Journal of Official Statistics, Vol. 27, No. 4, 2011, pp. 553–567

An Index Number Formula Problem: The Aggregation of Broadly Comparable Items

Mick Silver[1]

Index number theory informs us that if data on matched prices and quantities are available, a superlative index number formula is best to aggregate *heterogeneous* items, and a unit value index is best to aggregate *homogeneous* ones. The formulas can give very different results. Neglected is the practical case of broadly comparable items. This article provides a formal analysis as to why such formulas differ and proposes a solution to this index number problem.

Key words: Unit value index; superlative index; Consumer Price Index (CPI); Producer Price Index (PPI); hedonic regression; index numbers; price dispersion.

JEL Classification Numbers: C43; C81; E31; L15.

1. Introduction

Price index numbers are used to measure aggregate changes in prices, usually over many goods or services. There is a consensus as to which price index number formulas are best when price and quantity/value (weight) information are available, as given by the internationally-accepted manuals on consumer, producer, and trade price indexes (ILO et al. 2004a; 2004b; 2009). In Section 2, it is argued that superlative price index formulas are best for heterogeneous items, and unit value indexes are best for homogeneous ones. Given that the results from such formulas can be very different we consider the formal decomposition, and thus understanding, of the factors underlying their difference. A natural question that then arises is "What if the goods are neither homogenous nor heterogeneous, what we term *broadly comparable*?" This article outlines the limited work in this area, primarily by de Haan (2004; 2007), and makes further proposals.

The potential error in superlative index numbers for homogeneous goods or services is a neglected and important index number issue. If, for example, the prices of goods A and B were 10 and 12, respectively, in both the reference and current periods, but there was a shift in quantities from 6 for both A and B in the reference period to 8 for A and 4 for B in the current period, a superlative or any other price index number formula for heterogeneous goods would give an answer of unity, that is no overall price change. However, the correct answer for homogeneous goods would be a unit value *fall* of 3 per

[1] International Monetary Fund, Statistics Department, 1900 Pennsylvania Avenue, Washington, DC 20431, U.S.A. Email: msilver@imf.org
Disclaimer: The views expressed herein are those of the author and should not be attributed to the IMF, its Executive Board, or its management.

cent, appropriately reflecting the shift in the quantity basket in the current period from the higher price *level* of 12 for B to the lower price *level* of 10 for A. The good, and cost of living with regard to this good, is, on average, now cheaper.

The *CPI Manual* (ILO et al. 2004a, Chapter 20) and the *2008 SNA* advocate the use of unit value indexes for homogeneous goods and services:

> 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. 2009, Paragraph 15.68).

Index number theory recognizes that the appropriateness of each formula depends on whether the items aggregated are homogeneous or otherwise (Diewert 1995; Balk 1998; 2005). As matters stand, the advice is to simply determine whether or not items are homogeneous and apply the appropriate formula. But what if the items are broadly comparable, that is they are of different qualities such that some of the price dispersion is due to product differentiation, but some is due to search costs or price discrimination? A superlative index would wrongly ignore the effect on average prices of any shift in quantities of the same (or quality-adjusted) item to higher or lower average price levels, but a unit value index would wrongly treat changes in compositional mix of items of different quality as price changes, the familiar unit value bias. Given that these formulas will generally give quite different answers, it is important to determine why they differ, the conditions under which each is suitable, and what to do when – as is likely to be the case – neither is suitable.

The rationale for unit value and superlative indexes is outlined in Section 2, along with some discussion of the circumstances under which each of them is appropriate. Section 3 provides a formal analysis of how the unit value and Fisher price index differ. In Section 4 a solution is proposed: a formula based on an average of (quality-adjusted) unit value and Fisher indexes whose respective weights are derived from the explanatory power of a hedonic regression. An application using scanner data is provided in Section 5, with conclusions in Section 6.

The use of the results of this article is in the determination of price and volume measures at the national and micro level for economic aggregates. It applies to consumer, commodity, producer (input and output), import, and export price indexes, as well as price indexes of capital goods, such as house price indexes. Since price indexes are used as deflators, there is a concomitant application to volume indexes. The concern is with aggregation where price and quantity/value information is available for broadly comparable items. A good example would be measuring the aggregate price and volume change of matched (over time) models of different qualities of automobiles, but not aggregating price or volume changes over automobiles and beef.

2. Superlative and Unit Value Indexes

2.1. Superlative Index Numbers

The Fisher, P_F, and the Törnqvist, P_T, index number formulas are both commonly-used superlative indexes. 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). The Fisher price index is a geometric mean of Laspeyres, P_L, and Paasche, P_P, price indexes 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^t_m and q^t_m for period t, and p^0_m and q^0_m for period 0, by:

$$P_F \equiv \sqrt{\frac{\sum_{m=1}^{M} p^t_m q^0_m}{\sum_{m=1}^{M} p^0_m q^0_m} \times \frac{\sum_{m=1}^{M} p^t_m q^t_m}{\sum_{m=1}^{M} p^0_m q^t_m}} = \sqrt{P_L \times P_P} \tag{1}$$

The Törnqvist price index is defined as

$$P_T = \prod_{m=1}^{M} \left(\frac{p^t_m}{p^0_m}\right)^{(s^0_m + s^t_m)/2} \text{ where } s^t_m = p^t_m q^t_m / \sum_m p^t_m q^t_m \text{ and } s^0_m = p^0_m q^0_m / \sum_m p^0_m q^0_m \tag{2}$$

Both P_F and P_T make symmetric use of each period's price and quantity information. Diewert (1976; 1978), from an approach based on economic theory, demonstrated that both Fisher and Törnqvist indexes belong to a class of *superlative indexes* that have the desirable property of incorporating substitution effects, that is, the effect of consumers substituting from their basket of goods towards those with relatively low price increases, thus lowering their cost of living. Laspeyres and Paasche indexes are fixed (quantity) basket price indexes and allow no such substitution. His more formal analysis is based on the properties of the underlying aggregator functions. Aggregator functions underlie the definition of indexes in economic theory; an example is a utility function 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. Laspeyres, for example, corresponds to a highly restrictive Leontief form. The underlying functional forms for superlative indexes, including Fisher and Törnqvist, are flexible: they are second-order approximations to other (twice-differentiable) homothetic forms around the same point. It is the generality of functional forms that superlative indexes represent that allows them to accommodate substitution behavior and be desirable indexes.

In the test or axiomatic approach, desirable properties for an index number are chosen, and different formulas are evaluated against them. Fisher described his index as "ideal" because it satisfied the tests proposed, including the "time reversal" and "factor reversal" tests. 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 proportionate change in the current to reference period values, and that the two indexes are symmetric in prices and quantities (see Balk 2008, 84–85). In practice, Laspeyres-type indexes, such as the Lowe index (see ILO et al. 2004a, Chapter 15), are often calculated, because data on current period information are unavailable in real time. The arguments presented in this article apply as much to the use of unit value indexes against Laspeyres-type price index formulas as they do against superlative index number formulas.

2.2. *Unit Value Indexes*

A unit value index, P_{UV}, is given by

$$P_{UV}(t,0) \equiv \left(\frac{\sum_{m=1}^{M} p_m^t q_m^t}{\sum_{m=1}^{M} q_m^t}\right) \Big/ \left(\frac{\sum_{m=1}^{M} p_m^0 q_m^0}{\sum_{m=1}^{M} q_m^0}\right) \tag{3}$$

If the items whose prices are being aggregated are identical, that is perfectly homogeneous, then a unit value index has desirable properties. ILO et al. (2004a; 2004b; 2009) and Balk (2005) identify it as the target index for homogeneous goods. For example, peak-period electricity for which the service flow is constant, is a homogeneous good. Aizcorbe and Nestoriak (2008) used unit values for the measurement of the price of health care of well-defined outcomes as the result of bundles of treatment in the U.S., using a sample of 700 million health claim records.

Consider the case where the exact same item is sold at different prices during the same period, say lower sales and higher prices in the first week of the month and higher sales and lower prices in the last week of the month. The unit value for the monthly index solves the time aggregation problem and appropriately gives more weight to the lower prices than the higher ones in the aggregate. Furthermore, 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 that is intuitively appropriate, that is

$$\frac{\sum_{m=1}^{M} p_m^1 q_m^1}{\sum_{m=1}^{M} p_m^0 q_m^0} \Big/ \left[\left(\frac{\sum_{m=1}^{M} p_m^1 q_m^1}{\sum_{m=1}^{M} q_m^1}\right) \Big/ \left(\frac{\sum_{m=1}^{M} p_m^0 q_m^0}{\sum_{m=1}^{M} q_m^0}\right)\right] = \frac{\sum_{m=1}^{M} q_m^1}{\sum_{m=1}^{M} q_m^0} \tag{4}$$

Note that the summation of quantities in the top and bottom of the right-hand side of Equation (2) must be of the exact same type of item for the expression to make sense.

Balk (1998; 2008, page 73) showed that the unit value index satisfies the conventional index number tests, with the important 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*: $P(p,p,q^0,q^t) = 1$ is a special case of the proportionality test–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 relative quantities, that is, the composition of the products compared, do not change;

and (iii) *Invariance to Changes in the Units of Measurement (commensurability) Test*:

$$P\left(\alpha_1 p_1^0, \ldots, \alpha_m p_m^0; \alpha_1 p_1^t, \ldots, \alpha_m p_m^t; \alpha_1^{-1} q_1^0, \ldots, \alpha_m^{-1} q_m^0; \alpha_1^{-1} q_1^t, \ldots, \alpha_m^{-1} q_m^t\right)$$
$$= P\left(p_1^0, \ldots, p_m^0; p_1^t, \ldots, p_m^t; q_1^0, \ldots, q_m^0; q_1^t, \ldots, q_m^t\right) \text{ for all } \alpha_1 > 0, \ldots, \alpha_m > 0;$$

that is, the price index does not change if the units of measurement for each product are changed. Changes in units of measurement *de facto* arise when the quality of items changes.

However, these tests were devised for the aggregation of heterogeneous items and are not always meaningful for homogeneous items. For example, in the introduction we outlined the case where prices do not change, but a shift in quantities to lower level prices leads to a fall in the overall price level, which is 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, the homogeneous case, then the unit value (plug-in) index will not be biased against the theoretical index.

There is a literature on bias in import and export price indexes that use, as proxies for price changes, changes in unit value indexes for goods in a commodity group used for customs documents. Such groups can be too widely defined to ensure homogeneity, and the findings in this literature are that such unit value indexes substantially misrepresent price changes due to compositional changes in quantities and quality mix of what is exported and imported in the category concerned (Angermann 1980; Alterman 1991; Ruffles and Williamson 1997; Silver 2009a). Having outlined the nature and differences between Fisher price and unit value indexes, it is necessary to consider the factors determining such differences.

3. The Difference Between a Unit Value and a Fisher Price Index

Párniczky (1974) and von der Lippe (2007) compare unit value indexes to the Paasche and Laspeyres price indexes, and Balk (1998) compares unit value indexes to Fisher indexes. More recently, Diewert and von der Lippe (2010) identify a "bias" formula between unit value indexes and Paasche, Laspeyres, and Fisher price indexes, though in terms of single covariance terms that do not distinguish between levels and substitution effects. Further, the analysis is not directed to the issue of homogeneity, with differences between unit value and price index formula described as "bias," when for aggregation over homogeneous items the bias is in the price index formulas. These seminal decompositions, while useful, suffer from either (i) undertaking the decomposition in terms of quantity-weighted covariances of changes, quantity weighting implicitly assumes homogeneity though, or (ii) not distinguishing switches in quantities to lower or higher price levels. Both are issues of concern here. We provide a new decomposition.

We define a Laspeyres price index by

$$P_L(t,0) \equiv \frac{\sum_{m=1}^{M} p_m^t q_m^0}{\sum_{m=1}^{M} p_m^0 q_m^0} = \sum_{m=1}^{M} s_m^0 \frac{p_m^t}{p_m^0} \tag{5}$$

and the Laspeyres quantity index by

$$Q_L(t,0) \equiv \frac{\sum_{m=1}^{M} p_m^0 q_m^t}{\sum_{m=1}^{M} p_m^0 q_m^0} = \sum_{m=1}^{M} s_m^0 \frac{q_m^t}{q_m^0} \tag{6}$$

where s_m^0 was defined in Equation (2) above as period 0 value shares. The Paasche price index is defined by

$$P_P(t,0) \equiv \frac{\sum_{m=1}^{M} p_m^t q_m^t}{\sum_{m=1}^{M} p_m^0 q_m^t} \tag{7}$$

The Fisher price index is defined as Equation (1) above, the unit value index as Equation (3) above, and the Dutot quantity index based on Equation (4) above as

$$Q_D(t,0) \equiv \frac{\sum_{m=1}^{M} q_m^1}{\sum_{m=1}^{M} q_m^0} \tag{8}$$

From Equation (4) it can be seen that the unit value index equals the value ratio divided by the Dutot quantity index, and the value ratio in turn can be written as the Paasche price index times the Laspeyres quantity index, that is

$$P_{UV}(t,0) \equiv \frac{\sum_{m=1}^{M} p_m^t q_m^t / \sum_{m=1}^{M} p_m^0 q_m^0}{Q_D(t,0)} = \frac{P_P(t,0) \times Q_L(t,0)}{Q_D(t,0)} \tag{9}$$

Now the value ratio can be decomposed as

$$\begin{aligned} \frac{\sum_{m=1}^{M} p_m^t q_m^t}{\sum_{m=1}^{M} p_m^0 q_m^0} &= \frac{\sum_{m=1}^{M} s_m^0 \left(p_m^t/p_m^0\right)\left(q_m^t/q_m^0\right)}{\sum_{m=1}^{M} s_m^0 \left(p_m^t/p_m^0\right)} \sum_{m=1}^{M} s_m^0 \left(p_m^t/p_m^0\right) \\ &= \operatorname{cov}^{s_0}\left(p_m^t/p_m^0, q_m^t/q_m^0\right) + \sum_{m=1}^{M} s_m^0 \left(p_m^t/p_m^0\right) \sum_{m=1}^{M} s_m^0 \left(q_m^t/q_m^0\right) \\ &= \operatorname{cov}^{s_0}\left(p_m^t/p_m^0, q_m^t/q_m^0\right) + P_L(t,0) Q_L(t,0) \end{aligned} \tag{10}$$

where $\operatorname{cov}^{s_0}\left(p_m^t/p_m^0, q_m^t/q_m^0\right)$ is the s_0-weighted covariance between p_m^t/p_m^0 and q_m^t/q_m^0. In the literature, Equation (10) is referred to as a Bortkiewicz decomposition, see Bortkiewicz (1923, 374–375).

Using Equations (1), (4), (9), and (10) we derive the ratio of a unit value index and Fisher price index as

$$\frac{P_{UV}(t,0)}{P_F(t,0)} = \left[\frac{P_{UV}(t,0)}{P_L(t,0)} \times \frac{P_{UV}(t,0)}{P_P(t,0)}\right]^{\frac{1}{2}}$$

$$= \left[\frac{\text{cov}^{s_0}\left(p_m^t/p_m^0, q_m^t/q_m^0\right) + P_L(t,0)Q_L(t,0)}{Q_D(t,0)P_L(t,0)} \times \frac{P_P(t,0)Q_L(t,0)}{Q_D(t,0)P_P(t,0)}\right]^{\frac{1}{2}} \tag{11}$$

$$= \left[\frac{\text{cov}^{s_0}\left(p_m^t/p_m^0, q_m^t/q_m^0\right)}{P_L(t,0)Q_L(t,0)} + 1\right]^{\frac{1}{2}} \frac{Q_L(t,0)}{Q_D(t,0)}$$

Thus the ratio of a unit value index to a Fisher price index can be split into two parts, the first measuring the *substitution effect* and the second the *levels effect*. The substitution effect depends on the correlation between price and quantity changes. The levels effect measures the extent to which quantities shift towards higher or lower prices *even if there are no price changes*. The levels effect can be decomposed further by noticing that

$$\frac{Q_L(t,0)}{Q_D(t,0)} = \frac{\sum_{m=1}^{M} p_m^0 q_m^t / \sum_{m=1}^{M} q_m^t}{\sum_{m=1}^{M} p_m^0 q_m^0 / \sum_{m=1}^{M} q_m^0} \tag{12}$$

Equation (12) shows the levels effect as the ratio of two unit values, both of which are valued at period 0 prices, but the numerator contains the quantities in period t while the denominator has period 0 quantities. The levels effect is the change in unit values that arises solely as the result of changes in the quantities sold/purchased of a homogeneous item at different price levels. For example, assume peak-period electricity is sold in period 0 to M $= 1, \ldots, 9$ different regions at the same price and to Region 10 at a higher price due to price discrimination, as opposed to quality consideration. If in peak-period t some of the quantity sold to the 9 regions is switched to Region 10, then the revenue received by the electricity-supplying establishment from its fixed inputs increases. The producer gains from an overall price increase even though no price change may have taken place between periods 0 and t. If all quantities changed by the same proportion, substituting $\lambda q_m^0 = q_m^t$ in Equation (12), the levels effect is rightly zero.

The purpose of this section was to provide a formal analysis as to why unit value and Fisher price indexes might differ. Equation (11) successfully decomposed the difference between the formulas into two effects, a substitution and a levels effect, as we will illustrate in Section 6.

It is also apparent from (11) that the unit value index will be equal to the Fisher price index if the covariance between price and quantity changes is zero AND quantity changes for each of the M items are the same, or all weights are equal. These are extreme conditions. Having no dispersion in prices, quantities or their changes is a negation of the index number problem, and while we do not expect the laws of economics to work perfectly, there is expected to be some relationship between price and quantity changes.

4. What to Do for Broadly Comparable Items

For homogeneous items there is no problem: the answer is a unit value index. For heterogeneous items there is no problem: the answer is a superlative index such as a Fisher price index. However, many goods or services are broadly comparable: consumers in most product markets have a selection of differentiated items available to them, even if the differentiation is only due to the services provided by different outlets providing the same item. There is much concern in price index literature about including "outlet effects," such as the fall in the cost of living from a shift to lower-priced stores such as Wal-Mart (see Hausman and Leibtag 2008).

Our concern is with items that are comparable such as models of television sets, washing machines, laptop computers, automobiles, whose price dispersion due to product differentiation is significant, as is the price dispersion due to factors that cannot be accounted for by the characteristics of the item. Markets are often segmented, a practice encouraged by marketing professionals and marketing texts such as Kotler and Keller (2009). Consumers may consider only 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 common characteristics, will be conditioned on their belonging to the same market segment. We noted above that there is much empirical evidence that in many markets the law of one price does not hold; reasons for this include price discrimination, menu costs, search costs, signal extraction, inventory holdings, and strategic pricing. There is an element in the price comparison between differentiated varieties that is due to measurable quality differences for which the quality-adjusted unit value reduces the problem to one of homogenous items, but there is also an element for which a Fisher price index given by (1) is appropriate.

Quality-adjustment factors can be applied to prices to render the comparison of prices of differentiated items akin to one of homogeneous items. We make use of (hedonic) quality-adjusted unit value indexes 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). Silver and Heravi (2002) used hedonic regressions to control for heterogeneity in a Dutot index, see also Silver and Heravi (2007). Since a unit value index is appropriate for homogeneous items, a quality-adjusted unit value index must be appropriate for broadly comparable items. We consider such a measure.

A hedonic regression (see Triplett 2006) 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

$$p_m^\tau = \beta_0^\tau + \sum_{k=1}^{K} \beta_k^\tau z_{km}^\tau + u_m^\tau \tag{13}$$

where u_m^τ are assumed to be normally distributed with mean and variance δ^τ and ξ_τ^2, respectively. 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

$$\hat{p}_m^\tau = p_m^\tau - \sum_{k=1}^{K} \beta_k^\tau \left(z_{km}^\tau - \bar{z}_k^\tau \right) \tag{14}$$

Bear in mind that the models in each period are matched, so that $z^{\tau}_{km} = z^{0}_{km} = z^{t}_{km}$. Note also that β^{0}_{k} may or may not equal β^{t}_{k}, and (13) can be estimated on pooled data with a dummy variable for time and with the constraint that $\beta^{t}_{k} = \beta^{0}_{k} = \beta^{\tau}_{k}$, though it is preferable to estimate (13) separately for each time period without the constraint. The heterogeneity-adjusted unit value index is

$$P^{*}_{UV} = \left(\frac{\sum_{m=1}^{M} \hat{p}^{t}_{m} q^{t}_{m}}{\sum_{m=1}^{M} q^{t}_{m}} \right) \Big/ \left(\frac{\sum_{m=1}^{M} \hat{p}^{0}_{m} q^{0}_{m}}{\sum_{m=1}^{M} q^{0}_{m}} \right) \tag{15}$$

For goods or services with slight product differentiation we would advise a measure based on (15). Of course the quality adjustments need not use hedonic regressions. They may be much simpler due to the addition of a single feature or option for which cost or market estimates of their value are available. Bear in mind that the items are matched in each period. The problem is quite different from the usual use of hedonic indexes as proposed by Pakes (2003), Silver and Heravi (2003), and Triplett (2006), and for which hedonic indexes of various formulations are used to control for quality variation over time, as new higher-quality models of, say, personal computers replace older ones. Here the problem is of cross-sectional quality variation and the need to make models sold at the same time homogeneous, that is, of similar quality. The argument is phrased in terms of matched samples, so the quality-adjustment argument in the aforementioned papers does not apply, but the unit value argument remains.

There is a need for a weighted average of (15) and (1), but a problem as to what the weights should be. One approach is to consider what we mean by "comparable." Consider an example of measuring price changes of a sample of different models of cars over time. Given that the quality of models of cars sold in a specific period differs, a hedonic regression with prices on the left-hand side and price-determining characteristics on the right may be estimated on a cross-sectional basis and may explain much of the cross-sectional price variation. There is a sense in which features of the car – horsepower, coating, miles per gallon, type of entertainment system, and so forth – explain much of the price variation. Consider another data set of a similar model of automobile sold by different dealers for which a hedonic regression would explain a much smaller proportion of the price variation. The cars in the first case are relatively heterogeneous, while those in the second are relatively homogeneous, and the explanatory power of a hedonic regression provides an indicator of the extent of any heterogeneity.

Thus, the weight for the heterogeneity-adjusted unit value index (15) might be the ratio of the sum of squared errors from the (assumed well-specified) hedonic regression (SSE) to the total sum of squares (SST), and the weight for the Fisher price index given by (1) the ratio of the (explained) regression sum of squares (SSR) to SST. The weighted average is given by

$$P^{*}_{UV}\bar{w}_{U} + P_{F}(1 - \bar{w}_{U}) = \frac{\sum_{m=1}^{M} \hat{p}^{t}_{m} q^{t}_{m} / \sum_{m=1}^{M} q^{t}_{m}}{\sum_{m=1}^{M} \hat{p}^{0}_{m} q^{0}_{m} / \sum_{m=1}^{M} q^{0}_{m}} \times \bar{w}_{U} + \sqrt{\frac{\sum_{m=1}^{M} p^{t}_{m} q^{t}_{m}}{\sum_{m=1}^{M} p^{0}_{m} q^{t}_{m}} \times \frac{\sum_{m=1}^{M} p^{t}_{m} q^{0}_{m}}{\sum_{m=1}^{M} p^{0}_{m} q^{0}_{m}}} \times (1 - \bar{w}_{U}) \tag{16}$$

where

$$\bar{w}_U = \frac{SSE}{SST} \text{ and } (1 - \bar{w}_U) = \frac{SSR}{SST} = R^2$$

Note that the weights in (16) have a bar over them; they are an arithmetic mean of the weights from period 0 and period t for hedonic regressions in Equation (13) for $\tau = 0, t$.

An appropriate index should have the property that if all price variation is explained by the hedonic regression, the index is a Fisher index; if none of the price variation is explained by the hedonic regression, the index is a unit value index; as the percentage of price variation explained by the hedonic regression increases, so too will the weight given to the Fisher component. Equation (16) satisfies all these criteria.

The use of such weights is but one proposal. Consider the case of television sets. A hedonic regression could be estimated over all screen sizes, with dummies for these screen sizes and variables for the other quality characteristics. But while the regression might attribute a 30% premium for a 32-inch screen over a 14-inch one, while controlling for other variables, it is unlikely that consumers would consider the two models as substitutes even when an allowance has been made for the screen size and other variables. Thus, the regression should be undertaken for similar goods in the sense that there is some substitutability.

It might be argued that substitutability should be the main concept behind the weighting system. However, this is problematic. Substitutability can exist for goods and services for which quality-adjustments for unit values are not feasible and the concept of an average price not meaningful, for example beef and chicken. However, as indicated above, the first step should be to identify a cluster of goods that are comparable and substitutable or exchangeable, for example television sets of a similar screen size, and then use the hedonic framework. An approach to clustering consumer goods and services has been introduced as the concept of *consumption segment by purpose* (Commission of the European Communities 2007).

5. An Empirical Example: Using Scanner Data

The empirical work utilizes monthly scanner data for television sets (TVs) from the bar-code readers of UK retail outlets from January 2001 to December 2001. The scanner data were supplemented by data from price collectors from outlets without bar-code readers, though sales from outlets without bar-code readers were negligible. Each observation is a model of a TV in a given month sold in one of four different outlet types: multiple chains, mass merchandisers (department stores), independents and catalogue stores. The sample was devised to include only models of TVs that were sold in all twelve months. This has the advantage of replicating the matched model methodology employed by statistical offices for consumer and producer price indexes, as well as clarifying that the effects we identify as not being due to new and old models of differing quality entering and leaving the market (Silver and Heravi 2005). We further limit the sample to a narrow range of screen size, that is, 10- and 14-inch TV screens, since it may be argued that larger TVs with larger screen sizes serve a different consumer need. The data set included series for 94 such models in each month, accounting for sales of 0.37 million TVs worth £49 million.

Table 1. Price index formulas, unit value indexes and the substitution effect

	Laspeyres	Paasche	Fisher	Unit value	Substitution effect
January	1.000	1.000	1.000	1.000	
February	0.998	0.998	0.998	1.060	1.000
March	0.996	1.001	0.999	1.049	1.003
April	0.996	0.982	0.989	0.968	0.993
May	0.999	0.983	0.991	0.985	0.992
June	0.964	0.939	0.951	0.957	0.987
July	0.959	0.922	0.940	0.994	0.980
August	0.945	0.917	0.931	0.990	0.985
September	0.941	0.916	0.928	0.972	0.987
October	0.950	0.935	0.942	0.981	0.992
November	0.948	0.925	0.936	1.000	0.987
December	0.933	0.925	0.929	0.962	0.996

Hedonic regressions were estimated for each month to remove the effect on price of cross-sectional product heterogeneity. The variable set for the regressions included: (i) 17 brand dummies; (ii) size of screen, 10- and 14-inch; (iii) Nicam stereo sound; (iv) on-screen text retrieval news and information panels from broadcasting companies, in order of sophistication: teletext and fastext; (v) three types of reception systems; (vi) continental monitor style; (v) flat & square and super-planar tubes; (vi) s-vhs socket; (vii) DVD playback or DVD recording; and (viii) the outlet types, multiple chains, mass merchandisers (department stores), independents and catalogue stores.

Table 1 and Figure 1 show Laspeyres, Paasche, Fisher price and unit value indexes. Laspeyres and Paasche generally act as upper and lower bounds on Fisher, as expected from a negative correlation between relative price and quantity changes. From April to November all values of a weighted correlation coefficient between relative prices and quantities were negative, reaching a maximum of -0.218 in July. The magnitude of the difference between both Fisher and Laspeyres and Paasche and Fisher price indexes is given by the substitution effect (Equation 11). The magnitude of the substitution effect, measured as deviation from unity, can be seen from the calculated values in Table 1 to be generally small. An exception is July, during which the negative correlation increases in

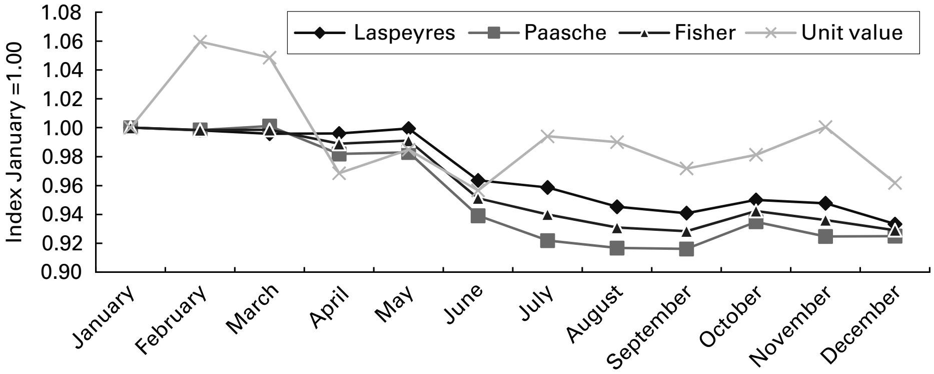

Fig. 1. Unit value and price indices for 14in TVs

Table 2. *Decomposition of ratio of unit value to price index number formulas*

	Ratio of unit value index to				
	Laspeyres	Fisher	Paasche	Substitution effect	Level effect
January	1.000	1.000	1.000		
February	1.061	1.061	1.061	1.000	1.062
March	1.053	1.050	1.047	1.003	1.048
April	0.972	0.979	0.986	0.993	0.986
May	0.985	0.994	1.002	0.992	1.002
June	0.993	1.006	1.019	0.987	1.019
July	1.037	1.057	1.078	0.980	1.079
August	1.047	1.064	1.080	0.985	1.081
September	1.033	1.047	1.061	0.987	1.062
October	1.033	1.041	1.050	0.992	1.050
November	1.056	1.069	1.082	0.987	1.083
December	1.031	1.035	1.040	0.996	1.040

magnitude as more price-conscious consumers hit the summer sales, leading to an increase in the Laspeyres-Paasche gap. The unit value index for these differentiated products has quite dissimilar changes to the price indexes.

Table 2 shows unit value indexes to generally exceed Fisher price indexes, with the levels effect as a significant and generally positive factor in this regard. The levels effect is somewhat volatile, over eight percent in August and November, but had little influence in May. Table 2 shows substitution to generally have a countervailing effect to that of the levels effect, bringing the unit value index closer to the price indexes, but the levels effect generally dominates.

The unit value index is compiled for heterogeneous TVs comprising 17 brands and several characteristics as detailed above. They are broadly comparable items. We estimated hedonic regressions for each month and quality-adjusted the prices as outlined in Equations (14) to (16). The quality-adjusted unit value index is given in Figure 2. Between January and February the unit value index increased by about six percent, reflecting an increasing quantity of purchases directed to more expensive sets and

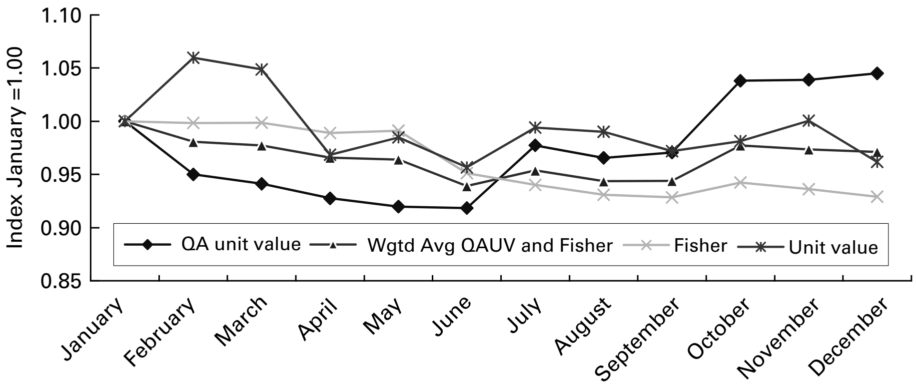

Fig. 2. *Quality adjusted unit value and Fisher price indices*

better brands. Yet when we take out the effect of such quality differences, the change in the mix of the characteristics purchased, there is a *fall* in the prices of about five percent. Consumers are paying more on average for better sets, but given their valuation of what the improved mix in characteristics is worth, the result is an overall fall in average (unit) prices. Similarly, from November to December the unit value index fell by about four percent, as the bundle of TVs purchased included a higher proportion of cheaper models, but the quality-adjusted unit value index actually increased, reflecting the fact that the fall in average prices did not compensate for the fall in quality that gave rise to it.

Also given in Figure 2 are the Fisher index, Equation (1), and the weighted average of the quality-adjusted unit value index and the Fisher index (Equation 16). The average value of R^2 for the hedonic regressions was 0.63, with the quality-adjusted unit value index receiving the lower weight of 0.37. As a result, the weighted average tracks the Fisher price index more closely than the quality-adjusted unit value index.

6. Conclusions

For the aggregation of homogeneous items, the unit value index is the best index, superlative index numbers are biased, and for the aggregation of heterogeneous items, superlative index numbers are the best index and unit value index numbers are biased.

A contribution of this article has been that the factors determining the difference between unit value indexes and Laspeyres, Paasche, and Fisher price indexes were established in a formal derivation in Section 3. They comprise a substitution bias and a levels effect. The conditions for the unit value index to equal the price indexes were established as implausible.

For items that are very similar, a unit value index remains appropriate, for it is necessary to capture the effect of a shift in quantities to higher/lower price levels, and price indexes do not properly do this. Quality adjustments to the prices to mitigate price dispersion due to the slight product heterogeneity and quality-adjusted unit value indexes would be appropriate.

The determination of whether or not an item is homogeneous is critical to the choice of index number formula, but in practice many items are broadly comparable, and neither a unit value nor a Fisher price index is appropriate. This article has drawn attention to the problem with an illustrative proposal using scanner data on television sets. The more similar the aggregated items, the stronger the case for a heterogeneity-adjusted unit value. It follows that an appropriate formula may be based on an average of a heterogeneity-adjusted unit value index and a Fisher price index. The weighting ascribed to each should be an indicator of the similarity of the items. A possible indicator explored in this article is the extent to which the price variation can be explained by price-determining characteristics: the (explained) sum of squares from a hedonic regression. While the discussion has been phrased in terms of hedonic regression analysis, the principles apply to simpler quality-adjustment procedures.

7. References

Aizcorbe, A. and Nestoriak, N. (2008). 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.

Alterman, W. (1991). Price Trends in U.S. Trade: New Data, New Insights. In International Economic Transactions, Peter Hooper and J. David Richardson (eds). Chicago: University of Chicago Press, 109–143.

Angermann, O. (1980). External Terms of Trade of the Federal Republic of Germany Using Different Methods of Deflation. Review of Income and Wealth, 26, 367–385.

Balk, B.M. (1998). On the Use of Unit Value Indexes as Consumer Price Subindexes. Proceedings of the Fourth Meeting of the International Working Group on Price Indexes, 112–120. Washington DC: U.S. Bureau of Labor Statistics. Available at: http://www.ottawagroup.org

Balk, B.M. (2005). Price Indexes for Elementary Aggregates: The Sampling Approach. Journal of Official Statistics, 21, 675–699.

Balk, B.M. (2008). Price and Quantity Index Numbers. Cambridge: Cambridge University Press.

Bortkiewicz, L.v. (1923). Zweck und Struktur einer Preisindexzahl. Nordisk Statistisk Tidsskrift, 2, 369–408. [In German]

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

Commission of the European Communities (2007). Commission Regulation (EC) No 1334/2007, amending Regulation (EC) No 1749/96.

Commission of the European Communities, International Monetary Fund, Organisation for Economic Co-operation and Development, United Nations, World Bank (2009). System of National Accounts 2008. Brussels/Luxemburg, New York, Paris, Washington: EC, IMF, OECD, UN, World Bank. New York: United Nations. Available at: http://unstats.un.org/unsd/nationalaccount/sna2008.asp

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

de Haan, J. (2004). Estimating Quality-Adjusted Unit Value Indexes: 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.

de Haan, J. (2007). Hedonic Price Indexes: A Comparison of Imputation, Time Dummy and Other Approaches. Paper presented at the Seventh EMG Workshop, Sydney, Australia, December 12–14.

Diewert, W.E. (1976). Exact and Superlative Index Numbers. Journal of Econometrics, 4, 114–145.

Diewert, W.E. (1978). Superlative Index Numbers and Consistency in Aggregation. Econometrica, 46, 883–900.

Diewert, W.E. (1995). Axiomatic and Economic Approaches to Elementary Price Indexes. Department of Economics, University of British Columbia Discussion Paper 95–01.

Diewert, W.E. and von der Lippe, P. (2010). Notes on Unit Value Index Bias. Unpublished report, revised, October.

Hausman, J. and Leibtag, E. (2008). CPI Bias from Supercenters: Does the BLS Know that Wal-Mart Exists? In Price Index Concepts and Measurement, National Bureau of Economic Research, E. Diewert, J. Greenlees, and C. Hulten (eds).

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

International Labour Office (ILO), IMF, OECD, UN ECE, World Bank (2004b). Producer Price Index Manual: Theory and Practice. Washington: International Monetary Fund. Available at http://www.imf.org/np/sta/tegppi/index.htm

International Labour Office (ILO), IMF, OECD, Eurostat, United Nations, World Bank (2009). Export and Import Price Index Manual: Theory and Practice. Washington: IMF.

Kotler, P. and Keller, K. (2009). A Framework for Marketing Management. New Jersey: Prentice-Hall.

Pakes, A. (2003). A Reconsideration of Hedonic Price Indexes with an Application to PCs. The American Economic Review, 93, 1576–1593.

Párniczky, G. (1974). Some Problems of Price Measurement in External Trade Statistics. Acta Oeconomica, 12, 229–240.

Ruffles, D. and Williamson, K. (1998). Deflation of Trade in Goods Statistics: Derivation of Price and Volume Measures from Current Price Values. Economic Trends, 521, April. Reproduced in Office for National Statistics (ONS), Economic Trends, Digest of Articles No. 1, London: ONS, 1998.

Silver, M. (2009). Do Unit Value Export, Import, and Terms of Trade Indexes Represent or Misrepresent Price Indexes? IMF Staff Papers, 56, 297–322.

Silver, M. and Heravi S. (2002). 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.

Silver, M. and Heravi, S. (2003). The Measurement of Quality-Adjusted Price Changes. Scanner Data and Price Indexes. In National Bureau of Economic Research, Studies in Income and Wealth, M. Shapiro and R. Feenstra (eds). 61. Chicago: University of Chicago Press, 277–317.

Silver, M. and Heravi, S. (2005) A Failure in the Measurement of Inflation: Results From a Hedonic and Matched Experiment Using Scanner Data. Journal of Business & Economic Statistics, American Statistical Association, 23, 269–281, July.

Silver, M. and Heravi, S. (2007). Why Elementary Price Index Number Formulas Differ: Price Dispersion and Product Heterogeneity. Journal of Econometrics, 140, 874–883.

Triplett, J.E. (2006). Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes. Directorate for Science, Technology and Industry. Paris: OECD.

Triplett, J.E. (2006). Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes. Special Application to Information Technology Products. Paris: OECD.

von der Lippe, P.M. (2007). Price Indexes and Unit Value Indexes in German Foreign Trade Statistics. Unpublished Paper, University of Duisburg-Essen. Available at http://mpra.ub.uni-muenchen.de/5525/

Received December 2009
Revised May 2011