampled and weighted based upon trade, these indexes in theory should most accurately reflect trade prices. Even if there are lower response rates in any part of the data collection process, whether sampling, initiation or on-going collection, one would hope the bias of not receiving what was sampled will be outweighed by the fact that the sample is trade-based and the indexes are trade weighted. However, there is a known weighting problem with the IPP BEA classification 00330 index, because the real world monthly seasonality of vegetable production is much more variable than the classification group trade weight – which is fixed constant value for the entire year. In a month where the trade is limited at the beginning or end of a particular vegetable's season a price change based upon limited trade still receives the annual weight. From a price setting standpoint for vegetables the price an exporter can charge is impacted by the world price. Given that domestically most vegetables are available year-round, the impact of this monthly weighting issue should be mitigated with the CPI index, thus justifying its use as a benchmark.

Results for BEA Classification 00310 (Dairy Products and Eggs)

Below are the results for the statistical test for the BEA classification 00310 using the XPI index as the benchmark. A CPI proxy index was also tested as a benchmark and was calculated using published CPI values and aggregated using IPP trade weights. The root mean squared errors and mean absolute errors values were about the same for the two benchmarks, however, the correlation coefficients that were calculated using the CPI index percentage changes were not supportive for using this data as a benchmark. It should be noted

that as outliers are removed one would expect these indexes to have lower root mean squared errors and mean absolute errors, since the variance is reduced. However, if bias is introduced when the outliers are removed, the root mean squared errors and mean absolute errors may not decrease as expected. This should be considered when reviewing these results.

Table 14 Correlation Coefficients for BEA Classification 00310 using an XPI Index Benchmark

| Index Key | ESZCDQRH | ESZQRH | ESQRH | EQRH | QRH | H |
|-----------|----------|--------|-------|------|------|------|
| LL | 0.39 | 0.33 | 0.34 | 0.35 | 0.18 | 0.46 |
| LL4 | 0.46 | 0.47 | 0.52 | 0.39 | 0.19 | 0.47 |
| LT4 | 0.61 | 0.59 | 0.60 | 0.58 | 0.50 | 0.48 |
| LT42 | 0.57 | 0.60 | 0.60 | 0.57 | 0.52 | 0.50 |
| LT43 | 0.59 | 0.60 | 0.62 | 0.60 | 0.51 | 0.50 |
| LT46 | 0.59 | 0.58 | 0.60 | 0.61 | 0.50 | 0.48 |
| LT462 | 0.57 | 0.53 | 0.60 | 0.64 | 0.52 | 0.50 |
| LT463 | 0.57 | 0.57 | 0.58 | 0.62 | 0.53 | 0.50 |

Table 15 Root Mean Squared Errors for BEA Classification 00310 using an XPI Index Benchmark

| Index Key | ESZCDQRH | ESZQRH | ESQRH | EQRH | QRH | H |
|-----------|----------|--------|-------|------|------|------|
| LL | 2.62 | 2.82 | 2.64 | 2.73 | 4.89 | 2.67 |
| LL4 | 2.56 | 2.49 | 2.28 | 2.60 | 4.72 | 2.66 |
| LT4 | 1.82 | 1.90 | 1.91 | 2.00 | 2.61 | 2.71 |
| LT42 | 1.96 | 1.97 | 2.00 | 2.07 | 2.58 | 2.61 |
| LT43 | 1.88 | 1.96 | 1.90 | 2.02 | 2.55 | 2.61 |
| LT46 | 1.96 | 2.03 | 2.04 | 1.99 | 2.59 | 2.72 |
| LT462 | 2.04 | 2.22 | 2.07 | 1.99 | 2.56 | 2.61 |
| LT463 | 2.07 | 2.08 | 2.08 | 2.05 | 2.56 | 2.61 |

Table 16 Mean Absolute Errors for BEA Classification 00310 using an XPI Index Benchmark

| Index Key | ESZCDQRH | ESZQRH | ESQRH | EQRH | QRH | H |
|-----------|----------|--------|-------|------|------|------|
| LL | 2.11 | 2.08 | 2.06 | 1.99 | 3.97 | 2.13 |
| LL4 | 1.97 | 1.77 | 1.59 | 1.83 | 3.78 | 2.12 |
| LT4 | 1.44 | 1.35 | 1.45 | 1.57 | 2.07 | 2.16 |
| LT42 | 1.60 | 1.50 | 1.47 | 1.53 | 2.09 | 2.10 |
| LT43 | 1.50 | 1.50 | 1.43 | 1.50 | 2.06 | 2.10 |
| LT46 | 1.54 | 1.48 | 1.52 | 1.53 | 2.10 | 2.18 |
| LT462 | 1.57 | 1.65 | 1.52 | 1.52 | 2.10 | 2.11 |
| LT463 | 1.58 | 1.67 | 1.63 | 1.53 | 2.11 | 2.11 |

As mentioned previously in the paper, an imputation bias was discovered in the index values when no limits were placed on how long an item could be imputed. Using a Laspeyres index for classification group index calculation with similar weighting, one would expect a closer correlation with the XPI indexes. However, when there are no limits to imputation, these indexes received the lowest correlation coefficients, and the root mean squared errors and mean absolute errors were generally the higher values. All other data used the 4 month imputation limitation, so this bias should not be an issue of significance for the other index keys.

Analyzing the correlation coefficient data, it is evident that when the Tornqvist index formula is used, the coefficients are higher, and the root mean squared errors and mean absolute errors are lower, all results supporting the use of the Tornqvist index formula. With the Tornqvist indexes, new items are being brought into the indexes on a regular basis, and the weights are also current. These results align the Tornqvist index closer with the benchmark index. The benchmark IPP index will incorporate new items as they enter the index and weight them with the lagged trade data that they represent. Additionally, when items are discontinued they may be replaced in the IPP's indexes. The Tornqvist indexes address a concern in the literature related to capturing substitution of goods in the indexes (Moulton 2018).

On the other hand, Laspeyres indexes versus the benchmark showed weaker results; this could be due to the fact that for the Laspeyres index formula, substitution is not occurring for new goods traded in 2015 and 2016. For the Tornqvist indexes, the root mean squared errors and mean absolute errors generally increased with the removal of outliers for the more detailed item keys. This is not the expected result when the variance is reduced. The result raises a concern that the information lost by removing the outliers is adding bias to the indexes.

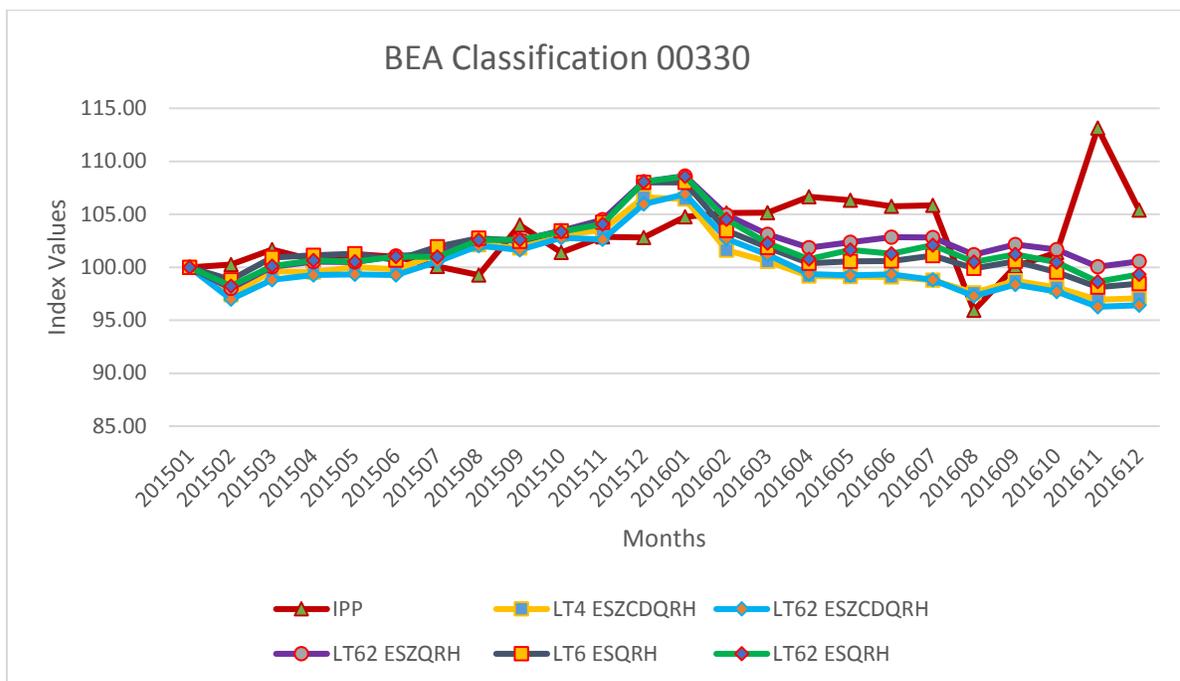
From an item key standpoint, the QRH and H keys which previously were identified as having a potential unit value bias issue are statistically not performing as well as other item keys for any of the statistical tests. This leaves the four more-detailed item key (ESZCDQRH – EQRH) Tornqvist indexes for further analysis. There are three index key and item key combinations that stand out. Given that the root mean squared error and mean absolute error are higher for LT462-EQRH (and these statistics should improve with the removal of outliers), this option should be eliminated even though it had the highest correlation coefficient. Across the board the removal of outliers did not have the expected impact on the root mean squared errors and mean absolute errors, which brings into question if outliers should be removed. Upon comparing the results, the following two indexes are recommended as the best aligned with the benchmark.

Table 17 Recommended Indexes for BEA Classification 00310

| | LT4- ESZCDQRH | LT43- ESQRH |
|-----------------|------------------|----------------|
| Corr. Coeff. | 0.61 | 0.62 |
| RMSE | 1.82 | 1.90 |
| MAE | 1.44 | 1.43 |

Results for BEA Classification 00330 (Vegetables and Vegetable Preparations and Juices)

A similar review was also performed on the BEA classification 00330 export index. For this index area the XPI did not prove to be a good benchmark with almost all correlation coefficients being negative and of higher values for the root mean squared errors and mean absolute errors. Given that there is significant monthly seasonality in the weights and the XPI index is constructed using a fixed annual weight, item prices that should receive minimal index weight when the product is out of the growing season in the U.S. are receiving greater weight than what is supported by the trade in the given month. This weighting issue is the likely explanation for this poor match. When the index values are reviewed graphically, one notices that the XPI indexes values for the first year reflect the test index values (that are later selected). In the second year of the pilot the data trends show significantly more divergence with generally higher index values and larger fluctuations in index values.



The CPI proved to be a better benchmark, by having consistent demand and steady weights, as consumers can buy most vegetables at any time during the year in the U.S. Below are the results for the BEA classification 00330.

Table 18 Correlation Coefficients for BEA Classification 00330 using a CPI Index Benchmark

| Index Key | ESZCDQRH | ESZQRH | ESQRH | EQRH | QRH | H |
|-----------|----------|--------|-------|------|------|------|
| LL4 | 0.26 | 0.26 | 0.24 | 0.29 | 0.39 | 0.35 |
| LT4 | 0.35 | 0.29 | 0.23 | 0.24 | 0.48 | 0.37 |
| LT42 | 0.39 | 0.35 | 0.37 | 0.35 | 0.31 | 0.32 |
| LT43 | 0.39 | 0.40 | 0.34 | 0.35 | 0.30 | 0.32 |
| LT46 | 0.33 | 0.37 | 0.35 | 0.32 | 0.37 | 0.26 |
| LT462 | 0.47 | 0.46 | 0.45 | 0.40 | 0.33 | 0.33 |
| LT463 | 0.41 | 0.43 | 0.38 | 0.38 | 0.32 | 0.33 |

Table 19 Root Mean Squared Errors for BEA Classification 00330 using a CPI Index Benchmark

| Index Key | ESZCDQRH | ESZQRH | ESQRH | EQRH | QRH | H |
|-----------|----------|--------|-------|------|------|------|
| LL4 | 2.16 | 2.15 | 2.18 | 2.09 | 2.01 | 2.42 |
| LT4 | 1.86 | 2.02 | 2.13 | 2.07 | 1.92 | 2.37 |
| LT42 | 1.79 | 1.84 | 1.82 | 1.82 | 2.03 | 2.02 |
| LT43 | 1.82 | 1.79 | 1.86 | 1.82 | 2.02 | 2.02 |
| LT46 | 1.92 | 1.79 | 1.84 | 1.92 | 2.07 | 2.50 |
| LT462 | 1.69 | 1.67 | 1.73 | 1.79 | 1.99 | 2.00 |
| LT463 | 1.84 | 1.75 | 1.82 | 1.83 | 1.98 | 2.00 |

Table 20 Mean Absolute Errors for BEA Classification 00330 using a CPI Index Benchmark

| Index Key | ESZCDQRH | ESZQRH | ESQRH | EQRH | QRH | H |
|-----------|----------|--------|-------|------|------|------|
| LL4 | 1.66 | 1.64 | 1.68 | 1.76 | 1.61 | 1.97 |
| LT4 | 1.34 | 1.60 | 1.68 | 1.67 | 1.51 | 1.94 |
| LT42 | 1.34 | 1.41 | 1.45 | 1.45 | 1.50 | 1.56 |
| LT43 | 1.39 | 1.42 | 1.49 | 1.41 | 1.49 | 1.56 |
| LT46 | 1.43 | 1.41 | 1.40 | 1.53 | 1.57 | 2.04 |
| LT462 | 1.33 | 1.31 | 1.40 | 1.44 | 1.47 | 1.55 |
| LT463 | 1.44 | 1.42 | 1.42 | 1.45 | 1.46 | 1.55 |

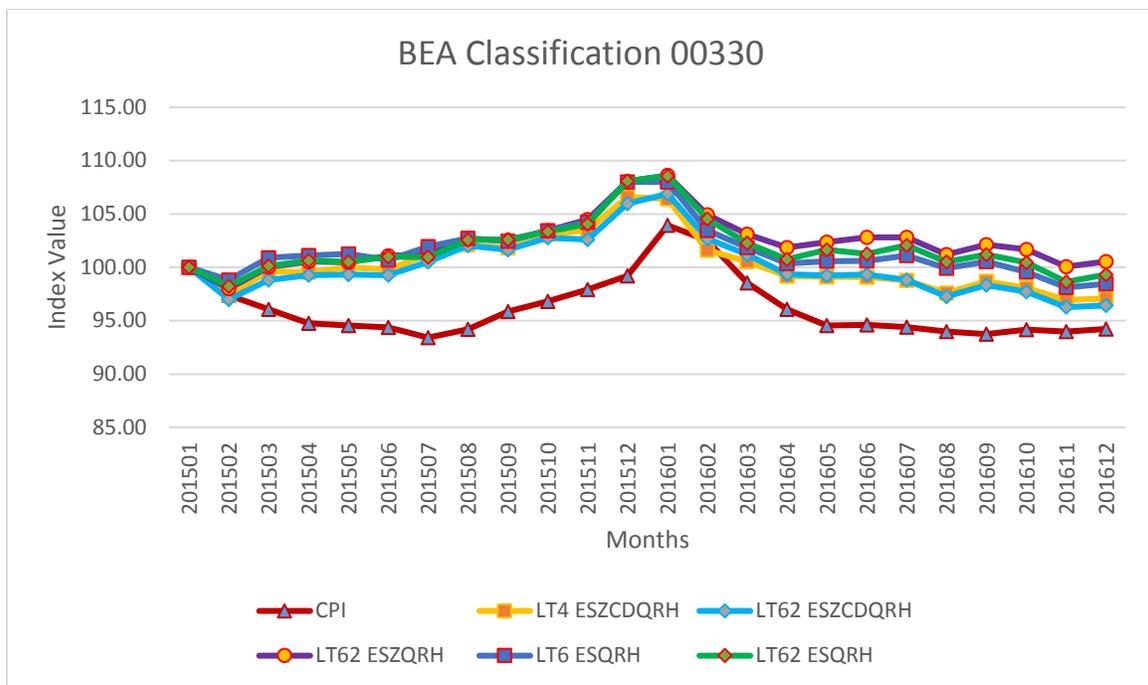
Evaluating the results, LT4-QRH has the highest correlation coefficient and the root mean squared error and mean absolute error perform better than other keys. However, based upon previous work in the paper, QRH appears to be impacted by unit value bias. Furthermore, it is surpassed by LT4-ESZCDQRH. Given these conditions, LT4-QRH is removed from selection. LT462-ESZCDQRH has the second highest correlation coefficient followed by LT462-ESZQRH and LT462-ESQRH, respectively. LT4-ESZCDQRH and LT46-ESQRH stand out, in that without removing outliers they have lower mean absolute error and root mean squared errors, only insignificantly surpassed by indexes that have outliers removed. LT462-ESQRH has a better correlation coefficient than both of these. This leaves the following five index keys to consider:

Table 21 Recommended Indexes for BEA Classification 00330

| | LT4- ESZCDQRH | LT462- ESZCDQRH | LT462- ESZQRH | LT46- ESQRH | LT462- ESQRH |
|-----------------|------------------|--------------------|------------------|----------------|-----------------|
| Corr. Coeff. | 0.35 | 0.47 | 0.46 | 0.37 | 0.45 |
| RMSE | 1.86 | 1.69 | 1.67 | 1.79 | 1.73 |
| MAE | 1.34 | 1.33 | 1.31 | 1.41 | 1.40 |

Unlike the 00310 BEA classification, the root mean squared errors and mean absolute errors more often exhibited lesser values when outliers were removed, as would be expected.

Examining the data graphically for these various indexes against the benchmark, they exhibit the same overall trend. The CPI index has a more pronounced rise to the high point then falls thereafter. At the end of this movement the index values show limited dispersion.



Special Case: Detailed Study of BEA Classification 00310

Given the challenges of finding a solid benchmark against which to compare the administrative price indexes, there is an alternative way to see how well the BEA classification 00310 index fits the benchmark. An analysis of the largest classification group in the index - “Concentrated milk and cream in solid form with a fat content less than or equal to 1.5%” – shows that this classification group carries significantly more weight than any other classification group index. The indexes from this milk and cream classification group will be compared to IPP’s unpublished comparable classification group index as the benchmark.

Table 22 Correlation Coefficients for Concentrated Milk and Cream in Solid Form 10-digit classification group

| Index Key | ESZCDQRH | ESZQRH | ESQRH | EQRH | QRH | H |
|-----------|----------|--------|-------|------|------|------|
| LL | 0.70 | 0.56 | 0.52 | 0.56 | 0.15 | 0.62 |
| LL4 | 0.69 | 0.54 | 0.53 | 0.56 | 0.24 | 0.62 |
| LT4 | 0.59 | 0.64 | 0.65 | 0.65 | 0.63 | 0.62 |
| LT42 | 0.62 | 0.66 | 0.66 | 0.65 | 0.64 | 0.62 |
| LT43 | 0.60 | 0.65 | 0.66 | 0.65 | 0.64 | 0.62 |
| LT46 | 0.59 | 0.63 | 0.64 | 0.65 | 0.63 | 0.62 |
| LT462 | 0.61 | 0.63 | 0.64 | 0.65 | 0.64 | 0.62 |
| LT463 | 0.60 | 0.63 | 0.65 | 0.65 | 0.64 | 0.62 |

Table 23 Root Mean Squared Errors for Concentrated Milk and Cream in Solid Form

| Index Key | ESZCDQRH | ESZQRH | ESQRH | EQRH | QRH | H |
|-----------|----------|--------|-------|------|------|------|
| LL | 2.93 | 3.55 | 3.83 | 3.68 | 6.06 | 3.54 |
| LL4 | 2.97 | 3.60 | 3.79 | 3.66 | 5.37 | 3.54 |
| LT4 | 3.26 | 3.10 | 3.20 | 3.22 | 3.49 | 3.54 |
| LT42 | 3.12 | 3.04 | 3.18 | 3.19 | 3.42 | 3.54 |
| LT43 | 3.21 | 3.07 | 3.18 | 3.19 | 3.42 | 3.54 |
| LT46 | 3.42 | 3.17 | 3.26 | 3.22 | 3.42 | 3.54 |
| LT462 | 3.29 | 3.16 | 3.30 | 3.20 | 3.43 | 3.54 |
| LT463 | 3.36 | 3.21 | 3.24 | 3.21 | 3.43 | 3.54 |

Table 24 Mean Absolute Errors for Concentrated Milk and Cream in Solid Form

| Index Key | ESZCDQRH | ESZQRH | ESQRH | EQRH | QRH | H |
|-----------|----------|--------|-------|------|------|------|
| LL | 2.48 | 2.88 | 3.02 | 2.93 | 3.94 | 2.81 |
| LL4 | 2.50 | 2.95 | 3.03 | 2.94 | 3.68 | 2.81 |
| LT4 | 2.79 | 2.57 | 2.59 | 2.58 | 2.75 | 2.81 |
| LT42 | 2.74 | 2.45 | 2.52 | 2.54 | 2.71 | 2.81 |
| LT43 | 2.82 | 2.47 | 2.53 | 2.55 | 2.71 | 2.81 |
| LT46 | 3.05 | 2.62 | 2.63 | 2.60 | 2.71 | 2.81 |
| LT462 | 2.91 | 2.57 | 2.64 | 2.56 | 2.73 | 2.81 |
| LT463 | 2.98 | 2.62 | 2.58 | 2.58 | 2.73 | 2.81 |

Analyzing the data, it must be noted that ESZCDQRH using a Laspeyres index at the classification group level (LL and LL4) best matches the XPI classification group index. Given that IPP methodology uses a Laspeyres index formula at the classification group level and similar lagged weights, this should not be a surprise. What is of significance is how the detailed item key ESZCDQRH outperforms the other Laspeyres item keys. This strengthens the argument that a detailed item key better emulates the matched-item model.

It should be noted that the quality of the classification group Tornqvist indexes is overall closer to the XPI classification group benchmarks than the Laspeyres classification group indexes. Laspeyres classification group indexes and QRH and H Laspeyres and Tornqvist classification group indexes in general have the worst root mean squared errors. When the significance of outliers is minimized using the mean absolute error, ESZCDQRH indexes joins the Laspeyres classification group indexes (excluding Laspeyres classification group indexes for item key ESZCDQRH) and QRH and H Laspeyres and Tornqvist classification group indexes with higher values.

When the milk and cream classification group index is compared to the BEA End Use 00310 index (the parent), there are significant similarities in the data result. One of the most noticeable differences is for the Laspeyres index with the ESZCDQRH item key. The BEA classification index 00310 did not exhibit as strong of a correlation coefficient as did the milk and cream classification group. This is likely an artifact of the accuracy of the IPP milk and cream classification group based upon its sampled coverage. Not all the classification groups in the index are given the support that the milk and cream classification group receives, and that is possibly why the parent BEA classification group does not have as high of a correlation coefficient. One would expect a comparison of two similar Laspeyres indexes to produce similar results.

Lessons Learned, Potential Scope, and Homework

The results of this research creating detailed prices and indexes based on administrative trade data are promising. Indexes derived and tests conducted as part of the study, substantiate the potential for use of export transaction data in lieu of directly-collected survey information for semi-homogeneous areas with the use of appropriate index methodologies. Based upon the limited index areas that were studied (BEA classification 00310 and 00330), the following general concepts are proposed given the results:

- The Tornqvist index formula should be used instead of the Laspeyres index formula for calculating classification group indexes due to use of current period weights, which allows for almost immediate product substitution.
- More detailed item keys show closer approximation to the matched item model, and show less risk of unit value bias. The most promising item keys ESZCDQRH, ESZQRH, ESQRH and EQRH.
- Additional research and analysis should be performed on measuring the impacts of eliminating outliers and on consistent exporters. It is difficult to conclusively make a determination on these methodologies based upon the results.
- More semi-homogeneous areas should be tested and analyzed to ascertain if there is one approach that can be applied all semi-homogeneous area.

A key objective of this pilot is to determine if this methodology can be expanded to other five-digit BEA classifications to expand index publication and quality of deflators. Only one of the three statistical tests of the 00310 and 00330 BEA End Use indexes, the coefficient of variation test, had previously demonstrated that unit value bias may exist within certain item keys. Given the effectiveness of the coefficient of variation test, it will be used to determine the potential reach and scope of other export five-digit BEA classifications – to see whether the approach could potentially apply the pilot’s methodology. The estimates provided give an approximate idea for how extensively these concepts may be expanded and provides some evidence that these indexes are not subject to unit value bias.

Starting at the low end of acceptability, given the volatility inherent in the 00330 classification of vegetables, the coefficient of variation values for item key ESZCDQRH will be used as foundational benchmark to test the potential scope of indexes that could be covered by this new approach – the assumption being that the natural price variation in 00330 puts its values at the outer limits of acceptable values. We lowered the

original cumulative percent values for 00330 by five to add flexibility for other volatile classifications that have similar cumulative percentages but exhibit slightly different rates of increase between coefficient of variation categories. If the five-digit BEA classification had a cumulative percent that is greater than all the cumulative percent thresholds for the coefficient of variation categories it would pass the test.

Table 25 Coefficient of Variation Cumulative Percent Thresholds

| Coefficient of Variation Category | Cumulative Percent Threshold based upon Item Key ESZCDQRH Coeff. of Var. Values – 5 |
|-----------------------------------|---|
| 0-2.49 (0 category) | 26.0 |
| 2.5-7.29 (5 category) | 44.0 |
| 7.5-12.49 (10 category) | 56.0 |
| 17.5-22.49 (20 category) | 72.0 |
| 27.5-32.49 (30 category) | 80.0 |
| 42.5-47.49 (45 category) | 87.0 |

Additional tests were done to see if some strata met a reduced benchmark standard using a lower rate of increase (original cumulative percent values were lowered by ten) for the following categories: 0 and 5, 10 and 20 or 30 and 45. These four tests were run for the following item keys ESZCDQRH, EQRH, and QRH with the following results summed by the 1-digit BEA 0, 1, 2, 3 and 4 classifications. Each table lists the number of 5-digit indexes, by BEA classification and test constraint, whether it would pass the coefficient of variation test. The subtotal of each column sums to 129 in each table, the total number of 5-digit BEA End Use export classifications.

Table 26 Coefficient of Variation Tests for Item Key ESZCDQRH within Export Five-Digit BEA Classifications

| Five-digit BEA Classification | Pass Coefficient of Variation - 5 tests | Pass C. of Var. - 5 tests with 0 & 5 using -10 | Pass C. of Var. - 5 tests with 10 & 20 using -10 | Pass C. of Var. - 5 tests with 30 & 45 using -10 | Fail all Coefficient of Variation Tests |
|--|---|--|--|--|---|
| BEA 0 Classification - Foods, Feeds, & Beverages | 16 | 0 | 1 | 0 | 1 |
| BEA 1 Classification - Industrial Supplies & Materials | 27 | 0 | 0 | 3 | 17 |
| BEA 2 Classification - Capital Goods | 2 | 0 | 0 | 1 | 30 |
| BEA 3 Classification - Automotive Vehicles, Parts & Engines | 3 | 0 | 0 | 0 | 3 |
| BEA 4 Classification - Consumer Goods, Excluding Automotives | 2 | 0 | 0 | 0 | 23 |
| Total | 50 | 0 | 1 | 4 | 74 |

Table 27 Coefficient of Variation Tests for Item Key EQRH within Export Five-Digit BEA Classifications

| Five-digit BEA Classification | Pass Coefficient of Variation - 5 tests | Pass C. of Var. - 5 tests with 0 & 5 using -10 | Pass C. of Var. - 5 tests with 10 & 20 using -10 | Pass C. of Var. - 5 tests with 30 & 45 using -10 | Fail all Coefficient of Variation Tests |
|--|---|--|--|--|---|
| BEA 0 Classification - Foods, Feeds, & Beverages | 9 | 0 | 1 | 0 | 8 |
| BEA 1 Classification - Industrial Supplies & Materials | 14 | 0 | 0 | 0 | 33 |
| BEA 2 Classification - Capital Goods | 0 | 0 | 0 | 0 | 33 |
| BEA 3 Classification - Automotive Vehicles, Parts & Engines | 0 | 0 | 0 | 0 | 6 |
| BEA 4 Classification - Consumer Goods, Excluding Automotives | 0 | 0 | 0 | 0 | 25 |
| Total | 23 | 0 | 1 | 0 | 105 |

Table 28 Coefficient of Variation Tests for Item Key QRH within Export Five-Digit BEA Classifications

| Five-digit BEA Classification | Pass Coefficient of Variation - 5 tests | Pass C. of Var. - 5 tests with 0 & 5 using -10 | Pass C. of Var. - 5 tests with 10 & 20 using -10 | Pass C. of Var. - 5 tests with 30 & 45 using -10 | Fail all Coefficient of Variation Tests |
|--|---|--|--|--|---|
| BEA 0 Classification - Foods, Feeds, & Beverages | 0 | 0 | 0 | 0 | 18 |
| BEA 1 Classification - Industrial Supplies & Materials | 0 | 2 | 0 | 0 | 45 |
| BEA 2 Classification -Capital Goods | 0 | 0 | 0 | 0 | 33 |
| BEA 3 Classification - Automotive Vehicles, Parts & Engines | 0 | 0 | 0 | 0 | 6 |
| BEA 4 Classification - Consumer Goods, Excluding Automotives | 0 | 0 | 0 | 0 | 25 |
| Total | 0 | 2 | 0 | 0 | 127 |

Analyzing the results, based upon the coefficient of variation tests for item key ESZCDQRH, 50 (55 based upon the criteria with greater flexibility) five-digit BEA classifications are not demonstrating any clear unit value bias. That is a significant number of export classifications that potentially could be calculated using this methodology. Based upon the data, the higher technology goods are much less suited for this methodology. For item key EQRH, 23 five-digit BEA classifications, and for item key QRH, 0 five-digit BEA classifications are also not demonstrating any clear unit value bias. Looking at these results, it is clear that with a more detailed item key, significantly less unit value bias is observed.

Table 29 Publication and Sampling Status for Item Key ESZCDQRH for the 50 Export Five-Digit BEA Classifications that passed the Cumulative Coefficient of Variation Tests

| Five-Digit BEA Classification | Unpublished with only sampled items | Unpublished and contains non-sampled items | Published with only sampled items | Published and contains non-sampled items |
|--|-------------------------------------|--|-----------------------------------|--|
| BEA 0 Classification - Foods, Feeds, & Beverages | 4 | 0 | 8 | 4 |
| BEA 1 Classification - Industrial Supplies & Materials | 10 | 1 | 10 | 6 |
| BEA 2 Classification - Capital Goods | 2 | 0 | 0 | 0 |
| BEA 3 Classification - Automotive Vehicles, Parts & Engines | 3 | 0 | 0 | 0 |
| BEA 4 Classification - Consumer Goods, Excluding Automotives | 1 | 0 | 1 | 0 |
| Total | 20 | 1 | 19 | 10 |

Analyzing the 50 five-digit BEA classifications that met the coefficient of variation's thresholds, 21 of these strata are unpublished and 29 are already published. Of these 50 strata, 11 already use some form of non-sampled data, and ten of the 11 are published indexes. This leaves 40 indexes as potential candidates for administrative trade data index estimation.

Conclusion

This study shows that the proxy item indexes with more detailed item keys exhibit less price dispersion, smaller cross-month price differences, and greater clustering, which provide reasonable evidence that unit value bias can be addressed by creating 'quasi-products' with detailed item keys. The second dimension of index quality was more challenging to address, as we identified a systemic approach to compare indexes. But the flawed decision to select product categories that had poor quality indexes meant that benchmark comparison results were not definitive. The BEA End Use 00310 Dairy index - with limits on imputation and with a Tornqvist index formula at the classification level - were a better fit with the comparable published export price index benchmark. The BEA End Use 00330 Vegetable index constructed only with consistent exporters where price outliers were excluded had a better overall fit compared to the CPI benchmark. However, greater root mean squared errors and mean absolute errors for Dairy 00310 were evidenced when excluding outliers. This outlier related discrepancy deserves further analysis with more product categories to determine whether excluding outliers typically creates bias.

IPP plans on expanding this pilot beyond the two studied index areas into other homogeneous and semi-homogeneous areas for exports and imports. Additionally, we will collaborate with the BEA to calculate and analyze the impact of multi-year detailed export price indexes not presently published as deflators for GDP. This research would require the expansion of the test index time frame beginning with 2012. Given the potential increase in the number of detailed BEA indexes the IPP could publish, there may be significant improvement in accuracy when deflating the import and export GDP numbers, demonstrating the potential significance of this work. Based on preliminary results, this study suggests that the unit value bias can be overcome, making administrative trade data for semi-homogeneous products a possible source for calculation of import and export price indexes.

- Alterman, William. 1991. "Price Trends in U.S. Trade, New Data, New Insights." Chapter in NBER book International Economic Transactions: Issues in Measurement and Empirical. Edited by Peter Hooper and J. David Richardson. National Bureau of Economic Research, Studies in Income and Wealth. Volume 55. University of Chicago Press 109-143 (January).
- Clausing, Kimberly A. 2003. "Tax-motivated transfer pricing and US intrafirm trade prices." Journal of Public Economics. Volume 87, Issues 9-10, 2207-2223 (September).
- Diewert, W.E. 1976. "Exact and Superlative Index Numbers. Journal of Econometrics" Volume 4, Issue 2. North-Holland Publishing Company (May).
- Gopinath, Gita. 2010. "Currency Choice and Exchange Rate Pass-through," American Economic Review 100 (1): 304-336.
- Houseman, Susan, Christopher Kurz, Paul Lengermann, and Benjamin Mandel. "Offshoring Bias in U.S. Manufacturing." Journal of Economic Perspectives. Volume 25, Number 2, 111-132 (Spring).
- Moulton, Brent R. 2018. "The Measurement of Output, Prices, and Productivity: What's Changed Since the Boskin Commission?" Hutchins Center on Fiscal & Monetary Policy at Brookings, The Brookings Institution (July). <https://www.brookings.edu/research/the-measurement-of-output-prices-and-productivity>.
- Silver, Mick. 2010. "The Wrongs and Rights of Unit Value Indices." Review of Income and Wealth. Series 56, Special Issue 1, 206-223 (June).
- Triplett, Jack E. 1992. "Economic Theory and BEA's Alternative Quantity and Price Indexes." Survey of Current Business. 49-52 (April). <https://www.bea.gov/scb/pdf/national/nipa/1992/0492trip.pdf>.