Review of Income and Wealth
Series 56, Special Issue 1, June 2010

THE WRONGS AND RIGHTS OF UNIT VALUE INDICES

BY MICK SILVER*

International Monetary Fund

There are two sides to unit value indices. The first is their widespread and unwarranted use based on customs data as surrogates for price relatives of product groups of exports and imports. For the aggregation of heterogeneous items, superlative price indices are best and unit value indices are biased. The second is the widespread and unwarranted use of price indices for homogeneous items, for which unit value indices are best and superlative price indices are biased. Then there is the "in-between": the case of broadly comparable items. This paper examines such issues.

1. INTRODUCTION

The problem with unit value indices is that we hear a lot about their use for trade price indices and little about their use elsewhere, when in fact we should be hearing much less about their use for trade price indices and a lot more elsewhere. This paper first argues that unit value indices based on customs information and widely used as surrogates for export and import price indices can seriously misrepresent inflation in traded products. Moreover, the use of unit value indices leads to even more serious errors in terms of trade indices. Unit value indices should appropriately be used only for homogeneous items. It will be argued that with increasing product differentiation as well as the dwindling availability of customs data, the use of unit value indices based on customs data is a disservice to price measurement. This position is in line with the recommendations for best practice given in ILO *et al*. (2009), the *Export and Import Price Index Manual*, where it is suggested that countries using unit value indices move to establishment survey-based price indices, and provides a strategy for doing so. This is a departure from international recommendations (United Nations, 1981) that are nearly 30 years old.[1]

Note: The views expressed herein are those of the author and should not be attributed to the IMF, its Executive Board, or its management. Any errors or omissions are the author's responsibility.

*Correspondence to Mick Silver, Statistics Department, IMF, 1900 Pennsylvania Avenue, Washington DC 20431, USA (msilver@imf.org).

[1]The United Nations (1983) provides case studies on the development and implementation of the two main approaches—the survey pricing approach as used by the Federal Republic of Germany and the unit-value approach as used by Norway—to assist countries in initiating and developing their trade price change measures.

The advice given in United Nations (1981) is reiterated in United Nations (2004, paragraph 15, customs documents and price indices based on survey data). The relative strengths and weaknesses of these two approaches to index number compilation are described in United Nations (1981): "Although price indices are generally preferred, in practice countries may not have the resources available to compile that information. *It is recommended that all countries produce and publish volume (quantum) indices and either unit value or price indices for their total imports and exports on a monthly, quarterly and annual basis*" (emphasis in original).

Second, the paper argues that there are many areas in index number measurement where the use of unit value indices *is* appropriate. Index number theory advocates the use of superlative price index number formulas, including Fisher and Törnqvist price indices, as target indices for heterogeneous products and services. However, for homogeneous products and services,[2] it is well recognized that superlative price indices can be misleading and unit value indices are the appropriate target index.[3] There is, however, a continuum between homogeneous and heterogeneous products and a need to examine the issue of which index number formulas are appropriate for products within this continuum. For example, there are many products that are only slightly differentiated to meet niche markets or to attain competitive advantage. The paper discusses the case of broadly comparable items and takes up the problem of the aggregation of their price changes. The paper is organized as follows. Section 2 outlines the properties of unit value and superlative index numbers. The use of unit value indices for trade price indices is considered in Section 3. Section 4 considers the use of superlative indices, or price indices such as the Laspeyres and Paasche, for homogeneous products. The interesting case of broadly-comparable products is considered in Section 5. Section 6 concludes.

2. The Properties of Unit Value and Superlative Indices in Index Number Construction

2.1. *Superlative Index Numbers*

The Fisher, P_F, and Törnqvist, P_T, index number formulas are two commonly used superlative indices.[4] The Fisher price index is a geometric mean of Laspeyres, P_L, and Paasche, P_P, price indices and is defined for a price comparison between the current period t and a reference period 0, over $m = 1, \ldots, M$ matched items whose respective prices and quantities are given by p_m^t and q_m^t for period t, and p_m^0 and q_m^0 for period 0, by:

$$P_F \equiv \sqrt{\frac{\sum_{m=1}^{M} p_m^t q_m^0}{\sum_{m=1}^{M} p_m^0 q_m^0} \times \frac{\sum_{m=1}^{M} p_m^t q_m^t}{\sum_{m=1}^{M} p_m^0 q_m^t}} = \sqrt{P_L \times P_P}. \tag{1}$$

The Törnqvist price index is defined as:

[2]Hereafter "products" and "commodities" include goods and services.

[3]As outlined in *CPI*, *PPI*, and *XMPI Manuals* (ILO *et al*., 2004a, chapters 17 and 20; ILO *et al*., 2004b, chapters 17–18 and 20; ILO *et al*., 2009, chapters 18 and 21; Commission of the European Communities *et al*., 2008, chapter 15).

[4]The Walsh price index is a less commonly used superlative index that is similar to a Laspeyres or Paasche price index, but uses a geometric mean of period 0 and t quantities as the fixed basket quantities (ILO *et al*., 2004a, chapter 15, paragraphs 15.24–32).

$$(2) \qquad P_T = \prod_{m=1}^{M} \left(\frac{p_m^t}{p_m^0} \right)^{(s_m^0 + s_m^t)/2} \quad \text{where } s_m^t = p_m^t q_m^t \Big/ \sum\nolimits_m p_m^t q_m^t \text{ and } s_m^0 = p_m^0 q_m^0 \Big/ \sum\nolimits_m p_m^0 q_m^0.$$

Both P_F and P_T make symmetric use of each period's price and quantity information. Diewert (1976, 1978), using an approach based on economic theory, demonstrated that both Fisher and Törnqvist indices belong to a class of *superlative indices*[5] that have the desirable property of incorporating substitution effects. That is, these indices can accommodate the effect of consumers, say, shifting more of their purchases toward goods with relatively low price increases, thereby lowering their cost of living. Laspeyres and Paasche indices are associated with aggregator functions that cannot accommodate substitution effects.

In the test or axiomatic approach, desirable properties for an index number are chosen and different formulas are then evaluated against them. Fisher described his index as "ideal" because it satisfied the tests proposed, including the "time reversal" and "factor reversal" tests.[6] The Fisher index has also been justified from a fixed basket approach. It is apparent from the Laspeyres and Paasche price indices that constitute equation (1) that both indices hold fixed the basket of quantities. The formulas differ in that the Laspeyres holds the basket fixed in the reference period while the Paasche holds the basket fixed in the current period. Neither formula can be judged superior to the other; yet they can yield different results. A compromise solution for the price index is to use a formula that makes symmetric use of the base and current period information on quantities. The Fisher index can be shown to be the most suitable in this regard (ILO *et al.*, 2004a, chapter 15).

Thus all three approaches favor the Fisher index, in particular, and superlative indices more generally, since they produce very similar results to the Fisher index. In practice of course Laspeyres-type indices are often calculated because data on current period information are not available.[7] The arguments presented in this paper apply as much to the use of unit value indices versus Laspeyres-type price indices as to the Fisher price index.

[5]Aggregator functions underlie the definition of indices in economic theory. For example, a utility function is used to define a constant utility cost of living index. Different index number formulas can be shown to correspond to different functional forms of the aggregator function. For example, a Laspeyres index corresponds to a highly restrictive Leontief aggregator function. The underlying functional forms for superlative indices, including Fisher and Törnqvist, are flexible: they are second-order approximations to other (twice-differentiable) functional forms around the same point. It is the generality of functional forms that superlative indices represent that allows them to accommodate substitution behavior and hence to be considered desirable index number formulas.

[6]The time reversal test requires that the index for period t compared with period 0, should be the reciprocal of that for period 0 compared with t. The factor reversal test requires that the product of the price index and the volume index should be equal to the value ratio.

[7]In practice, especially for CPIs where timeliness is of the essence, the price reference period 0 differs from the earlier weight reference period, say b, since it takes time to compile the results from the survey of households, establishments, or other sources for the weights used in the index, or because the survey that the weights are based on is conducted less frequently than the interval at which prices are collected. The Laspeyres index given by the first component in the right-hand-side expression in equation (1) may have quantities in period b instead of 0. This index is a Lowe index (see ILO *et al.*, 2004a, chapter 15).

2.2. *Unit Value Indices*

A unit value index, P_U, for commodity group i, for period t relative to a reference period 0 is given for comparison over $m = 1, \ldots, M$ prices, p_m^t, and quantities, q_m^t, in period t and over $n = 1, \ldots, N$ prices, p_n^0, and quantities, q_n^0, in period 0 where $m \in G_i$ and $n \in G_i$ by:

$$P_{U,i}\left(p^0, p^t, q^0, q^t\right) \equiv \left(\frac{\sum_{m=1}^{M} p_m^t q_m^t}{\sum_{m=1}^{M} q_m^t}\right) \Bigg/ \left(\frac{\sum_{n=1}^{N} p_n^0 q_n^0}{\sum_{n=1}^{N} q_n^0}\right). \tag{3}$$

Higher level indices aggregate $P_{U,i}$ over the i commodity groups use standard index number formulas such as Laspeyres and Fisher indices (ILO *et al.*, 2004a, chapter 15). If the items whose prices are being aggregated are identical—that is, perfectly homogeneous—a unit value index has desirable properties. Balk (2005) identifies it as the target index for homogeneous goods.

Consider the case where the exact same item is sold at different prices during different times of year, say lower sales volumes and higher prices in the first week of each month and higher sales volumes and lower prices in the last week of each month. The unit value for the monthly index solves the time aggregation problem and appropriately gives more weight in the aggregate to the lower prices than the higher ones. If the elementary unit value index in equation (3) is used as a price index to deflate a corresponding change in the value, the result is a change in total quantity which is intuitively appropriate, i.e.

$$\frac{\sum_{m=1}^{M} p_m^t q_m^t}{\sum_{n=1}^{N} p_n^0 q_n^0} \Bigg/ \left[\left(\frac{\sum_{m=1}^{M} p_m^t q_m^t}{\sum_{m=1}^{M} q_m^t}\right) \Bigg/ \left(\frac{\sum_{m=1}^{M} p_m^0 q_m^0}{\sum_{m=1}^{M} q_m^0}\right)\right] = \frac{\sum_{m=1}^{M} q_m^t}{\sum_{m=1}^{M} q_m^0}. \tag{4}$$

Note that the summation of quantities in the top and bottom of the right-hand-side of equation (4) must be of the exact same type of item for the expression to make sense.

Balk (1998) showed that the unit value index satisfies the conventional index number tests with the exceptions of:

(i) The *Proportionality Test*: $P(p,\lambda p,q^0,q^t) = \lambda$ for $\lambda > 0$; that is, if all prices are multiplied by the positive number λ, then the new price index is λ. The unit value index only satisfies the proportionality test in the unlikely event that relative quantities do not change.

(ii) The *Identity or Constant Prices Test*: $P(p,p,q^0,q^t) = 1$, a special case of (i); that is, if the price of every good is identical during the two periods, then the price index should equal unity, no matter what the quantity vectors are. The unit value index only satisfies the identity test if the relative quantities stay the same—that is, the composition of the products compared does not change.

(iii) The *Invariance to Changes in the Units of Measurement (commensurability) Test*: $P(\alpha_1 p_1^0, \ldots, \alpha_n p_n^0; \alpha_1 p_1^1, \ldots, \alpha_n p_n^1; \alpha_1^{-1} q_1^0, \ldots, \alpha_n^{-1} q_n^0; \alpha_1^{-1} q_1^1, \ldots, \alpha_n^{-1} q_n^1) = P(p_1^0, \ldots, p_n^0; p_1^1, \ldots, p_n^1; q_1^0, \ldots, q_n^0; q_n^1 \ldots, q_n^1)$ for all $\alpha_1 > 0, \ldots, \alpha_n > 0$; that is, the price index does not change if the units of measurement for each product are changed.

Changes in units *de facto* arise when the quality of items change. These tests were devised for the aggregation of heterogeneous items and are not meaningful for homogeneous items. However, the commensurability test *is* satisfied in the homogeneous case when product items are truly identical. For example, in Section 4.1 we outline the case where prices do not change, but a shift in quantities switches the average price to a lower level, leading to a fall in the overall price level: a meaningful failure of the identity test.

Bradley (2005) takes a cost-of-living index defined in economic theory and compares the bias that results from using unit values as "plug-ins" for prices. He finds that if there is no price dispersion in either the current or reference period compared, the unit value (plug-in) index will not be biased against the theoretical index. This case can be subsumed under that of homogeneous items. If there is price dispersion in the reference (current) period, but not the current period, a unit value "plug-in" would have an upwards (downward) bias, and if there is price dispersion in both periods, there is a guarantee (there is a zero probability that the condition of no bias holds for any arbitrary data generating process) that there will be a bias in the "plug-in" unit value index but one cannot sign that bias. Balk (1998) finds the unit value index to be appropriate for a cost of living index only if the underlying preference ordering can be represented by a restrictive simple sum utility function in which the utility from $q_1^\tau, q_2^\tau, \ldots, q_m^\tau$ for periods $\tau = 0,t$ is given by the unweighted $\sum_m q_m^\tau$.

Given these properties of superlative and unit value indices, we now turn to examine the context within which each should be used. Section 3 considers the improper use of unit value indices, as opposed to superlative or other price indices, in trade price indices based on customs data, while sections 4 and 5 consider more generally for price indices the improper use of superlative or other price indices, as opposed to unit value indices.

3. Unit Value Indices Based on Customs Data

3.1. *Sources of Bias*

Unit value indices derived from data collected by customs authorities are frequently used by countries as surrogates for price changes at the elementary level of aggregation (United Nations, 2005). The following are grounds upon which unit value indices based on customs data might be deemed unreliable:

- Bias arises from compositional changes in quantities and quality mix of what is exported or imported. Even with best practice stratification,[8] the scope for reducing this sort of bias is limited due to the sparse variable

[8]Párniczky (1974) shows that it does not follow that such breakdowns are always beneficial to a unit value index.

list—class of (quantity) size of the order and trader, country of origin/destination—available on customs documents. Product differentiation and branding and a turnover in the composition of differentiated items traded each month can lead to bias.

- For unique and complex commodities such as ships and specialized machinery, model pricing can be used in establishment-based surveys where the respondent is asked each period to price a commodity such as a machine with specified fixed characteristics. Characteristics information on products cannot be incorporated into unit value indices and hence changes in the specification of the commodity for a unit value index will bias the index.
- Methods for appropriately dealing with quality change,[9] temporarily missing values, and seasonal goods can be employed using establishment-based survey information and index measures that allow for substitution but not with unit value indices.
- The information on quantities in customs returns, and the related matter of the choice of units in which the quantities are measured, have been found in practice to be seriously problematic, though the former has improved with the adoption of ASYCUDA.[10]
- With customs unions, countries often collect very limited intra-area trade data.
- An increasing proportion of trade is in services and by e-trade and is not subject to customs documentation.
- Unit value indices rely to a large extent on outlier detection and deletion. Given the stickiness of many price changes, such deletions run the risk of missing large price catch-ups when they take place, and understating inflation.
- Valuation issues can be directly addressed at the establishment level to an extent not possible using customs data.

A main advantage of the use of unit value indices is their coverage and relatively low resource cost. However, the unit values used are drawn as non-random samples in the sense that they often exclude: products traded irregularly; products that have no quantity reported; products that have low-value shipments; and, where outlier detection is automatically and badly applied,[11] erratic month-to-month changes.

The alternative data source is the well established method used for consumer and producer price indices. This is to use as source data the observed/reported prices of a sample of representative items from representative establishments for which detailed information on their quality characteristics and terms of sale are recorded so that the prices of like in each month are compared with like—the matched model method. Establishment-based surveys can be quite representative,

[9]Von der Lippe (2007) shows that adjustments for quality change are one reason why price indices are less volatile than unit value indices.

[10]The Automated System for Customs Data (ASYCUDA) project of the United Nations Conference on Trade and Development (UNCTAD).

[11]It is recognized that the appropriate response is good outlier detection, which in turn requires collaboration with establishments for confirmation as to whether a large change is real or not.

especially if a small number of wholesalers or establishments are responsible for much of the total value of imports or exports and, with the important assumption of cooperation, are a cost-effective source of reliable data.[12] Further, good sampling, can, by definition, realize accurate price change measures and finally, the *value* shares of exports and imports, obtained from customs data, can form the basis of information for weights for establishment-based survey price data.

CPI and PPI compilation practices have benefited from much research and experience since the publication of United Nations (1981), and extensive guidelines on good practice are available in the *CPI* and *PPI Manuals* (ILO *et al*., 2004a and 2004b, respectively) and the forthcoming *XMPI Manual* (ILO *et al*., 2009).[13]

3.2. *Evidence of Bias*

We briefly outline comparisons of price and unit value indices when they co-exist. Evidence is presented below, first at the aggregate level, and then, in Section 3.3 (under hybrid indices) evidence at a disaggregated level is considered.

Angermann (1980) compared price index number changes with unit value changes for exports from and imports to the Federal Republic of Germany. Between 1970 and 1976 the Paasche-version price index for exports increased by 38.6 percent compared with 34.3 percent measured by the corresponding unit value indices. The discrepancy between import price indices and corresponding unit value indices was greater, at 45.8 and 33.1 percent, respectively. Angermann also found that when unit value indices were used to calculate the terms of trade effect, there was a gain in 1976 of 1.4 billion DM to real income, at 1970 prices, compared with a loss of 6.6 billion DM when using a price index that has been reworked into a Paasche form.

Alterman (1991) compared export and import price changes between March 1985 and June 1989 for the U.S. as measured by unit value indices and by the price indices based on establishment surveys that replaced them.[14] For imports, over this period, the price index increased by 20.8 percent and the unit value index increased by 13.7 percent. For exports, the figures were much closer; 13.0 and 12.2 percent for the price and unit value index, respectively. Some of the difference between the two series may be attributed to their use of different periods for weights. However, when price indices were recalculated using the same weights as the unit value indices, the differences were exacerbated: a 20.6 percent and 16.4 percent increase for the import and export price indices, respectively.

A review in 1992 of the unit value methodology used by the U.K. led to their change in May 1996 to establishment-based trade price indices, following similar changes in methodology by the U.S., Japan, and Germany (Ruffles and William-

[12]Of course, such a group of wholesalers or other establishments will only be a reliable source of data in the absence of collusion among these establishments.

[13]Chapters 2–14 of ILO *et al.* (2009) provides practical advice on the compilation of export and import price indices (XMPIs), and chapter 16 onwards provides a detailed account of the theory of price indices and issues in their construction including choice of formula, the treatment of transfer prices, seasonal goods and services, quality change, and terms of trade measurement. Chapter 1 provides an overview of the manual and, more particularly, its annex provides a step by step account of how to start compiling XMPIs.

[14]The official U.S. unit value indices were discontinued by the Bureau of the Census in 1989, so these figures are the latest available estimates.

son, 1997). The annual averages of export prices in 1995, compared with 1994, increased by 6.6 percent using price indices compared with 8.1 percent using unit value indices.

Such findings are not new. Kravis and Lipsey (1971) found that the prices of manufactured goods exported by developed countries to developing countries had risen over about 20 years by 75 per cent, as compared to 14 per cent using unit value indices. Kravis and Lipsey (1985) found a decrease in the terms of trade of manufactured relative to all primary products between 1953 and 1976 of over 36 per cent, using price indices, which is almost a quarter greater than that suggested by the unit value indices (28 per cent). With a further correction for quality change, the price data suggested a fall in manufacture terms of trade of over 45 per cent, which is more than 50 per cent greater than the corresponding figure based on unit value indices.

Silver (2009a) compared unit value indices and the corresponding price indices for both Germany and Japan for exports and imports using monthly data for 1996:7 to 2006:9. Unit value indices were found to seriously misrepresent price indices in the sense that discrepancies between unit value indices and the price indices were substantial; changes could not be relied upon to have the same sign; there was no evidence of long-run (cointegrating) relationships between the values for the price indices and unit value indices; and the unit value indices were of little help in predicting the price index values. The findings held for both month-on-month and month-on-12 month changes. Moreover, the marked unreliability of unit value indices as measures of export and import price inflation was surpassed by the unreliability of the terms of trade (ToT) indices based upon them. For example, the mean month-on-month discrepancy for ToT changes for Germany was 1.3 percent compared with discrepancies of 1.1 and 0.9 percent, respectively, for imports and exports. A discrepancy of, for example, 1.1 percent implies that if the month-on-month change in the price index was zero, indicating no change, then the unit value index would indicate a change of 1.1 percent on average. Such discrepancies can be regarded as seriously misleading for economists. The discrepancy for individual months can, of course, be much larger than the mean discrepancy, as reflected by an associated standard deviation of 1.0 to a maximum of 7.3 percentage points for Germany's month-on-month index changes for imports. The problem is not just an issue of the magnitude of error, but also of direction. For about 25 percent of month-on-month comparisons, the signs differ; that is, in one quarter of comparisons the economist would read that prices were rising (falling) when they were falling (rising).

Silver (2009a) also found that the discrepancy arising from the use of unit value indices to calculate the terms of trade *effect*, rather than a price index, was most marked. For example, Japan's trade balance in 2005 of 6956 billion Yen was eliminated by the adverse change in its terms of trade when using price indices, but only halved when using unit value indices. The values of exports and imports of Germany and Japan were deflated over the period from 1999 to 2005 by corresponding unit value indices and then price indices, and the results were compared. It was found, for example, that the volume of exports by Japan increased by 50 percent when a unit value deflator was used, but the increase was halved when a price index was used.

3.3. *What Can Be Done?*

Improve Unit Value Indices

Strategies for improving unit value indices by using stratification by trader, country of origin/destination, size of shipment, and type of commodity, as well as improved detection and deletion (after follow-up) of outliers can improve price measurement that is based on unit value indices. However, the nature of customs documentation and the natural volatility of sticky price changes and pass through rates constrain the benefits of increased stratification and outlier detection.

Hybrid Indices: A Staged Gradualist Approach

Countries whose resources are limited and that use unit value indices may compile hybrid indices that utilize unit value indices only for those subaggregates with little compositional change and price indices elsewhere. There is thus a need to identify the particular commodity classes less prone to unit value bias. Reports by PLANISTAT Europe that were commissioned by Eurostat to examine the methodology for import price indices used by EU member states,[15] analyzed disaggregated CPA[16] data for Finland, the Netherlands, and Sweden and found that:

> Any list of CPA categories for which UVIs [unit value indices] are a priori acceptable as proxies for SPIs [import price indices] would be very short, especially as regards monthly data. It would include almost only aggregates and raw materials, even if sizable discrepancies between unit value indices and SPIs are deemed acceptable. (Decoster, 2003b, p. 9)

Silver (2009a) came to an even less sympathetic conclusion using disaggregated German data.

It may be thought that unit values of detailed commodity groups will be well defined and have little dispersion. The most detailed level for XMPIs is the ten-digit level of the Harmonized System (HS).[17] An example is the Harmonized Code 6402991815: *tennis shoes, basketball shoes, gym shoes, training shoes and the like*: a group whose unit value change is highly susceptible to composition changes. Note that detailed specifications for a price index would enable the average price changes to be measured for specific, representative brands and models of shoes exported/imported by establishments under specified terms of sale. The price dispersion of unit values within detailed HS groups is substantial. WTO *et al.* (2007, technical appendix B) cites as an example the product 04041006,[18] imported

[15]See Decoster (2003a, 2003b).

[16]Statistical Classification of Products by Activity in the European Economic Community, 1996 version, available at: http://forum.europa.eu.int/Public/irc/dsis/bmethods/info/data/new/classifications/cpa_en.pdf.

[17]The Harmonized Commodity Description and Coding System (HS) or Harmonized System, as it is commonly known, was introduced in 1988 by the World Customs Organization, as a replacement for the Standard International Trade Classification (SITC). Descriptions are common across all countries down to the six-digit level; however, for statistical or tariff purposes, countries are allowed to include additional digits on a country specific basis.

[18]Whey and modified whey, in powder, granules, or other solid forms, without added sugar or other sweetening matter, of a protein content nitrogen content X 6.38′ OF ≤15 per cent by weight and a fat content by weight of >27 per cent.

by the EU. Before the exclusion of extreme values, the 65 unit values in this group ranged from 39 US$/tonne to 111,250 US$/tonne, with an average of 4021 US$. After the elimination of six extreme values, the range was reduced to a still substantial 39–3212 US$/tonne, with an average of 1282 US$/tonne. WTO *et al*. (2007) provides summary information on the dispersion of unit values for goods subject to non-ad valorem duties (NAV). There were about 28,000 NAVs present in the tariff schedule in 2006 for which unit values were calculated. The summary results for the ratio of the lower (Q_1) to upper (Q_3) quartiles for the about 15,000 unit values for NAVs calculated at the more tightly specified tariff line level[19] are given in WTO *et al*. (2007, technical appendix B). About 54 per cent of such unit values had a ratio that was 0.5 or less, with this figure falling to 33 per cent when outliers were removed. Bear in mind that this tells us that the 75th observation in a ranked ordering of unit values, for what is supposed to be the same good, had a price that was at the very least twice that of the 25th price in the ordering.

However, resource constraints may require a staged gradualist approach using hybrid indices. The gradualist approach requires, as a first step, a rigorous evaluation for each commodity group of the relative pay-off of abandoning unit value indices. The initial aim would be to identify important commodity groups whose current series are deemed unreliable and for which there are readily available alternative sources: the "low-hanging fruit." However, changes should be limited to staged steps with good metadata and backdated series so that analysts do not confuse price changes with changes in methodology. The culmination of a program of use of hybrid indices should be an index in which unit values have little or no place.

A One-Off Switch

A preferable, though resource-intensive, approach is a one-off switch from unit value indices at the elementary level to an index based on establishment-based price surveys, possibly supplemented by global/mirror series, world commodity price indices, and domestic (export) series as applicable. That is, to stop compiling the XMPIs based on unit values and commence compiling XMPIs based on price indices, though a short period of overlap will be required. Such a switch may be prompted by a country joining a customs/monetary union. While the main problem with simply introducing a new program is the resource cost, there will be natural synergies between the administrative offices responsible for XMPIs and a producer price index (PPI) program. Given a reliable PPI program, the development of an establishment-based XMPI program would involve a stratification of existing responses by domestic/foreign markets and an extension of the sample to exporters/importers where necessary.

[19]National customs tariffs contain a list of all products that can be imported. Within the tariff, products are grouped according to the material they are made of, or according to the industrial sector to which they pertain either as input or as output materials (HS six-digit headings). Within these product groups customs tariffs contain as many tariff lines as there are different levels of customs duties. In other words, each duty rate is attached to a tariff line. Thus tariff line data are more detailed than the six-digit HS levels.

4. Where Unit Value Indices Work and Price Indices Do Not Work

4.1. *That Unit Value Indices and Price Indices Differ*

Consider the case of the same model of a television set sold at different outlets, A and B, between two periods, March and April, with broadly the same quality of service facilities. Say, in spite of the law of one price,[20] the price in March at outlet A was 425 and at outlet B was 500, and *prices do not change* between March and April. However, quantities change, say, from 100 in March to 140 in April for outlet A, with customers being attracted there by the lower price, and from 100 in March to 60 in April at outlet B. Now Laspeyres, Paasche, Fisher, and Törnqvist or unweighted price index numbers for a CPI would all show no price change in spite of the clear fact that the average cost of living (at least with regard to watching this television), of maintaining a constant level of utility, has fallen. A unit value index would take account of this change in the average *level* of prices and would register a fall of 3.2 percent. A similar argument holds for a PPI, with different producers selling the same product at different price levels. This sort of argument holds as well for an export price index, with exporters switching the balance of their exported quantities from lower to higher paying countries. And it holds for an import price index (MPI), with the importers switching from countries selling at higher prices to those selling at lower prices.

Aggregation in index number work is usually finely stratified across suppliers/countries supplying the same commodity and ignores the unit value effect. In some cases the argument is that different suppliers have different terms of sales/tax rates, though PPIs and MPIs should be valued at basic prices and not purchasers' prices, or slight quality variations. We address this issue in the next section.

There are many particular areas of application of unit values. For example, Aizcorbe and Nestoriak (2008) used unit values for the measurement of the price of health care in the U.S. for a sample of 700 million health claim records covering the period 2003–05. They defined a "good" as the bundle of completed episodes of treatments required to fix a medical condition: the (assumed) homogeneous outcome. They tracked the price of the input bundle as a whole rather than tracking prices of treatments separately. Thus shifts in the bundle to using more of cheaper (level) treatments changed the price of the bundle even if the prices of the individual treatments were unchanged. As a result, the measured price of treating diseases grew an average of 12% over this period instead of by 17% using standard index numbers.

That price indices and unit value indices give different results is well recognized. Párniczky (1974) and Balk (1998), respectively, compared unit value indices to the Paasche and Fisher price indices. These seminal decompositions, while

[20]It may be argued that the law of one price precludes such price dispersion, but there is sufficient theory and evidence of substantial price dispersion of the same item to debunk this. For example, Yoskowitz's (2002) study for raw water found one supplier discriminating against a private customer, charging $500 an acre foot (AF), while a municipality was charged $20 per AF, though there was some evidence of arbitrage and learning. Sorenson (2000) examined prices of identical prescription drugs from different pharmacies with similar sales service in upstate New York. He found, on average, the highest posted price to be over 50 percent above the lowest available price; see also Lach (2002). Many factors are cited for such differences including search cost, menu costs, signal extraction problems, price discrimination, and strategic random pricing behavior.

useful, formulated the decomposition in terms of quantity-weighted covariances. However, quantity weighting implicitly assumes homogeneity and thus cannot be used to determine unit value bias from product heterogeneity. Furthermore, the decompositions do not distinguish between levels and changes. Silver (2009b) derived a value-weighted decomposition to show how unit value indices differed from Laspeyres, Paasche, and Fisher price indices by including among other factors,[21] the effect of switches between the levels of prices.

4.2. *Unit Value Indices and Homogeneous Products*

The principle that a price index for homogeneous commodities should include changes in the levels is well enshrined in price index number theory. Index number theory conceptualizes a CPI (and MPI) in terms of a cost of living index (COLI) defined as the ratio of minimum expenditures in the periods compared required to maintain a given level of utility, assuming fixed preferences (ILO *et al*., 2004a, chapter 17). The economic theory of producer price index numbers (and XPIs) defines a (fixed input) output price index as the ratio of the two revenues in the periods compared, assuming fixed technologies and inputs (ILO *et al*., 2004b, chapter 17). A COLI will decrease if consumption is switched from higher to lower priced (identical) products, and there will be a theoretical output price index increase if output is switched from lower to higher priced (identical) products. A unit value index will take account of such switches in levels between homogeneous products, but a price index will not. The *CPI*, *PPI*, and *XMPI Manuals* (ILO *et al*., 2004a, 2004b, chapter 20; 2009, chapter 21) advocate the use of unit value indices for homogeneous goods and services, as does the *2008 SNA*:[22]

> When there is price variation for the same quality of good or service, the price relatives used for index number calculation should be defined as the ratio of the weighted average price of that good or service in the two periods, the weights being the relative quantities sold at each price. Suppose, for example, that a certain quantity of a particular good or service is sold at a lower price to a particular category of purchaser without any difference whatsoever in the nature of the good or service offered, location, timing or conditions of sale, or other factors. A subsequent decrease in the proportion sold at the lower price raises the average price paid by purchasers for quantities of a good or service whose quality is the same and remains unchanged, by assumption. It also raises the average price received by the seller without any change in quality. This must be recorded as a price and not a volume increase. (Commission of the European Communities *et al*., 2008, paragraph 15.69)

The *SNA 2008* holds that if the price dispersion in a period is not due to quality differences—the homogeneous case—a unit value index should be used. Yet it notes an important exception regarding the case of institutionalized price discrimination. If different purchasers of the same good or service, say water or electricity, face different prices and the individual purchasers, say commercial

[21]These other factors were different powers of the substitution effect. Note that the substitution effect is concerned with the response of quantity *changes* to price *changes*; not *levels*.

[22]Similar guidance can be found in the *1993 SNA* (Commission of the European Communities *et al*., 1993).

customers and private households, are unable to change from one price to another, then price indices should be used. The constraint on the availability to the purchaser of different prices must be institutional and not simply an income constraint. Yet even this stance is problematic. As noted above, the economic theory of producer price index defines a (fixed input) output price index as the ratio of the two revenues in the periods compared, assuming fixed technologies and inputs. A theoretical cost of living index (COLI) is defined as the ratio of minimum expenditures in the periods compared that are required to maintain a given level of utility, assuming fixed preferences (ILO *et al.*, 2004a, chapter 17). From the producer's perspective, a shift in the quantities of identical items sold at differentiated prices results in a change in revenue from fixed inputs[23]—the institutional arrangements matter and indeed were likely devised to enable revenue to be maximized. Unit values should be used. From the purchaser's perspective, it make no difference to the ratio of expenditures for a, say, commercial customer if the producer shifts some of its quantities to private households. In this case, the institutional arrangements do not matter and unit values should not be used. In other words, from the perspective of the purchasers of a homogeneous product, what counts is his or her (separate from other purchasers) unit value price; not the overall unit value price across all purchasers, which would be the relevant price for the seller.

Thus the principles are quite clear. Unit value indices should be used for homogeneous products for which price indices would be misleading and, as outlined in Section 3 in the context of XMPIs, price indices should be used for heterogeneous products for which unit value indices are misleading. But then there is the question of "broadly comparable" goods and services.

5. Broadly Comparable Goods and Services

Groups of goods and services cannot be simply and meaningfully divided into those that are homogeneous and those that are heterogeneous. In differentiated markets, many products are very similar, but slightly differentiated to gain market advantage and/or to appeal to a particular niche. This differentiation involves not only product characteristics but also the terms of the sale including warranty, service level, return policy, geographical convenience, parking facilities at the outlet, and so forth. There is a continuum between strictly homogeneous products sold at the same terms and conditions and heterogeneous products which have no substitutability or characteristics in common.

Consider again the case outlined in Section 4.1 of the price in March of identical goods sold at outlet A for a price of 425 and at outlet B for 500. Suppose that *prices do not change at either outlet* between March and April. Again, let the quantities change from 100 in March to 140 in April for outlet A, with consumers being attracted by the lower price at outlet A, whereas sales fall at outlet B from 100 in March to 60 in April. We established above that a unit value index would correctly show a fall of 3.2 percent while a price index would be unchanged. Next, consider the case in which the product items sold at outlet B in both March and

[23]We assume the costs of serving the different purchasers are significantly different.

April are differentiated from the product items sold at outlet A. In particular, suppose that the items sold at outlet B have an additional feature that similar such goods in the market show is worth a 5 percent premium. We could quality adjust the items sold in B. Quality adjustment factors can thus be applied to prices to render a comparison of prices of differentiated items akin to a comparison of homogeneous items.

The requirement here, for an operational adjustment for quality differences to be made to a price, is that the extent of product differentiation must be measurable. A useful concept of "broadly similar items" here is one in which the items share similar characteristics, but differ in the amount to which the characteristics are incorporated. This concept would include the absence or presence of a quality feature, say a built in camera for a personal computer (PC). This consideration of price-determining quality characteristics sets the stage for the use of hedonic regressions as a means by which quality adjustments can be undertaken and the use of the explanatory power, $\bar{R}^2$, of the hedonic regression as a measure of the extent of the product differentiation, as explained below.

It can be argued that substitutability should be the main concept behind defining a continuum between homogenous and heterogeneous items. However, operationalizing substitutability in this context is problematic. Substitutability exits for goods and services for which quality adjustments for unit values are not feasible and the concept of an average (unit value) price not meaningful, for example, beef or chicken, or going to a restaurant or movies. However, the concept of substitutability has some application. Consider, for example, the choice between television sets of different makes and characteristics. A hedonic regression will identify the marginal value of different brands against each other, screen sizes, and so forth. Yet markets are often segmented, a practice encouraged by marketing professionals and marketing texts such as Kotler and Keller (2009). Consumers may, for example, only consider certain brands as belonging to the segment they wish to purchase from and, indeed, the marketing and pricing of the brands may well have been with this intention in mind. Thus the continuum of homogeneous to heterogeneous television sets, as being those sharing a common characteristics will be conditioned on their belonging to the same market segment.[24]

We can make use of (hedonic) quality-adjusted unit value indices that remove the effects on prices of product heterogeneity; a proposal that goes back to Dalén (2001) and is formalized and empirically examined in de Haan (2004) and reiterated in de Haan (2007).[25] De Haan's (2004) quality-adjusted unit value index solution to the problem of price measurement of broadly comparable items is most useful and instructive. Since a unit value index is appropriate for homogeneous items, a quality-adjusted unit value index might be appropriate for broadly comparable items. We consider such a measure.

[24]A yet further criterion for homogeneity, advocated by Balk (1998), is whether *quantities* of homogeneous items can be meaningfully added. Thus a PC with a built-in camera cannot be meaningfully added to one without to make two PCs. It should be apparent from equation (4) that if a quality adjustment was made to the price of one PC for the value of the camera, then the unit values for the two PCs are well defined, and thus, the summation of the quantities meaningful.

[25]Silver and Heravi (2002) used hedonic regressions to control for heterogeneity in a Dutot index; see also Silver and Heravi (2007).

A hedonic regression (see Triplett, 2004) using data on $m = 1, \ldots, M$ *matched* models for periods $\tau = 0,t$, of the price, p_m^τ, on $k = 1, \ldots, K$ quality characteristics, z_{km}^τ is calculated as follows:

$$p_m^\tau = \beta_0^\tau + \sum_{k=1}^{K} \beta_k^\tau z_{km}^\tau + u_m^\tau \tag{5}$$

where u_m^τ are assumed to be normally distributed with mean δ^τ and variance ξ_τ^2. The heterogeneity-adjusted prices in each period relative to a reference numeraire item with mean characteristics $\bar{z}_{km}^\tau$ in each period are given by:

$$\hat{p}_m^\tau = p_m^\tau - \sum_{k=1}^{K} \beta_k^\tau \left(z_{km}^\tau - \bar{z}_{km}^\tau \right). \tag{6}$$

Bear in mind the models in each period are matched so that $z_{km}^\tau = z_{km}^0 = z_{km}^t$. Note also that β_k^0 may or may not equal β_k^t and (5) can be estimated on pooled data with a dummy variable for time and with the constraint that $\beta_k^t = \beta_k^0 = \beta_k^\tau$, though it is preferable to estimate (5) separately for each time period without the constraint. The heterogeneity adjusted unit value index is:

$$P_U^* = \left(\frac{\sum_{m=1}^{M} \hat{p}_m^t q_m^t}{\sum_{m=1}^{M} q_m^t} \right) \Bigg/ \left(\frac{\sum_{m=1}^{M} \hat{p}_n^0 q_n^0}{\sum_{m=1}^{M} q_n^0} \right). \tag{7}$$

For commodities with slight product differentiation we would advise a measure of (7) based on (6). For commodities that are more broadly differentiated, Silver (2009b) proposed the use of a weighted average of (7) and a Fisher price index (1) with the weights based on the explanatory power, $\bar{R}^2$, of the hedonic regression (5). As the differentiation decreases, in terms of the explanatory power of the regression, more weight would be applied to (7) than (1). Of course the quality adjustments need not use hedonic regressions. For instance, a simpler method of quality adjustment might be used if the heterogeneity is due to the addition of a single feature or option for which cost or market estimates of value are available.

6. Conclusions

Price index number theory tells us that superlative index numbers such as Fisher and Törnqvist are superior for heterogeneous products. Unit value indices were evaluated against the axiomatic and economic theoretic approaches to index numbers and found to have major shortcomings when applied to heterogeneous products. Yet the United Nations (1981) strategy advised that unit value indices based on data from customs administrative documents be used by countries with

limited resources as surrogates for narrow specification price indices, mainly in view of the low cost of the data. The bias in unit value indices was judged by the United Nations (1981) to be tolerable.

We argue in this paper that, by now, unit values based on customs data are a poor choice because of: (i) the increased product differentiation and turnover of differentiated products that will aggravate bias due to compositional quality mix changes; (ii) the lack of ways for dealing with quality change with unit value indices; (iii) the increase in trade in services coupled with the lack of customs data for many sorts of service products; (iv) the impossibility of dealing appropriately with "unique" goods such as ships with customs data and unit value indices; and (v) the fact that unit values cannot be tailored to meet valuation needs.

Evidence on the extent of the resulting bias from the use of unit value indices was outlined in Section 3.2 where it was found that export and import unit value indices are inadequate surrogates for their price index counterparts. The evidence is that unit value indices can be seriously misleading. A one-off change to establishment based price indices is advised for countries still using customs unit value indices. Many countries still rely on unit value indices, and many are likely to continue to do so. Thus strategies of improving unit value indices are also considered in this paper, including increased stratification, by company/country of origin/destination, size of shipment, and type of commodity, and improved detection and deletion (after follow-up) of outliers. The nature of customs documentation and the natural volatility of sticky price changes and pass through rates as well as the expansion of customs unions seriously constrain the possibilities for improving unit value indices.

In addition, therefore, the benefits of hybrid indices are discussed. A move to hybrid indices constitutes a staged, gradualist program to price index program improvement. The gradualist approach requires as a first step, a rigorous evaluation of each commodity group of the relative pay-off of abandoning unit value indices. The initial aim is to identify important commodity groups whose current series are deemed unreliable and for which there are readily available alternative sources: the "low-hanging fruit."

Given the problem with unit value indices based on customs data, the paper considers the usefulness of unit value indices for homogenous products and the potential bias in the use of superlative and other price indices for such products. Concern arises from the fact that price indices do not pick up compositional (quantity) switches when different batches of truly homogeneous products are sold, contemporaneously at *different prices*, while unit value indices can deal with this situation. A switch to buying at outlets with lower prices by utility-maximizing consumers should lead to a decrease in the average cost of living while a switch to outlets that can sell their products for relatively higher prices can lead to higher revenue for revenue maximizing producers. In situations like these, the CPI should decrease and the PPI should fall, respectively. A price index need not behave in this manner while a unit value index would.

We then turned in Section 5 to the case of broadly-comparable products. There are many products that are only slightly differentiated. Given that a price index may yield quite different results than a unit value index, there is the obvious question as to whether small differences in quality can justify the loss of the

benefits of a unit value index. Quality adjusted prices are an accepted part of index number methodology[26] and such adjustments enable unit value indices to be compiled for items that are not quite homogeneous. This significantly extends the potential use of unit value indices and raises concerns about the use of price index number formulas for aggregating price changes over such products.

References

Aizcorbe, Ana and Nicole Nestoriak, "The Importance of Pricing the Bundle of Treatments, U.S. Bureau of Economic Analysis," Paper presented at the 2008 World Congress on National Accounts and Economic Measurement, Washington DC, May 2008.

Alterman, William, "Price Trends in U.S. Trade: New Data, New Insights," in Peter Hooper and J. David Richardson (eds), *International Economic Transactions*, University of Chicago Press, Chicago, 109–43, 1991.

Angermann, Oswald, "External Terms of Trade of the Federal Republic of Germany Using Different Methods of Deflation," *Review of Income and Wealth*, 26, 367–85, 1980.

Balk, B. M., "On the Use of Unit Value Indices as Consumer Price Subindices", in *Proceedings of the Fourth Meeting of the International Working Group on Price Indices*, U.S. Bureau of Labor Statistics, Washington DC, 112–20, 1998 (http://www.ottawagroup.org).

———, "Price Indices for Elementary Aggregates: The Sampling Approach," *Journal of Official Statistics*, 21, 675–99, 2005

Bradley, Ralph, "Pitfalls of Using Unit Values as a Price Measure or Price Index," *Journal of Economic and Social Measurement*, 30, 39–61, 2005.

Commission of the European Communities, International Monetary Fund, Organisation for Economic Co-operation and Development, United Nations, World Bank, *System of National Accounts 1993*, EC, IMF, OECD, UN, World Bank, Brussels/Luxembourg, New York, Paris, Washington, 1993.

———, *System of National Accounts 2008*, pre-edit version of Volume 1, IMF and United Nations, Washington DC and New York, 2008 (http://unstats.un.org/unsd/sna1993/snarev1.asp).

Dalén, J., "Statistical Targets for Price Indices in Dynamic Universes," Paper presented at the Sixth Meeting of the International Working Group on Price Indices, Canberra, April 2–6, 2001.

Decoster, Renaud, *Import and Export Price Indices: A Comparison of Import Unit Value Indices and Import Specific Price Indices for Four EU Member States, Second Report*, Study Performed for Eurostat Unit C-4: Methodology, Nomenclature and Statistics of External and Intra-Community Trade, PLANISTAT Europe, Paris, 2003a.

———, *Import and Export Price Indices: Final Report*; Study Performed for Eurostat Unit 0074 C-4: Methodology, Nomenclature and Statistics of External and Intra-Community Trade, PLANISTAT Europe, Paris, 2003b.

de Haan, J., "Estimating Quality-Adjusted Unit Value Indices: Evidence from Scanner Data," Paper presented at the Seventh EMG Workshop, Sydney, Australia, December 12–14, SSHRC International Conference on Index Number Theory and the Measurement of Prices and Productivity, Vancouver, June 30–July 3, 2004.

———, "Hedonic Price Indices: A Comparison of Imputation, Time Dummy and Other Approaches," Paper presented at the Seventh EMG Workshop, Sydney, Australia, December 12–14, 2007.

Diewert, W. E., "Exact and Superlative Index Numbers," *Journal of Econometrics*, 4, 114–45, 1976.

———, "Superlative Index Numbers and Consistency in Aggregation," *Econometrica*, 46, 883–900, 1978.

ILO (International Labour Office), IMF, OECD, Eurostat, United Nations, World Bank, *Consumer Price Index Manual: Theory and Practice*, ILO, Geneva, 2004a (http://www.ilo.org/public/english/bureau/stat/guides/cpi/index.htm).

———, *Producer Price Index Manual: Theory and Practice*, International Monetary Fund, Washington DC, 2004b (http://www.imf.org/external/data.htm#guide).

———, *Export and Import Price Index Manual: Theory and Practice*, International Monetary Fund, Washington DC, 2009 (http://www.imf.org/external/data.htm#guide).

Kotler, Philip and Kelvin Keller, *A Framework for Marketing Management*, Prentice-Hall, Upper Saddle River, NJ, 2009.

Kravis, Irving B. and Robert E. Lipsey, *Price Competitiveness in World Trade*, Studies in International Economic Relations, Vol. 6, National Bureau of Economic Research, New York, 1971.

[26]See chapter 7 of ILO *et al*. (2004a, 2004b).

———, "Prices and Terms of Trade for Developed-Country Exports of Manufactured Goods," National Bureau of Economic Research Working Paper No. 0774, 1985.
Lach, S., "Existence and Persistence of Price Dispersion," *Review of Economics and Statistics*, 84, 433–44, 2002.
Párniczky, G., "Some Problems of Price Measurement in External Trade Statistics," *Acta Oeconomica*, 12, 229–40, 1974.
Ruffles, David and Kevin Williamson, "Deflation of Trade in Goods Statistics: Derivation of Price and Volume Measures from Current Price Values," *Economic Trends*, 521, April 1997.
Silver, Mick, "Do Unit Value Export, Import, and Terms of Trade Indices Represent or Misrepresent Price Indices?" *IMF Staff Papers*, 56, 297–322, 2009a.
———, "An Index Number Formula Problem: The Aggregation of Broadly Comparable Items," IMF Working Paper Series, WP/09/19, Washington DC, January 2009b.
Silver, Mick and Saeed Heravi, "Why the CPI Matched Models Method May Fail Us: Results from an Hedonic and Matched Experiment Using Scanner Data," European Central Bank Working Paper Series, 144, 2002.
———, "Why Elementary Price Index Number Formulas Differ: Price Dispersion and Product Heterogeneity," *Journal of Econometrics*, 140, 874–83, 2007.
Sorenson, A. T., "Equilibrium Price Dispersion in Retail Markets for Prescription Drugs," *Journal of Political Economy*, 108, 833–50, 2000.
Triplett, J. E., *Handbook on Hedonic Indices and Quality Adjustments in Price Indices*, Directorate for Science, Technology and Industry, OECD, Paris, 2004.
United Nations, Department of International Economic and Social Affairs, *Strategies for Price and Quantity Measurement in External Trad*e, Statistical Papers, Series M, No. 69, United Nations, New York, 1981.
———, Department of International Economic and Social Affairs, *Price and Quantity Measurement in External Trade: Two Studies of National Practice*, Statistical Papers, Series M, No. 76, United Nations, New York, 1983.
———, Department of Economic and Social Affairs, Statistics Division, *International Merchandise Trade Statistics: Compilers Manual*, Series F, No. 87, United Nations, New York, 2004.
———, Department of Economic and Social Affairs, Statistics Division, *National Practices in Compilation and Dissemination of External Trade Index Numbers*, Statistical Papers, Series M, No. 76, United Nations, New York, 2005.
Von der Lippe, Peter, "Price Indices and Unit Value Indices in German Foreign Trade Statistics," Unpublished Paper, University of Duisburg-Essen, 2007.
WTO (World Trade Organization), United Nations Conference on Trade and Development (UNCTAD), and International Trade Centre (ITC), *World Tariff Profiles 2006*, World Trade Organization, Geneva, 2007.
Yoskowitz, D. W., "Price Dispersion and Price Discrimination: Empirical Evidence from a Spot Market for Water," *Review of Industrial Organization*, 20, 283–9, 2002.