DISKUSSIONSBEITRAG

aus der Fakultät für
WIRTSCHAFTSWISSENSCHAFTEN

der

UNIVERSITÄT DUISBURG - ESSEN
Campus Essen

Nr. 185
Oktober 2010

Price Indices on the Basis of Unit Values – Unit Value Indices as Proxies for Price Indices

Peter von der Lippe

Universitätsstraße 12
45117 Essen

Price Indices on the Basis of Unit Values
Unit Value Indices as Proxies for Price Indices

Peter M. von der Lippe

Abstract

In some countries a Paasche price index based on unit values is compiled as a proxy for a true Paasche (or Laspeyres) price index on the basis of prices. This is for example the case in German foreign trade statistics. Unit values are average prices referring to a "commodity number" (CN), that is an aggregate of (more or less homogeneous) commodities defined by a commodity classification. They are often easily available as free by-products of other statistics (foreign trade or wage statistics for example) and therefore less costly than true transaction prices of well-defined specific products as they are in general reported in price statistics.

Changes in unit values between two points in time, however, do not only reflect a price movement but also changes in the quantities transacted. They are, in other words, affected by a structural component, the changing mix of commodities within CNs and therefore biased relative to genuine price indices. The focus of the paper is on explaining this bias. It is shown that amount and sign of the "unit value bias" depends on the correlation between the change of quantities of the goods included in the CNs and their respective *base period* prices, while current period prices do not matter. This result is useful as it may help to define "homogeneity" with respect to CNs and thus conditions under which unit values may be acceptable as (cost-effective) substitutes for prices.

This paper is a revised version of my contribution to the 11th Ottawa Group Meeting in Neuchatel (Switzerland) 27th to 29th May 2009 http://www.ottawagroup2009.ch/

Key words: Price index, unit value index, unit values, axioms, foreign trade statistics, Bortkiewicz, Drobisch.

JEL: C43, C80, E01, F10

1. Introduction

Only few countries (among which Germany and Japan) are able to provide on a monthly basis both, a unit value index (UVI) and a true price index (PI) for measuring the price development in export and import. UVIs are internationally not uncommon and increasingly relevant. Some recent empirical studies of the impact of the methodological differences between these two indices (Silver (2007), Silver (2008), von der Lippe (2007b)),[1] however, gave rise to concerns in that UVIs can be viewed only as an unsatisfactory surrogate of PIs.

The problem with UVIs is, however that the term is used for quite different indices. On the one hand there are indices actually compiled in official statistics as for example the German export and import[2] and which are in the focus of this paper (and of v.d.Lippe (2007b)). On the other hand the term UVI is also in use for another index, primarily of theoretical interest, that should preferably be called "Drobisch index" in the honour of W. M. Drobisch (1871)[3], and to which most of the present literature under the key word "unit value index" is devoted. This applies for example to Párniczky (1974), Balk (1994), (1998), (2005), (2008; 72-74) and Diewert (1995), (2004). Diewert (2010) and v.d.Lippe (2007a; 18-20, 415-428) examine both types of unit value indices (or "price indices on the basis of unit values"). We will refer to Diewert's paper in particular in section 3.

It is therefore necessary to start with some (more than usual) remarks concerning definitions, notation and terminology, and this will be done in section 2. Unit values are a sort of average prices (for a *group* of goods) and in an index on the basis of unit values they take the part prices of individual goods have in the case of a price index (which thus uses data on a much more disaggregated level).

Section 3 compares the Drobisch index to the "normal" Paasche and Laspeyres price index and introduces the method of describing biases in terms of covariances which also will be amply used in the next sections. This part of the paper is based on Diewert (2010). We briefly report some equations of Diewert in which he explored earlier "bias" results (relative to the index of Drobisch) of Balk and Párniczky together with some new results of Diewert.

In section 4 the "unit value index" (UVI) as actually compiled in Germany for exports and imports is compared to a price index of Paasche. This UVI differs from the Drobisch index and may be viewed as a Paasche index compiled in two stages where unit values instead of prices are used in the first, or "low level aggregation" stage. Again some interesting covariance expressions are found for the "bias" showing that the bias is apparently closely related to the heterogeneity of the aggregate underlying the calculation of unit values.

It has been suggested[4] that all bias formulas developed in sec. 3 and 4 may be viewed simply as special cases of a general theorem on linear indices found by L. v. Bortkiewicz. In section 5 it is shown that this is in fact the case.

---

[1] I am very grateful to Erwin Diewert for his valuable remarks to my former draft of this paper and to the earlier texts quoted above. Also Jens Mehrhoff was very helpful. Some of the hypotheses examined in v.d.Lippe (2007b) as well as conceptual and empirical differences between customs-based UVIs as opposed to survey-based price indices (PIs) are described in the annex of this paper. The empirical part in particular is owed to Jens Mehrhoff. See von der Lippe and Mehrhoff (2008).

[2] The method of a UVI is also quite common in the case of indices of wages or prices for certain services (air transport for example). Unit values will definitely gain importance because it can be expected that scanner data will be ever more frequently used in statistics.

[3] See the paper of L. von Auer in this special edition of the Jahrbücher.

[4] It was actually Bert Balk who expressed this idea in the discussion of my presentation in a conference in Nuremberg (Sept. 2010).

In section 6 some conclusions and suggestions for further research are given and in the annex we briefly present some basic information concerning the German foreign trade statistics as well as some results of our empirical study comparing the UVI of exports and imports to the respective (Laspeyres) price indices.

2. Unit values and price indices on the basis of unit values

2.1. Definition and some properties of unit values

It is important to realize that unit values are defined only for several goods grouped together in a sub collection of goods defined by a classification of products (e.g. of commodities for production or for foreign trade statistics). The relevant unit of the classification is called "commodity number" (CN) and the unit value of the k^{th} CN (k = 1, …, K) in period t is a kind of average price of the n_k goods of this CN

(1) $$\tilde{p}_{kt} = \frac{\sum_j p_{kjt} q_{kjt}}{\sum_j q_{kjt}} = \frac{V_{kt}}{Q_{kt}} = \sum_{j=1}^{n_k} p_{kjt} \frac{q_{kjt}}{Q_{kt}} = \sum p_{kjt} m_{kjt} \quad \text{in periods } t = 0, 1$$

where the summation takes place over the j = 1, …, n_k ($n_k < n$) goods of a CN and refers to periods 0 (base period), or 1 (reference period) respectively. Hence the unit *value* of an aggregate (for example a CN or any *collection* of goods) is simply the *value* V_{kt} divided by the total *quantity* of the aggregate under consideration. The notion of "unit value" is quite relevant in statistics because situations are not infrequent in which data on both, values as well as the underlying quantities are readily available. They may be even more common than situations in which one can dispose of all n_k individual prices.

The properties of unit values are, however, not entirely satisfactory.[5]

If no data on quantities is given but only the number of units (n_k) instead, the unit value $\tilde{p}_{kt}$ reduces to the average price $\bar{p}_{kt} = \sum_j p_{kjt} / n_k$ of the k-th CN. We also have $\tilde{p}_{kt} = \bar{p}_{kt}$ if all n_k prices in t are equal $p_{kjt} = \bar{p}_{kt}$ $(\forall j = 1,...,n_k)$.[6]

If unit values can be calculated for the same (k-th) CN for the two periods t = 0, and t = 1 the *ratio of unit values*

(2) $$\frac{\tilde{p}_{k1}}{\tilde{p}_{k0}} = \frac{V_{k1}}{V_{k0}} \frac{Q_{k0}}{Q_{k1}} = \frac{V_{k1}}{V_{k0}} : \frac{Q_{k0}}{Q_{k0}} = \frac{V_{01}^k}{\tilde{Q}_{01}^k}$$

can be regarded as a "low level" price index. Note, however, that this ratio is not a mean of price relatives

(2a) $$\frac{\tilde{p}_{k1}}{\tilde{p}_{k0}} = \frac{Q_{k0}}{Q_{k1}} \sum_j \frac{p_{kj1}}{p_{kj0}} \left(\frac{p_{kj0} q_{kj1}}{\sum_j p_{kj0} q_{kj0}} \right) = \sum_j \frac{p_{kj1}}{p_{kj0}} \left(\frac{p_{kj0} m_{kj1}}{\sum_j p_{kj0} m_{kj0}} \right)$$

---

[5] From eq. 1 follows that unit values violate proportionality (and therefore also identity). If all n_k individual prices change λ–fold ($p_{kj1} = \lambda p_{kj0}$ ∀j) the unit value as a rule does not change λ–fold unless the quantity-structure coefficients m are constant ($m_{kj1} = m_{kj0}$). Unit values also violate commensurability which is due to the fact that Q_{kt} is affected from changes in the quantity units to which the price quotations refer.

[6] In what follows we will have both types of covariances, those in which the unweighted arithmetic mean of prices appears and also those in which the unit value appears. The fact that unit values may coincide with arithmetic means irrespective of the quantities involved implies that different bias-formulas in terms covariances may well coincide. This applies in particular to the case of a vanishing covariance, which is usually the interesting situation of "no bias".

because the sum of the weights (in brackets in the rightmost expression) do not add up to unity[7] unless $m_{kj1} = m_{kj0}$ for all j. This sum is rather

(2b) $$\sum_j \frac{p_{kj0}q_{kj1}}{\tilde{p}_{k0}Q_{k1}} = \frac{Q_{01}^{L(k)}}{\tilde{Q}_{01}^{k}} = S_{01}^{k}$$

where $Q_{0t}^{L(k)}$ is the Laspeyres quantity index of the k[th] CN. When no price changes within each CN and therefore the unit value ratios $\tilde{p}_{k1}/\tilde{p}_{k0}$ reflect only a structural change (in the absence of a price change) we get

(2c) $$\tilde{p}_{k1}/\tilde{p}_{k0} = Q_{01}^{L(k)}/\tilde{Q}_{01}^{k} = S_{01}^{k}$$

for each k instead of the general formula of eq. 2 and it will be shown (sec. 4) that the S_{01}^{k} terms are the key figures to explain the "unit value bias".

2.2. Price indices on the basis of unit values

It appears useful to make a distinction between two types of "price indices based on unit values" (instead of prices). In addition to the well known Drobisch index (part 2.2.1) we have a less well understood class of indices (see part 2.2.2) which are necessarily compiled in two steps and make use of unit values as building blocs in the first step.

2.2.1. Drobisch index

This index defined by

(3) $$P_{01}^{D} = \frac{\sum_k \sum_j p_{kj1}q_{kj1} \Big/ \sum_k \sum_j q_{kj1}}{\sum_k \sum_j p_{kj0}q_{kj0} \Big/ \sum_k \sum_j q_{kj0}} = \frac{Q_0}{Q_t}\frac{\sum_k \sum_j p_{kj1}q_{kj1}}{\sum_k \sum_j p_{kj0}q_{kj0}} = \frac{V_{01}}{\tilde{Q}_{01}} = \frac{\tilde{p}_1}{\tilde{p}_0}$$

is unfortunately more often than not called "unit value index" although it is quite different from an index defined later by eq. 4 (the index PU^P instead of P^D) which is also called "unit value index". To avoid confusion and this ambiguity the index P^D should better be called "Drobisch index" as it was being proposed by Drobisch (1871).[8]

It should be noted, that the problem with the Drobisch index is that it is in general not possible - let alone meaningful - to summate over the quantities of all $n = \Sigma n_k$ commodities, as required in the compilation of "Drobisch's" index. Hence unlike the K terms $Q_{kt} = \sum_j q_{kjt}$ the term $Q_t = \sum_k \sum_j q_{kjt} = \sum_k Q_{kt}$ (and therefore also $\tilde{p}_t$) is in general not defined. Even if there were a common unit of measurement for all quantities sums of quantities have a meaningful interpretation only if they are taken over related goods such as the commodities of a commodity number (CN). Unsurprisingly the Drobisch index is *not* compiled in the practice of official statistics and primarily of theoretical interest. This applies also to $Q_{01}^{D} = \tilde{Q}_{01} = Q_1/Q_0$ which may be called "Drobisch quantity index".

---

[7] By the same reason unit value ratios as opposed to price relatives p_{kj1}/p_{kj0} violate proportionality.

[8] The label "Drobisch's index" is, however, uncommon which is possibly due to the fact that it is already in use for another index also advocated by M. W. Drobisch, viz. the *arithmetic* mean of a Laspeyres and a Paasche price index. For more details concerning his index P^D see also the contribution of von Auer in this journal. However von Auer does not mention the "unit value index" of official statistics which is pr.

Note also that the P^D index is not simply a weighted mean of ratios of unit values $\tilde{p}_{k1}/\tilde{p}_{k0}$ because

(3a) $$P_{01}^D = \sum_k \frac{\tilde{p}_{k1}}{\tilde{p}_{k0}} \left(\frac{\tilde{p}_{k0}\sigma_{k1}}{\sum_k \tilde{p}_{k0}\sigma_{k0}} \right) \text{ where } \sigma_{kt} = Q_{kt} \Big/ \sum_k Q_{kt} = Q_{kt}/Q_t$$

which means that in general (with changing structure of quantities between CNs, that is $\sigma_{k1} \neq \sigma_{k0}$) the index P^D can *not* be viewed as being aggregated over "low level" unit value ratios $\tilde{p}_{k1}/\tilde{p}_{k0}$. Results gained from studying the "elementary" or low level ratios $\tilde{p}_{k1}/\tilde{p}_{k0}$ cannot simply be translated into statements relevant for the Drobisch index. Eq. 3a also shows that P^D not only reflects changes of the structure of quantities *within* CNs (via $\tilde{p}_{k1}/\tilde{p}_{k0}$ which depends on the m_{kj1} and m_{kj0} coefficients)[9] but also *between* CNs (unless $\sigma_{k1} = \sigma_{k0}$ holds for all k = 1, …, K).

2.2.2. Paasche index of unit values

According to eq. 3a the Drobisch index is not a weighted arithmetic mean of K ratios of unit values $\tilde{p}_{k1}/\tilde{p}_{k0}$. However, such a mean is in fact the following index[10]

(4) $$PU_{01}^P = \sum_k \frac{\tilde{p}_{k1}}{\tilde{p}_{k0}} \frac{\tilde{p}_{k0}Q_{k1}}{\sum_k \tilde{p}_{k0}Q_{k1}} = \frac{\sum_k \tilde{p}_{k1}Q_{k1}}{\sum_k \tilde{p}_{k0}Q_{k1}} = \frac{\sum_k \tilde{p}_{k1}\sigma_{k1}}{\sum_k \tilde{p}_{k0}\sigma_{k1}}.$$

By contrast to the Drobisch index this index does not require the calculation of a total unit value of all goods (and maybe also services) at two points in time, 0 (base period) and 1 (present period), that is $\tilde{p}_0$ and $\tilde{p}_1$. It necessitates only the calculation of unit values for specific CNs, $\tilde{p}_{k0}$ and $\tilde{p}_{k1}$ respectively.

It is in particular useful that it is not necessary to add quantities across all CNs in order to get a total quantity $Q_t = \sum_k Q_{kt}$. The need for such quantities is just the reason for the fact that the Drobisch index formula is of limited use only for the official price statistics (at least when a broader aggregate of goods is concerned). In practice we cannot meaningfully add kilograms of farm products to kilograms of cars, or kilograms of TV-sets let alone "add" over quantities without referring to a common unit of measurement. Thus it clearly is an advantage that the index PU^P requires only quantities Q_{kt}. They can more reasonably be established than the all-items sum Q_t because they are defined for a CN only, and the CNs are in turn explicitly formed to cover fairly similar products.

The index PU^P (eq. 4) is called "unit value index" in German official statistics. However, this name is notoriously mistaken for the index P^D (eq. 3). So presently the same name is in use for quite different formulas. To avoid ambiguity and confusion we propose to call

- the P^D index (eq. 3) *Drobisch index* instead of (the unfortunately very common term) "unit value index", and
- the PU^P index (eq. 4) *Paasche (price) index of unit values* (as opposed to the "normal" or "true" Paasche price index defined in eq. 7) or PU^P index for short instead of calling it (again) "unit value index"; to add "price" in the (unfortunately) rather long name for the PU^P index is necessary because there are also quantity indices in which use is made

---

[9] This applies for example to the indices discussed in sec. 2.2.2.
[10] This, of course applies also to PU^L defined in eq. 4a.

of unit values (as weights; see eq. 5) and the PU^P index can reasonably be compared only with a price index.[11]

There seems to be no obvious reason why the Paasche formula should be preferred to the Laspeyres formula, such that

$$\text{(4a)} \qquad PU_{01}^{L} = \frac{\sum \tilde{p}_{k1} Q_{k0}}{\sum \tilde{p}_{k0} Q_{k0}}$$

appears to be equally useful and should be called "Laspeyres (price) index of unit values". However, our motivation to study this type of "unit value" index was initially the German statistics of export and import prices, and there PU^P (rather than PU^L) indices are in use. We are therefore going study in what follows primarily the Paasche variant PU^P (eq. 4).

Three final remarks may be appropriate

1. Indices of the PU type, that is PU^P and PU^L may be viewed as two-stage or two-level index compilations where in the first (low) level use is made of unit values rather than prices. There are, however, some differences to the usual notion of "low level" aggregation which normally applies to situations in which no information about quantities is available, and therefore no weights can be established (unlike the upper level for which the introduction of weights is characteristic). Moreover in low-level aggregation prices usually are referring to the same commodity in different outlets. Here (and also in the case of using scanner data for the purposes of price statistics) quantities are known and unit values refer to different commodities grouped together by a classification.

2. In a unit value index for the measurement of *prices*, that is in PU indices quantities act as weights. It is also possible to measure the dynamics of *quantities* on the basis of sums of quantities Q_{kt} (t = 0, 1) which then yields QU-indices and where unit values consequently take the part of weights. So for example

$$\text{(5)} \qquad QU_{01}^{L} = \frac{\sum Q_{k1} \tilde{p}_{k0}}{\sum Q_{k0} \tilde{p}_{k0}}$$

   is a Laspeyres type quantity index of unit values.[12] Out of the many possible variants of PU and QU indices respectively, we choose only two on which we focus in what follows, viz. PU_{01}^{P} and QU_{01}^{L}.

3. We saw that the Drobisch index P^D (like the value index V_{0t} but unlike the PU^P index) is not a mean value of unit value ratios $\tilde{p}_{k1}/\tilde{p}_{k0}$ (unless $\sigma_{k1} = \sigma_{k0}$ holds for all K *between-CNs quantity shares* σ), and in the same way the PU^P index is (unlike the true Paasche price index P^P defined in eq. 7) not a weighted arithmetic mean of price relatives

[11] Diewert suggested two other names for this index, "hybrid Paasche" (P_{HP}) index, and "Drobisch-Paasche" index. P_{HP} does not seem to be satisfactory because there are indices in which use is made of both, prices or price indices from an outside source (for certain sub-aggregates) on the one hand and unit values (in the case of other sub-aggregates) on the other, and such indices may (rightly) be called "hybrid". This applies for example for Canada; see the information given by Statistics Canada in the Internet concerning the Canadian "International Merchandise Trade Price Index (IMTPI)" according to which this index "blends unit value indices with specified indexes", or the index is "based in part on actual unit values" and also makes "use of price relatives provided by other sources". The term "Drobisch-Paasche" on the one hand raises the problem of how to call the "true" Paasche price index (Paasche- Paasche?).

[12] Note that this differs from the "unit value quantity index" (better Drobisch's quantity index) $\Sigma_k Q_{k1}/\Sigma_k Q_{k0}$ as mentioned above. The QU^P (analogously to (5) for QU^L) makes use of period 1 unit values.

(6) $$PU_{01}^{P} = \sum_k \sum_j \frac{p_{kj1}}{p_{kj0}} \cdot \frac{p_{kj0} q_{kj1}}{\sum_k \tilde{Q}_{01}^{k} \sum_j p_{kj0} q_{kj0}} = \sum_k \sum_j \frac{p_{kj1}}{p_{kj0}} \cdot w_{kj}$$

unless the weights add up to unity ($\sum_k \sum_j w_{kj} = 1$) which is for example the case if quantities within each CN change at the same rate

(6a) $$q_{kj1} / q_{kj0} = \tilde{Q}_{01}^{k}$$

or don't change at all $q_{kj1} = q_{kj0}$ (so that $\tilde{Q}_{01}^{k} = 1$ for all k). This means in other words, that the *within-CN quantity shares*, reflecting the structure of quantities remains constant). Interestingly it turns out in sec. 4 that these are also the conditions under which no bias of the PU^P index occurs.[13]

2.2.3. True price indices and indices based on unit values

In this part indices we are going to introduce "true" price indices (PIs) in order to compare them to (price) indices on the basis of unit values (UVIs, equations 3, 4, 4a). For this purpose we make the assumption - unrealistic though[14] - that a price index is comprising all K CNs with all $n = \Sigma n_k$ commodities so that the same positive prices are involved in both types of indices, PIs and UVI. The difference is only that in the case of UVIs the underlying individual prices are not known. We then get for the indices of Paasche and Laspeyres

(7) $$P_{01}^{P} = \frac{\sum_{k=1}^{K} \sum_{j=1}^{n_k} p_{kj1} q_{k1}}{\sum_{k=1}^{K} \sum_{j=1}^{n_k} p_{kj0} q_{kj1}} = \frac{\sum_k \tilde{p}_{k1} Q_{k1}}{\sum_k \sum_j p_{kj0} q_{kj1}}, \text{ and}$$

(7a) $$P_{01}^{L} = \frac{\sum_{k=1}^{K} \sum_{j=1}^{n_k} p_{kj1} q_{k0}}{\sum_{k=1}^{K} \sum_{j=1}^{n_k} p_{kj0} q_{kj0}} = \frac{\sum_k \sum_j p_{kj1} q_{kj0}}{\sum_k \tilde{p}_{k0} Q_{k0}}$$

For some of the following considerations it is useful to note that the following decompositions of the value index V_{01} exist

(8) $$V_{01} = PU_{01}^{L} QU_{01}^{P} = PU_{01}^{P} QU_{01}^{L} = \sum p_1 q_1 / \sum p_0 q_0, \text{ and}$$

(8a) $$V_{01} = P_{01}^{L} Q_{01}^{P} = P_{01}^{P} Q_{01}^{L}.$$

From this follows that equality of PU^P and P^P amounts to s the for

$$PU_{01}^{P} = P_{01}^{P} \Rightarrow \frac{\sum_k \tilde{p}_{k1} Q_{k1}}{\sum_k \tilde{p}_{k0} Q_{k1}} = \frac{\sum_k \sum_j p_{kj1} q_{kj1}}{\sum_k \sum_j p_{kj0} q_{kj1}} = \frac{\sum_k \tilde{p}_{k1} Q_{k1}}{\sum_k \sum_j p_{kj0} q_{kj1}} \text{ or equivalently}$$

[13] It is interesting to note that condition (6a) amounts to equality of PU^P and the true Paasche price index P^P. This is also instantly shown in the next paragraph (see eq. 9).

[14] Strictly speaking the assumption is not justified, however, because in the case of foreign trade prices under consideration here price indices are based on a sample survey whereas unit value indices are resulting from a comprehensive customs statistics. This inaccuracy may be acceptable because our focus is on the formal aspects of the differences between the two types of indices. In addition to the coverage there are many more conceptual and methodological differences between UVIs and PIs for example in German foreign trade statistics. In the annex we try to give briefly an account of the differences and empirical findings as regards their consequences.

(9) $$\sum_k \sum_j p_{kj0} q_{kj0} \frac{q_{kj1}}{q_{kj0}} = \sum_k \sum_j p_{kj0} q_{kj0} \frac{Q_{k1}}{Q_{k0}}$$

a condition given if for all j and k holds $\frac{q_{kj1}}{q_{kj0}} = \frac{Q_{k1}}{Q_{k0}} = \tilde{Q}^k_{01}$, as stated in eq. 6a.

It is easy to verify that the corresponding condition for $PU^P_{01} = P^L_{01}$ which is interesting from the point of view of German price statistics[15] is less straightforward. However, we find for the condition $PU^L_{01} = P^L_{01}$ a condition similar to (9) using *reciprocal* quantity changes

(9a) $$\sum_k \sum_j p_{kj0} q_{kj1} \frac{q_{kj0}}{q_{kj10}} = \sum_k \sum_j p_{kj0} q_{kj1} \frac{Q_{k0}}{Q_{k1}}.$$

It should be kept in mind, that both types of indices using unit values, the Drobisch index P^D as well as the PU indices (such as PU^P and PU^L) have much less satisfactory axiomatic properties than the price indices introduced here (P^P and P^L).[16] All these indices violate proportionality (and thus also identity by implication), commensurability and the mean value property (regarding the price relatives). As shown in eq. 6a PU^P is a mean of price relatives only in the unbiased case (that is when $PU^P = P^P$). In contrast to "true" price indices P^L and P^P all indices based on unit values reflect changes in the structure of quantities within - and in the case of PD also between - CNs in addition to changes in prices.[17] In other words, they violate the principle of pure price comparison.[18]

3. Bias of the unit value index of Drobisch

For convenience of presentation we introduce the following notation to denote a covariance related to all $n = \Sigma_k n_k$ commodities (summation takes place over k and j)

(10) $$\mathrm{Cov}(x, y, w) = \sum\sum (x_{kjt} - \bar{x})(y_{kjt} - \bar{y}) w_{kj} = \sum\sum x_{kjt} y_{kjt} w_{kj} - \bar{x} \cdot \bar{y}$$

where x and y are variables with arithmetic means $\bar{x} = \sum\sum x_{kjt} w_{kj}$ and $\bar{y} = \sum\sum y_{kjt} w_{kj}$ and w are weights $\sum\sum w_{kj} = 1$. By the well known "shift theorem" of eq. 10 the covariance around the point (centre of gravity) $\bar{x}, \bar{y}$ is shifted to the point 0, 0 (origin) resulting in the covariance $\sum\sum x_{kjt} y_{kjt} w_{kj}$ around the origin. In the case of a covariance for the n_k elements of the k-th CN only we introduce the symbol

(10a) $$\mathrm{cov}_k(x, y, w^*) = \sum_{j=1}^{n_k} (x_{kjt} - \bar{x}_k)(y_{kjt} - \bar{y}_k) w^*_{kj} \quad \text{where } \sum_j w^*_{kj} = 1$$

In Diewert (2010) we find the following three bias formulas (bias of P^D in relation to P^P)[19]

---

[15] As mentioned above, we there have a price index of Laspeyres and a "unit value index" of Paasche.

[16] Unlike Diewert we do not consider Fisher's "ideal index" $(P^L P^P)^{1/2}$ here and in sec. 3.

[17] They therefore can indicate a change in prices although each price remained constant, and the index also does not necessarily represent an "average" price change (beteen the lowests and the highest price change).

[18] This has already been established by the SNA 1993 which states that unit values are "affected by changes in the mix of items as well as by changes in their prices. Unit value indices cannot therefore be expected to provide good measures of average price change over time" (§ 16.13). Interestingly the SNA did not seem to realize that the same argument (no pure price comparison) would apply also to chain indices.

[19] In what follows we only quote the results without showing how they were derived. The equations 11, 12, 13 and 14 correspond directly to the equations 20, 22, 25 and 29 in Diewert (2010).

(11) $$\frac{P_{01}^{D}}{P_{01}^{P}} - 1 = \text{bias}(P^{D}, P^{P}) = \frac{n}{\tilde{p}_0} \cdot \text{Cov}\left(p_{kj0}, s_{kj1} - s_{kj0}, 1/n\right) \text{ where}$$

(11a) $$s_{kjt} = \frac{q_{kjt}}{\sum_k \sum_j q_{kjt}} = \frac{q_{kjt}}{Q_t}$$

are quantity shares relative to the *total* quantity (so that $\sum\sum s_{kjt}/n = 1$), and the "unweighted" (weights uniformly 1/n) covariance is given by

(11b) $$\text{Cov}(p_{kj0}, s_{kj1} - s_{kj0}, \frac{1}{n}) = \sum_k \sum_j \left(p_{kj0} - \bar{p}_0\right)\left(\{s_{kj1} - s_{kj0}\} - 0\right)\frac{1}{n} \text{ and } \bar{p}_0 = \sum\sum p_{kj0}/n .$$

A second bias equation reads as follows

(12) $$\frac{P_{01}^{D}}{P_{01}^{P}} - 1 = \frac{\text{Cov}\left(p_{kj0}, G_{kj}, s_{kj0}\right)}{\tilde{p}_0}$$ using quantity shares s_{kj0} as weights and

(12a) $$G_{kj} = s_{kj1}/s_{kj0} - 1 \text{ because } \sum\sum s_{kj0}\frac{s_{kj1}}{s_{kj0}} = 1 \text{ we get } \sum\sum s_{kj0} G_{kj0} = 0 .$$

Furthermore a sort of "average" base period price is now $\sum\sum s_{kj0} p_{kj0} = \tilde{p}_0$ instead of $\bar{p}_0$ defined in (11b) and the relevant covariance is

(12b) $$\text{Cov}(p_{kj0}, G_{kj}, s_{kj0}) = \sum_k \sum_j \left(p_{kj0} - \tilde{p}_0\right)\left(G_{kj} - 0\right) \cdot s_{kj0} .$$

Finally the bias can also be described by (third bias equation)

(13) $$\frac{P_{01}^{D}}{P_{01}^{P}} - 1 = \frac{\text{Cov}\left(p_{kj0}, q_{kj1}/q_{kj0}, s_{kj0}\right)}{\tilde{p}_0 \tilde{Q}_{01}}$$ using

(13a) $$\sum\sum s_{kj0}\frac{q_{kj1}}{q_{kj0}} = \tilde{Q}_{01}$$ so that we get

(13b) $$\text{Cov}(p_{kj0}, q_{kj1}/q_{kj0}, s_{kj0}) = \sum_k \sum_j \left(p_{kj0} - \tilde{p}_0\right)\left(q_{kj1}/q_{kj0} - \tilde{Q}_{01}\right) \cdot s_{kj0}$$

for the relevant covariance. This third expression seems to be most interesting for two reasons:

1. It closely corresponds to our result in section 4 and
2. It is worth noting that a case in which the covariance (13b) is vanishing is given when all changes of quantities q_{kj1}/q_{kj0} are equal and therefore equal to $\tilde{Q}_{01}$, and this is precisely the situation of eq. 6a and 9 where $PU_{01}^{P} = P_{01}^{P}$ in which case (that is $q_{kj1}/q_{kj0} = \tilde{Q}_{01}$) PU_{01}^{P} is also equal to P_{01}^{D}.

It is also interesting to note that the three results are equivalent. It can, for example, easily be seen that

(13c) $$\text{Cov}(p_{kj0}, q_{kj1}/q_{kj0}, s_{kj0}) = \sum\sum p_{kj0}\frac{q_{kj1}}{q_{kj0}} s_{kj0} - \tilde{p}_0 \tilde{Q}_{01}$$

$$= \tilde{Q}_{01} \sum\sum p_{kj0} G_{kj} s_{kj0} \text{ using } \frac{q_{kj1}}{q_{kj0}} = (G_{kj} + 1)\tilde{Q}_{01} \text{, and}$$

(12c) $\mathrm{Cov}(p_{kj0}, G_{kj}, s_{kj0}) = \sum\sum (p_{kj0} - \tilde{p}_0)(G_{kj} - 0) \cdot s_{kj0}$

$= \sum\sum p_{kj0} G_{kj} s_{kj0}$ since $\sum\sum s_{kj0} G_{kj0} = 0$, and we then get

$$\mathrm{bias}(P^D, P^P) = \frac{P_{01}^D}{P_{01}^P} - 1 = \frac{\mathrm{Cov}(p_{kj0}, G_{kj}, s_{kj0})}{\tilde{p}_0} = \frac{\mathrm{Cov}(p_{kj0}, q_{kj1}/q_{kj0}, s_{kj0})}{\tilde{p}_0 \tilde{Q}_{01}} .$$

To sum up: the bias of P^D relative to P^P will be zero if

1. the covariance (11b), (12b) or (13b) vanishes, or if one or both of the following two special conditions for this situation occurs, viz.
2. all base period prices are equal, so that $p_{kj0} = \bar{p}_0 = \tilde{p}_0 \;\; \forall k, j$, or
3. quantity shares do not change ($s_{kj1} = s_{kj0}$ in (11b) or $s_{kj1}/s_{kj0} = 1$ in (12b) or as in (13b) *absolute* quantities within a CN change at the same rate $q_{kj1}/q_{kj0} = \tilde{Q}_{01}^k$ for all k and j.

A discussion of the direction and amount of bias in terms of covariances will be deferred to section 4, because we will there encounter basically the same covariances.

Diewert also derived three equations for the bias of P^D relative to P^L (corresponding to equations 11, 12 and 13). We only quote

(14) $\frac{P_{01}^L}{P_{01}^D} - 1 = \frac{\mathrm{Cov}(p_{kj1}, q_{kj0}/q_{kj1}, s_{kj1})}{\tilde{p}_1 (\tilde{Q}_{01})^{-1}}$ where the covariance is given by

(14a) $\mathrm{Cov}(p_{kj1}, q_{kj0}/q_{kj1}, s_{kj1}) = \sum_k \sum_j (p_{kj1} - \tilde{p}_1)\left(q_{kj0}/q_{kj1} - (\tilde{Q}_{01})^{-1}\right) \cdot s_{kj1}$,

which is the equivalent to eq. 13. Note that

- the bias is now $P^L/P^D - 1$ (by contrast to P^D/P^P -1),
- the current period prices p_{kj1} take the part of base period prices p_{kj0}, and
- we now consider reciprocal quantity relatives q_{kj0}/q_{kj1} around their mean $1/\tilde{Q}_{01}$ instead of quantity relatives around $\tilde{Q}_{01}$.

The last mentioned aspect is relevant in the following way: the covariance 14a will vanish if $\frac{q_{kj0}}{q_{kj10}} = \frac{Q_{k0}}{Q_{k1}}$ holds, a condition also stipulated in (9a) for $PU_{01}^L = P_{01}^L$. It should be kept in mind that though $q_{kj0}/q_{kj1} = 1/\tilde{Q}_{01}^k$ does not differ from $q_{kj1}/q_{kj0} = \tilde{Q}_{01}^k$, this does not mean that (14a) amounts to the same scenario as in (13b).

4. Bias of the Paasche index of unit values

We now express biases in terms of covariances of the type introduced in (10a), that is referring to specific CNs in which quantity shares m_{kjt} are defined relative to the total quantity Q_{kt} of the CN (within-CN-shares m_{kjt} as opposed to total shares s_{kjt} defined in (11a))

(15) $m_{kjt} = q_{kjt}/Q_{kt}$.

The various quantity shares used in this paper are related as follows

(15a) $s_{kjt} = m_{kjt} \sigma_{kjt}$.

We call the ratio of PU^P defined in (4) and P^P defined in (7) "structural effect" S (or S-effect for short) because it is the changing structure of quantities which makes the difference between PU^P and P^P

(16) $$S = \frac{PU_{01}^P}{P_{01}^P} = \text{bias}(PU^P, P^P) + 1 .$$

Due to (8) and (8a) S can also be expressed in terms of quantity indices

(17) $$S = \frac{Q_{01}^L}{QU_{01}^L} = \sum_k \frac{Q_{01}^{L(k)}}{\tilde{Q}_{01}^k} \cdot \frac{\tilde{Q}_{01}^k s_{k0}}{\sum_k \tilde{Q}_{01}^k s_{k0}} = \sum_k S_{01}^k \cdot \frac{\tilde{p}_{k0} Q_{k1}}{\sum_k \tilde{p}_{k0} Q_{k1}}$$

where $Q_{01}^{L(k)}$ is the Laspeyres quantity index for the k-th CN so that S is a weighted arithmetic mean of "contributions" (of CNs) to the structural effect. Our aim now is to explain the S_{01}^k "contributions" with reference to covariances. It can easily be shown that

(18) $$S_{01}^k = \frac{Q_{01}^{L(k)}}{\tilde{Q}_{01}^k} = 1 + \frac{\tilde{p}_{k0}\left(Q_{01}^{L(k)} - \tilde{Q}_{01}^k\right)}{\tilde{p}_{k0}\tilde{Q}_{0t}^k}$$

and that the numerator $\tilde{p}_{k0}\left(Q_{01}^{L(k)} - \tilde{Q}_{01}^k\right)$ is the following covariance

(19) $$\text{cov}_k(p_{kj0}, q_{kj1}/q_{kj0}, m_{kj0}) = c_k = \sum_j \left(p_{kj0} - \tilde{p}_{k0}\right)\left(\frac{q_{kj1}}{q_{kj0}} - \tilde{Q}_{01}^k\right) \cdot m_{kj0}$$

$$= \frac{\sum_j q_{kj1} p_{kj0}}{\sum_j q_{kj0}} - \tilde{p}_{k0}\tilde{Q}_{01}^k = \tilde{p}_{k0}\left(Q_{01}^{L(k)} - \tilde{Q}_{01}^k\right).$$

The "contribution" S_{01}^k as ratio of two indicators of quantity change can also be written as

(20) $$S_{01}^k = \frac{Q_{01}^{L(k)}}{\tilde{Q}_{01}^k} = \frac{\sum_j p_{kj0} q_{kj1} \big/ \sum_j p_{kj0} q_{kj0}}{\sum_j q_{kj1} \big/ \sum_j q_{kj0}} = \frac{\tilde{p}_{k1}/\tilde{p}_{k0}}{P_{01}^{P(k)}} = \frac{\sum_j p_{kj0} q_{kj1}}{\sum_j q_{kj1}} : \tilde{p}_{kj0}$$

that is as ratio of indicators of price movement ($P_{01}^{P(k)}$ is the Paasche price index for the k-th CN), or as two different expressions for the average base period price in absolute terms, and

(19a) $$\text{bias}(Q^L, \tilde{Q}) = \frac{\tilde{p}_{k0}\left(Q_{01}^{L(k)} - \tilde{Q}_{01}^k\right)}{\tilde{p}_{k0}\tilde{Q}_{0t}^k} = \frac{\text{cov}_k\left(p_{kj0}, q_{kj1}/q_{kj1}, m_{kj0}\right)}{\tilde{p}_{k0}\tilde{Q}_{01}^k} = S_{01}^k - 1$$

may be viewed as a low-level (referring to a specific CN only) counterpart of (13).

It was only when I presented an earlier version of this paper at the 11th Meeting of the Ottawa Group in Neuchâtel 2009 that I became aware of the fact that G. Párniczky (1974) had already mentioned the covariance (19), or more precisely (13), because he studied the ratio P_{01}^D / P_{01}^P instead of PU_{01}^P / P_{01}^P. So he examined $\frac{\tilde{p}_1/\tilde{p}_0}{P_{01}^P}$ rather than $\frac{\tilde{p}_{k1}/\tilde{p}_{k0}}{P_{01}^{P(k)}}$,[20] because his focus was on P_{01}^D, while mine was – and still is – on PU_{01}^P.

---

[20] Interestingly Párniczky presented his result with explicit reference to a theorem of L. v. Bortkiewicz (however without detailed derivation of his formula using the theorem), which was also our point of departure in order to arrive at (19). In the following section this theorem will be considered in a more systematic manner.

A major difference between the equations 19 and 13 is that PU^P is a mean of K ratios $\tilde{p}_{k1}/\tilde{p}_{k0}$ and P^P a mean of "low level" Paasche indices

(21) $$P_{01}^{P} = \sum_k P_{01}^{P(k)} \frac{Q_{k1}\sum_j p_{kj0}m_{kj1}}{\sum_k Q_{k1}\sum_j p_{kj0}m_{kj1}},$$

however, as shown in (3a) it is not possible to express in a similar fashion the Drobisch index $P_{01}^{D} = \tilde{p}_1/\tilde{p}_0$ (and the value index V_{01}) as a mean of $\tilde{p}_{k1}/\tilde{p}_{k0}$ ratios.

It is now appropriate to give an interpretation to the result (19). This equation tells us, that

- a commodity j tends to raise (lower) $S = Q_{01}^{L}/QU_{01}^{L}$ as a weighted sum of S_{01}^{k} = $Q_{01}^{L(k)}/\tilde{Q}_{01}^{k}$ terms whenever the covariance c_k is positive (negative) and the commodity j has a non-negligible weight given by the share $m_{kj0} = q_{kj0}/\Sigma q_{kj0}$ of the total quantity at the base period, and
- the covariance c_k will be negative[21] so that S tends to be less than unity (in short: $c_k < 0 \rightarrow S_{01}^{k} < 1 \rightarrow S < 1$) if quantities of goods with above average prices ($p_{kj0} > \tilde{p}_{k0}$) in the base period tend to change below average ($q_{kjt}/q_{kj0} < \tilde{Q}_{0t}^{k}$) or if goods with below average base period prices. Correspondingly one may infer: $c_k > 0 \rightarrow S_{01}^{k} > 1 \rightarrow S > 1$.

It is important to note that the covariance c_k is not reflecting a substitution process in response to a change in the price structure because the prices in t = 1 are irrelevant and prices may even remain constant so that $p_{kj1} = p_{kj0}$. Note that $c_k \neq 0$ only necessitates a positive variance of base period prices. What matters is not the change of prices but only the structure of base period prices. In other words, it is only important whether it is a quantity change $q_{kjt}/q_{kj0} \neq \tilde{Q}_{0t}^{k}$ of a good j of which the base period price is above average or below average.[22]

The interesting case of a zero covariance c_k as defined in (19) - that is no contribution to the S-effect - takes place if *within* each CN

1. all base period prices are equal, so that $p_{kj0} = \bar{p}_{k0} = \tilde{p}_{k0} \quad \forall j$, and/or
2. for all j = 1, …, n_k holds $m_{kj1} = m_{kj0}$ (no structural change within a CN), or quantities change at the same rate λ so that $\lambda = q_{kj1}/q_{kj0} = Q_{01}^{L(k)} = \tilde{Q}_{01}^{k}$,[23]
3. the CN consists of only one commodity so that $n_k = 1$.

which is a set of conditions obviously closely related to conditions mentioned in our comments to eq. 11b, 12b and 13b.

As to the first statement: Equal prices in 0 entail equality of quantity shares ($m = q_0/\Sigma q_0$) and expenditure shares ($p_0q_0/\Sigma\, p_0q_0$), or equivalently $Q_{01}^{L(k)} = \tilde{Q}_{01}^{k}$, because then

[21] The empirical studies I performed in cooperation with Jens Mehrhoff (see the annex) clearly demonstrated that a negative S-effect (S < 1) is much more likely than a positive S-effect (S > 1). It also revealed that the most frequently observed case is $PU^P < P^P < P^L$ (or equivalently $Q^P < Q^L < QU^L$).

[22] After all it seems therefore difficult to "explain" the sort of economic behaviour which gives rise to a negative and a positive covariance c_k in terms of utility maximizing behaviour similar to the well known microeconomic theoretical underpinning of the (negative) substitution-effect responsible for $P^P < P^L$. The S-effect implies that a change in quantities may even take place although all prices remain constant.

[23] or do not change (λ=1).

(22) $\tilde{Q}_{01}^{k} = \sum_j \frac{q_{kj1}}{q_{kj0}} \frac{q_{kj0}}{\sum_j q_{kj0}} = \sum_j \frac{q_{kj1}}{q_{kj0}} \cdot m_{kj0}$ coincides with

(22a) $Q_{01}^{L(k)} = \sum_j \frac{q_{kj1}}{q_{kj0}} \frac{q_{kj0} p_{kj0}}{\sum_j q_{kj0} p_{kj0}}$.

The second statement follows from $S = PU_{01}^{P} / P_{01}^{P}$ and comparing (21) to

(23) $PU_{01}^{P} = \sum_k P_{01}^{P(k)} \frac{Q_{k1} \sum_j p_{kj0} m_{kj1}}{\sum_k Q_{k1} \sum_j p_{kj0} m_{kj0}}$ or considering

(23b) $S = \frac{Q_{01}^{L}}{QU_{01}^{L}} = \frac{\sum_k Q_{k1} \sum_j m_{kj1} p_{kj0}}{\sum_k Q_{k1} \sum_j m_{kj0} p_{kj0}}$

shows that assuming $m_{kj1} = m_{kj0}$ for all j and k gives $P_{0t}^{P} = PU_{0t}^{P}$, or $Q_{0t}^{L} = QU_{0t}^{L}$ (thus S = 1). Note that $Q_{01}^{L(k)} = \tilde{Q}_{01}^{k}$ implies $\frac{V_{01}^{k}}{\tilde{Q}_{01}^{k}} = \frac{\tilde{p}_{k1}}{\tilde{p}_{k0}} = \frac{V_{01}^{k}}{Q_{01}^{L(k)}} = P_{01}^{P(k)}$.

This once more confirms that the (contribution to the) S-effect may described either in terms of two measures of quantity movement $Q_{01}^{L(k)}$ and $\tilde{Q}_{01}^{k}$ or alternatively of two measures of price change, $P_{01}^{P(k)}$ and $\tilde{p}_{k1} / \tilde{p}_{k0}$ respectively.

Note that absence of the S-effect implies $P_{0t}^{P} = PU_{0t}^{P}$ but this does not mean that the Drobisch index $PU_{01}^{D} = \sum_k \tilde{p}_{k1} \sigma_{k1} / \sum_k \tilde{p}_{k0} \sigma_{k0}$ is also equal to $PU_{01}^{P} = \sum_k \tilde{p}_{k1} \sigma_{k1} / \sum_k \tilde{p}_{k0} \sigma_{k1} = P_{01}^{P}$ unless for all k holds $\sigma_{k1} = \sigma_{k0}$. "No S-effect" is therefore to be kept distinct from "no bias of the Drobisch index", because PU^P (by contrast to P^P) is reflective of changes within CNs and P^D will also be affected by changes between the CNs.

Now to the third statement: The difference between PU^P and P^P will also diminish to the extent that CNs are formed in a way that they are becoming more "homogeneous". Obviously we get $c_k = 0$ also with n_k identical observations (with respect to the variables p_{kj0}, q_{kj1}/q_{kj0}), or equivalently with $n_k = 1$ (and therefore $n = \Sigma n_k = K$).[24]. Interestingly we then again have $PU^P = P^P$ and a Drobisch index P^D which will still be in general different from PU^P and P^P.[25]

This again shows that the Drobisch index P^D

- is not just an aggregation of the low-level unit value ratios $\tilde{p}_{k1} / \tilde{p}_{k0}$ (this applies to PU^P and PU^L but not to P^D), and that
- absence of the S-effect (that is unbiased PU indices) is not tantamount to an unbiased Drobisch index, and that finally
- axiomatic properties are quite different in the case of P^D as opposed to PU^P and PU^L.

An interesting consequence of this finding (see statement 1 above) is the following advice "for choosing how to construct the subaggregates: in order to minimize bias (relative to the Paasche price index), use unit value aggregation over products that sell for the same price in the base period" (Diewert 2010, p. 14). One might add (alluding to the other two conditions):

---

[24] In this situation $p_{kjt} = p_{kt}$ and $q_{kjt} = q_{kt}$ (t = 0, 1). We come close to this situation if a CN is clearly dominated by one specific good, and the other (similar) goods are more or less insignificant.

[25] It can easily be seen since $PU^P = P^P = \Sigma p_{k1} q_{k1} / \Sigma p_{k0} q_{k1}$ but $P^D = (\Sigma p_{k1} q_{k1} / \Sigma q_{k1}) / (\Sigma p_{k0} q_{k0} / \Sigma q_{k0})$.

where changes in the structure of quantities transacted appear less likely - or where products are "homogeneous" in the sense of similar quantity changes – or the CN is dominated by one specific good only.

In contrast to section 3 we do not consider here the bias of PU^L with respect to P^L, or – what is much more complicated – of PU^P with respect to P^L. Instead we mention briefly only another covariance equation which may be useful for understanding the difference between PU_{01}^P and P_{01}^P. As an alternative to (19) one my explain the contribution $1/S_{01}^k$ to the (inverse) structural effect 1/S instead of the contribution S_{01}^k to S. The relevant covariance is given by

$$\text{(24)} \qquad \operatorname{cov}_k\left(q_{kj1}/q_{kj0}, 1/p_{kj0}, w_{kj}\right) = c_k^* = \sum\left(\frac{q_{kj1}}{q_{kj0}} - Q_{01}^{L(k)}\right)\left(\frac{1}{p_{kj0}} - \frac{1}{\tilde{p}_{k0}}\right)\frac{p_{kj0}q_{kj0}}{\sum p_{kj0}q_{kj0}}$$

$$= \frac{\sum_j q_{kjt}}{\sum_j p_{kj0}q_{kj0}} - Q_{01}^{L(k)} \cdot \frac{1}{\tilde{p}_{k0}} = \left(\tilde{p}_{k0}\right)^{-1}\left(\tilde{Q}_{01}^k - Q_{01}^{L(k)}\right)$$

using expenditure weights $w_{kj0} = p_{kj0}q_{kj0}/\Sigma\, p_{kj0}q_{kj0}$ rather than $m_{kj0} = q_{kj0}/\Sigma q_{kj0}$, and c_k^* is related to $1/S_{01}^k$ as follows $\frac{1}{S_{01}^k} = 1 + \frac{\left(\tilde{p}_{k0}\right)^{-1}\left(\tilde{Q}_{01}^k - Q_{01}^{L(k)}\right)}{\left(\tilde{p}_{k0}\right)^{-1} Q_{01}^{L(k)}} = \frac{\tilde{Q}_{01}^k}{Q_{01}^{L(k)}}$, so that c_k^* explains $1/S_{01}^k$ rather than S_{01}^k. The inverse structural effect S^{-1} is a weighted sum of these $1/S_{01}^k$ terms with weights $Q_{0t}^{L(k)}s_{k0}/\sum_k Q_{0t}^{L(k)}s_{k0}$ (instead of the weights $\tilde{Q}_{01}^k s_{k0}/\sum_k \tilde{Q}_{01}^k s_{k0}$ in (17)) and

$$\text{(24a)} \qquad \left(\tilde{p}_{k0}\right)^2 c_k^* = -c_k\,.$$

shows that the covariances are closely related to one another.[26]

5. Bias formulas and a theorem of L. v. Bortkiewicz

It can be shown that all covariance equations introduced in the sections 3 and 4 are simply special cases of a theorem first derived by L. von Bortkiewicz (1923) in which two *linear* indices (ratios of scalar products) X_0 and X_1 are compared.[27] Given

$$\text{(25)} \qquad X_1 = \frac{\mathbf{x}_1'\mathbf{y}_1}{\mathbf{x}_0'\mathbf{y}_1} = \frac{\sum x_1 y_1}{\sum x_0 y_1} \quad \text{and} \qquad \text{(25a)} \quad X_0 = \overline{X} = \frac{\mathbf{x}_1'\mathbf{y}_0}{\mathbf{x}_0'\mathbf{y}_0} = \frac{\sum x_1 y_0}{\sum x_0 y_0}$$

the theorem states that the ratio of the two indices is given by

$$\text{(26)} \qquad \frac{X_1}{X_0} = 1 + \frac{c_{xy}}{\overline{X} \cdot \overline{Y}} \quad \text{with the covariance}$$

$$\text{(27)} \qquad c_{xy} = \sum\left(\frac{x_1}{x_0} - \overline{X}\right)\left(\frac{y_1}{y_0} - \overline{Y}\right) w_0 = \frac{\sum x_1 y_1}{\sum x_0 y_0} - \overline{X} \cdot \overline{Y} \quad \text{and weights } w_0 = x_0 y_0 / \sum x_0 y_0\,.$$

Using (25) through (27) and two indices derived from interchanging x and y, viz.

---

[26] The covariances necessarily have different signs. The covariance c_k relates to S rather than S^{-1}, however, on the other hand c_k^* can more readily be compared to the covariance C of eq. 30.

[27] This is a *generalized* theorem of Bortkiewicz for the ratio X_1/X_0 of two linear indices. See von der Lippe (2007), pp. 194 – 196. The best known special case of this theorem is $X_0 = P^L$ and $X_1 = P^P$ of eq. 30.

(28) $\overline{Y} = \sum \frac{y_1}{y_0} \cdot w_0 = \frac{\sum y_1 x_0}{\sum y_0 x_0} = \frac{\mathbf{x}_0^{'}\mathbf{y}_1}{\mathbf{x}_0^{'}\mathbf{y}_0} = Y_0$ and (28a) $Y_1 = \frac{\mathbf{x}_1^{'}\mathbf{y}_1}{\mathbf{x}_1^{'}\mathbf{y}_0} = \frac{\sum x_1 y_1}{\sum x_1 y_0}$

it follows

(26a) $\frac{X_1}{X_0} = \frac{Y_1}{Y_0} = \frac{\mathbf{x}_1^{'}\mathbf{y}_1}{\mathbf{x}_0^{'}\mathbf{y}_1} \cdot \frac{\mathbf{x}_0^{'}\mathbf{y}_0}{\mathbf{x}_1^{'}\mathbf{y}_0} = 1 + \frac{c_{xy}}{\overline{X} \cdot \overline{Y}}$ where $c_{xy} = \mathbf{x}_1^{'}\mathbf{y}_1 / \mathbf{x}_0^{'}\mathbf{y}_0 - X_0 Y_0$.

Note also that it turns out that the product-moment $\mathbf{x}_1^{'}\mathbf{y}_1 / \mathbf{x}_0^{'}\mathbf{y}_0$ around the origin is

(29) $\gamma_{xy} = \mathbf{x}_1^{'}\mathbf{y}_1 / \mathbf{x}_0^{'}\mathbf{y}_0 = X_0 Y_1 = Y_0 X_1$ by contrast to

(27a) $c_{xy} = \gamma_{xy} - \overline{XY} = X_0(Y_1 - Y_0) = Y_0(X_1 - X_0)$.

The best known special case of this theorem is of course[28]

$X_1/X_0 = P_{01}^P / P_{01}^L$ and $Y_1/Y_0 = Q_{01}^P / Q_{01}^L$ (or vice versa) leading to the covariance

(30) $C = \sum_i \left(\frac{p_{i1}}{p_{i0}} - P_{01}^L \right) \left(\frac{q_{i1}}{q_{i0}} - Q_{01}^L \right) \frac{p_{i0} q_{i0}}{\sum p_{i0} q_{i0}} = V_{01} - Q_{01}^L P_{01}^L = Q_{01}^L P_{01}^P - Q_{01}^L P_{01}^L$ so that

(30a) $C = Q_{01}^L (P_{01}^P - P_{01}^L) = P_{01}^L (Q_{01}^P - Q_{01}^L)$ and $\gamma_{xy} = V_{01}$ the value index

showing how Paasche and Laspeyres indices are interrelated (we usually expect $C < 0$ and therefore $P_{01}^P < P_{01}^L$).[29]

It can easily be seen that the eqs. 19, 24, 30 and 31 are simply special cases of the Bortkiewicz-theorem.[30] This is shown in table 1. In particular for the covariance of (19) it is also shown what happens when the vectors **x** and **y** are interchanged.

Note, however, that the theorem does not allow comparing *any* two indices. It is for example possible to compare $X_1 = PU_{01}^P = \frac{\sum_k \tilde{p}_{k1} \sigma_{k1}}{\sum_k \tilde{p}_{k0} \sigma_{k1}}$ to $X_0 = PU_{01}^L = \frac{\sum_k \tilde{p}_{k1} \sigma_{k0}}{\sum_k \tilde{p}_{k0} \sigma_{k0}}$ resulting in the covariance

(31) $\tilde{C} = \sum_k \left(\frac{\tilde{p}_{k1}}{\tilde{p}_{k0}} - PU_{01}^L \right) \left(\frac{Q_{k1}}{Q_{k0}} - QU_{01}^L \right) \frac{\tilde{p}_{k0} Q_{k0}}{\sum \tilde{p}_{k0} Q_{k0}}$

(31a) $= QU_{01}^L (PU_{01}^P - PU_{01}^L) = PU_{01}^L (QU_{01}^P - QU_{01}^L)$

in perfect analogy to (30) and (30a). In the first place it does not appear straightforward to compare $X_1 = PU_{01}^P$ to

(32) $X_0^* = P_{01}^D = \frac{\sum_k \tilde{p}_{k1} \sigma_{k1}}{\sum_k \tilde{p}_{k0} \sigma_{k0}}$.

28 We assume here that the index will be compiled in one stage only so that i = 1, …,n.

29 A negative covariance ($P^P < P^L$) may arise from rational substitution among goods in response to price changes on a given (negatively sloped) demand curve. The less frequent case of a positive covariance is supposed to take place when the demand curve is shifting away from the origin (due to an increase of income for example).

30 This result is anything but surprising because I derived the equations for c_k (eq. 19) and c_k^* (eq. 24) by explicitly using this theorem.

Table 1: Covariance equations and the theorem of L. v. Bortkiewicz

eq	$\mathbf{x}_0$	$\mathbf{x}_1$	$\mathbf{y}_0$	$\mathbf{y}_1$	$X_0 = \overline{X}$	X_1	$Y_0 = \overline{Y}$	Y_1	weights
19	1	p_0	q_0	q_1	$\tilde{p}_{k0}$	$\Sigma p_0 q_1 / \Sigma q_1$ a)	$Q_{01}^{L(k)}$	$Q_{01}^{L(k)}$	$q_0/\Sigma q_0$
19	q_0	q_1	1	p_0	$\tilde{Q}_{01}^{k}$	$Q_{01}^{L(k)}$	$\tilde{p}_{k0}$	$\Sigma p_0 q_1 / \Sigma q_1$	$q_0/\Sigma q_0$
24	q_0	q_1	p_0	1	$Q_{01}^{L(k)}$	$\tilde{Q}_{01}^{k}$	$1/\tilde{p}_{k0}$	$\Sigma q_1 / \Sigma p_0 q_1$	$p_0 q_0 / \Sigma p_0 q_0$
30	p_{kj0}	p_{kj1}	q_{kj0}	q_{kj1}	P_{01}^{L}	P_{01}^{P}	Q_{01}^{L}	Q_{01}^{P}	$p_0 q_0 / \Sigma p_0 q_0$
31	$\tilde{p}_{k0}$	$\tilde{p}_{k1}$	σ_{k0}	σ_{k1}	PU_{01}^{L}	PU_{01}^{P}	QU_{01}^{L}	QU_{01}^{P}	b)
34	Q_{k0}	Q_{k1}	1	$\tilde{p}_{k0}$	$\tilde{Q}_{01}$	QU_{01}^{L}	$\tilde{p}_0$	$\tilde{p}_0^{*}$ c)	$Q_{k0}/\Sigma Q_{k0}$

Diewert's equations

eq	$\mathbf{x}_0$	$\mathbf{x}_1$	$\mathbf{y}_0$	$\mathbf{y}_1$	$X_0 = \overline{X}$	X_1	$Y_0 = \overline{Y}$	Y_1	weights
13	1	p_{kj0}	q_{kj0}	q_{kj1}	$\tilde{p}_0$	$\tilde{p}_0^{**}$ d)	$\tilde{Q}_{01}$	Q_{01}^{L}	s_{kj0} e)
12	1	p_{kj0}	s_{kj0}	s_{kj1}	$\tilde{p}_0$	$\tilde{p}_0^{**}$	1	$Q_{01}^{L}/\tilde{Q}_{01}$	s_{kj0}
11	1	p_{kj0}	1	Δ_{kj} f)	$\bar{p}_0$	g)	$0 = \Sigma\Delta_{kj}/n$	h)	$1/n$
14	1	p_{kj1}	q_{kj1}	q_{kj0}	$\tilde{p}_1$	$\tilde{p}_1^{**}$ i)	$1/\tilde{Q}_{01}$	$1/Q_{01}^{P}$	s_{kj1} j)

a) Note that according to (20) the term $\sum p_0 q_1 / \sum q_1$ is equal to $S_{01}^{k} \tilde{p}_{k1}$

b) weights are $\tilde{p}_{k0}\tilde{q}_{k0} / \sum_k \tilde{p}_{k0}\tilde{q}_{k0} = \sum_j p_{kj0} q_{kj0} / \sum_k \sum_j p_{kj0} q_{kj0}$

c) $\tilde{p}_0^{*}$ as defined in (34b)

d) $\tilde{p}_0^{**} = \sum\sum p_{kj0} q_{kj1} / \sum\sum q_{kj1} = \left(Q_{01}^{L} / \tilde{Q}_{01}\right) \cdot \tilde{p}_0 = S^* \tilde{p}_0$

e) $s_{kj0} = q_0/\Sigma\Sigma q_0$

f) $\Delta_{kj} = s_{kj1} - s_{kj0}$

g) not defined (division by $\Sigma\Delta_{kj1}$)

h) $Y_1 = \sum\sum p_{kj0} \Delta_{kj0} / \sum\sum p_{kj0}$

i) $\tilde{p}_1^{**} = \sum\sum p_{kj1} q_{kj0} / \sum\sum q_{kj0} = \left(\tilde{Q}_{01} / Q_{01}^{P}\right) \cdot \tilde{p}_1$

j) $s_{kj1} = q_1/\Sigma\Sigma q_1$

However the ratio $\frac{X_1}{X_0^*} = \frac{PU_{01}^{P}}{P_{01}^{D}} = \frac{\sum_k \tilde{p}_{k1}\sigma_{k1}}{\sum_k \tilde{p}_{k0}\sigma_{k1}} \cdot \frac{\sum_k \tilde{p}_{k0}\sigma_{k0}}{\sum_k \tilde{p}_{k1}\sigma_{k1}} = \frac{\sum_k \tilde{p}_{k0}\sigma_{k0}}{\sum_k \tilde{p}_{k0}\sigma_{k1}}$ is a meaningful expression. Although the rightmost term is not simply $\left(QU_{01}^{L}\right)^{-1} = \sum_k \tilde{p}_{k0} Q_{k0} / \sum_k \tilde{p}_{k0} Q_{k1}$ but rather

(33) $$\frac{PU_{01}^{P}}{P_{01}^{D}} = \frac{\sum_k \tilde{p}_{k0}\sigma_{k0}}{\sum_k \tilde{p}_{k0}\sigma_{k1}} = \frac{\tilde{Q}_{01}}{QU_{01}^{L}},$$

numerator and denominator of this ratio can be viewed as linear indices if we assume for the four vectors: $x_0 = Q_{k0}$, $x_1 = Q_{kt}$, $y_0 = 1$ and $y_1 = \tilde{p}_{k0}$. Then the relevant covariance amounts to

(34) $$\tilde{C}^* = \mathrm{Cov}\left(\tilde{Q}_{01}^{k}, \tilde{p}_{k0}, \sigma_{k0}\right) = \sum_k \left(\frac{Q_{k1}}{Q_{k0}} - QU_{01}^{L}\right)\left(\tilde{p}_{k0} - \tilde{p}_0\right)\frac{Q_{k0}}{\sum Q_{k0}},$$ or equivalently

(34a) $\tilde{C}^* = \tilde{Q}_{01}\left(\tilde{p}_0^* - \tilde{p}_0\right) = \tilde{p}_0\left(QU_{01}^L - \tilde{Q}_{01}\right)$ where

(34b) $\tilde{p}_0^* = \sum \tilde{p}_{k0} Q_{k1} / \sum Q_{k1}$ as opposed to $\tilde{p}_0 = \sum \tilde{p}_{k0} Q_{k0} / \sum Q_{k0}$

It is interesting to compare this "between CNs" covariance with the "within CNs" covariance of (19):

(19) $c_k = \text{cov}_k(p_{kj0}, q_{kj1}/q_{kj0}, m_{kj0}) = \tilde{p}_{k0}\left(Q_{01}^{L(k)} - \tilde{Q}_{01}^k\right)$ and

(34) $\tilde{C}^* = \text{Cov}\left(\tilde{p}_{k0}, Q_{k1}/Q_{k0}, \sigma_{k0}\right) = \tilde{p}_0\left(QU_{01}^L - \tilde{Q}_{01}\right)$

$\tilde{Q}_{01}$, $\tilde{p}_0$ and σ_{k0} are the between CNs counterparts of the terms $\tilde{Q}_{01}^k$, $\tilde{p}_{k0}$ and m_{kj0} referring to the k-th CN. Furthermore $Q_{01}^{L(k)}$ (22a) corresponds to $QU_{01}^L = \sum_k \frac{Q_{k1}}{Q_{k0}} \frac{Q_{k0}\tilde{p}_{k0}}{\sum_j Q_{k0}\tilde{p}_{k0}}$.

So there is a sequence of biases from P_{01}^P to PU_{01}^P (explained by c_k) and then from PU_{01}^P to P_{01}^D (explained by $\tilde{C}^*$). This once more reaffirms our conclusion that even if PU_{01}^P were unbiased relative to P_{01}^P (because $c_k = 0$ for all k), the Drobisch index can still be biased to the extent to which $\tilde{C}^*$ differs from zero.

To this point we only revealed some obvious implications of the theorem. However it seems to be less obvious that also Diewert's equations, as they are reported in section 3 can be shown to follow from the theorem. We first consider

(13b) $\text{Cov}(p_{kj0}, q_{kj1}/q_{kj0}, s_{kj0}) = \sum_k \sum_j \left(p_{kj0} - \tilde{p}_0\right)\left(q_{kj1}/q_{kj0} - \tilde{Q}_{01}\right) \cdot s_{kj0}$

$$= \frac{\sum\sum p_{kj0} q_{kj1}}{Q_0} - \tilde{Q}_{01}\tilde{p}_0 = Q_{01}^L \tilde{p}_0 - \tilde{Q}_{01}\tilde{p}_0 = \tilde{p}_0\left(Q_{01}^L - \tilde{Q}_{01}\right) \quad .$$

Upon substituting this into (13) it remains to be seen that $\frac{P_{01}^D}{P_{01}^P} = \frac{Q_{01}^L \tilde{p}_0}{\tilde{p}_0 \tilde{Q}_{01}} = \frac{Q_{01}^L}{\tilde{Q}_{01}} = S^*$, and this can indeed easily be verified. So the interesting result is, that Diewert's (13) turns out as

- not only another example for Bortkiewicz's theorem (if the vectors **x** and **y** and therefore the X and Y indices are specified according to table 1) but also as
- closely related to eq. 19 so that $\text{Cov}(p_{kj0}, q_{kj1}/q_{kj0}, s_{kj0})$ seems indeed to be the all-item analogue to k^{th}-CN covariance $c_k = \text{cov}_k(p_{kj0}, q_{kj1}/q_{kj0}, m_{kj0})$.

It has already been shown that (13) can be translated into (12). It may nonetheless be interesting to see which specification will lead to (12). We see in table 1 that we can arrive from (13) to (12) by making only a slight modification of the definitions the **y**-vectors. The consequences regarding Y_0 and Y_1 are quite plausible taking into account that $\frac{s_{kj1}}{s_{kj0}} \cdot \tilde{Q}_{01} = \frac{q_{kj1}}{q_{kj0}}$.

It remains to be seen that Diewert's unweighted covariance (11) and also his formulas for the bias of the Drobisch index relative to the Laspeyres index can again be derived from the theorem.[31] With the specification for eq. (11) in table 1 we have $\text{Cov}(p_{kj0}, s_{kj1} - s_{kj0}, \frac{1}{n}) = Y_1 \bar{p}_0$

[31] Of the three bias-formulas Diewert derived we only quoted one here, viz. (14).

so that with (11) $\frac{P^D_{01}}{P^P_{01}}-1=\frac{Q^L_{01}}{\tilde{Q}_{01}}-1=\frac{n}{\tilde{p}_0}\cdot Y_1\bar{p}_0$ we get $\tilde{p}_0\left(\frac{Q^L_{01}}{\tilde{Q}_{01}}-1\right)==nY_1\bar{p}_0$ and both sides of this equation are equal to $\sum\sum p_{kj0}q_{kj1}/Q_1-\tilde{p}_0$.

So again this result can be regarded as specialisation of Bortkiewicz's theorem.[32] Finally we examine (14) and find with the specification in table 1

$$\text{(14)}\qquad \frac{P^L_{01}}{P^D_{01}}-1=\frac{\operatorname{Cov}(p_{kj1},q_{kj0}/q_{kj1},s_{kj1})}{\tilde{p}_1\left(\tilde{Q}_{01}\right)^{-1}}=\frac{\tilde{p}_1\left(\frac{1}{Q^P_{01}}-\frac{1}{\tilde{Q}_{01}}\right)}{\tilde{p}_1\left(\tilde{Q}_{01}\right)^{-1}}=\tilde{Q}_{01}\left(\frac{1}{Q^P_{01}}-\frac{1}{\tilde{Q}_{01}}\right)$$

also complies with the theorem because using $P^L_{01}Q^P_{01}=V_{01}=P^D_{01}\tilde{Q}_{01}$ it is easy to see that (14), that is $\frac{P^L_{01}}{P^D_{01}}-1=\frac{\tilde{Q}_{01}}{Q^P_{01}}-1$ is true.

Table 1 also reveals that (14) is in a way a sort of "inversion" of (13).

6. Conclusions and some additional remarks

The analysis has shown that the indices P^P_{0t}, PU^P_{0t}, P^D_{0t} are ordered according to decreasingly favourable axiomatic properties and a decreasing compliance with "pure price comparison".[33] Wile P^P_{0t} is a mean of price relatives, this does not apply to PU^P_{0t} which is "only" a mean of ratios of unit values $\tilde{p}_{k1}/\tilde{p}_{k0}$ (and therefore affected from the structural effect S, that is changes in the structure of quantities *within* CNs).

However, $P^D_{01}=\frac{\tilde{p}_1}{\tilde{p}_0}=\frac{\sum_k\tilde{p}_{k1}\sigma_{k1}}{\sum_k\tilde{p}_{k0}\sigma_{k0}}$ as opposed to $PU^P_{01}=\frac{\sum_k\tilde{p}_{k1}\sigma_{k1}}{\sum_k\tilde{p}_{k0}\sigma_{k1}}$ or $PU^L_{01}=\frac{\sum_k\tilde{p}_{k1}\sigma_{k0}}{\sum_k\tilde{p}_{k0}\sigma_{k0}}$ is not even a mean of the ratios $\tilde{p}_{k1}/\tilde{p}_{k0}$ and affected by changes in the structure of quantities *between* CNs) *in addition to* the S-effect. Therefore even if the S-effect were absent, that is in the case of $S=PU^P_{0t}/P^P_{0t}=1$ the Drobisch index can still be biased.

We explained the bias of PU^P_{0t} relative to P^P_{0t} with the S effect where $S=Q^L_{0t}/QU^L_{0t}$ can be regarded as a weighted mean of K terms $S^k_{01}=Q^{L(k)}_{01}/\tilde{Q}^k_{01}$ which in turn depend on the K covariances c_k (defined in (19)) between $q_{kj1}/q_{kj0}-\tilde{Q}^k_{01}$ and $p_{kj0}-\tilde{p}_{k0}$. This result is in line with some formulas of Diewert and a theorem of v. Bortkiewicz.

However, as mentioned above it is difficult to think of a microeconomic theory able to explain sign and amount of the covariance c_k as this covariance relates changes in quantities from period 0 to 1 to the structure of base period prices only irrespective of a possible change of prices between periods 0 and 1 (thus the change in quantities cannot be viewed as a substitution in response to changing prices). It nonetheless may be a challenge to explain S with reference to utility maximizing behaviour.

[32] This appears a bit far-fetched, however, because in contrast to the other formulas it is not possible here to express the theorem analogously to (26a) due to the fact that $X_1/X_0 = Y_1/Y_0$ is not defined in this case.
[33] Of course the same applies mutatis mutandis to the sequence P^L, PU^L and P^D.

Initially our aim was to compare PU_{0t}^{P} to P_{0t}^{L} because the former corresponds formally to the German customs based "unit value index" while the latter is the formula for the survey based "price index". We realised that formulas appropriate to compare PU_{0t}^{P} to P_{0t}^{L} turn out much more complicated and involved than the formulas we discussed in section 4 to compare PU_{0t}^{P} to P_{0t}^{P}. So we could not achieve more than a comparison in two steps decomposing the discrepancy D into two factors (or effects"), the well known "S-effect" (S) and a substitution or Laspeyres effect (L) with the covariance C as defined in (30)

(35) $$D = \frac{PU_{01}^{P}}{P_{01}^{L}} = \left(\frac{C}{Q_{01}^{L}P_{01}^{L}} + 1\right)\left(\frac{Q_{01}^{L}}{QU_{01}^{L}}\right) = \frac{P_{01}^{P}}{P_{01}^{L}} \cdot \frac{PU_{01}^{P}}{P_{01}^{P}} = L \cdot S.$$

Just like S can be broken down to S_{01}^{k} terms on the level of CNs we may also break down L to the level of individual commodities i = 1, … , n

$$L = \sum_{i=1}^{n} L_i = \sum_{i=1}^{n} \left(\frac{p_{i1}/p_{i0}}{P_{01}^{L}}\right)\left(\frac{q_{i1}/q_{i0}}{Q_{01}^{L}}\right)\frac{p_{i0}q_{i0}}{\sum p_{i0}q_{i0}} = 1 + \frac{C}{P_{01}^{L}Q_{01}^{L}}$$

where $C/P_{01}^{L}Q_{01}^{L}$ is a sort of a "centred" covariance (divided by the respective means).[34]

Note that if S vanishes (for example because all prices of a CN in 0 are equal) $L = Q_{01}^{P}/Q_{01}^{L}$ does not vanish but only reduces to $L = Q_{01}^{P}/\tilde{Q}_{01}$ (since in this case $S = Q_{0t}^{L}/QU_{0t}^{L} = 1$ and therefore $Q_{01}^{L} = \tilde{Q}_{01}$).

An interesting difference between the two effects S and L is that we may want a price index to reflect the substitution between quantities in response to changing prices, that is the L-effect as typically enough all "superlative" indices do in contrast to the non- superlative Laspeyres index. This, however, does not apply to the S-effect, which rather seems to be an unwanted disturbance, a phenomenon a price index better should *not* reflect.

Moreover, it the S-effect is also undesirable because it may amount to a violation of identity in the following way: while prices must be changing for the L-effect to occur, the S-effect ($PU_{01}^{P} \neq P_{01}^{P}$) is possible even with constant prices (in which case $P_{01}^{P} = 1$), provided only that the structure of quantities is changing.

So there are reasons to study the two components or distinct "effects" separately and it may be interesting to see how the effects, L and S work in the same or in opposite direction[35] Yet it seems to be worthwhile to try to compare directly PU_{01}^{P} to P_{01}^{L} with preferably only one single determinant brought into play.

In addition to the formal aspects regarding the difference between PU^{P} and P^{L} on which this paper focuses, there are many other aspects that should be considered when an assessment of unit value indices has to be made. Although they are standard practice in many countries and unit values gain importance with the increased use of scanner data there are strong

[34] Unlike the L-effect the S effect only exists when commodities are grouped together in CNs and the structural effect owes its existence to the two-stage compilation of the PU-type indices. If summation would take in one stage over all individual commodities (not grouped into CNs) or (equivalently) if CNs were perfectly homogeneous the S-effect would disappear. There can be no S-effect without heterogeneity and/or structural change within the CNs. It appears therefore sensible to study the S-effect by examining the situation *within* the CNs, and this is precisely what is done in section 4.

[35] It is of course also possible that either or both effects vanish. As aforementioned we found that mostly both effects were negate so that L < 1 and S < 1 unequivocally produced D < 1 (that is $PU^{P} < P^{L}$).

reservations about PU-indices for the principal reason that they do not compare like with like. They are as a rule compiled without quality adjustments, outlier detection and deletion or provisions for temporarily non available goods.[36] We therefore agree with Silver (2007, 2008) that they may be justified – if at all – only as low-budget proxies for survey-based price indices. It is, however, most likely that due to budget constraints and the desire to minimize response burdens in surveys this type of price index will – notwithstanding its shortcomings – will gain ground in the future, and it is therefore of some significance to better understand the differences between PU-indices and "true" price indices, and this is precisely at which this paper aims.

Appendix

A1. Formulas of indices of export and import in Germany

Unit values $\tilde{p}_{kt}$ take the part of prices in both price- and quantity indices; hence we have unit value indices on the level of price and of quantity indices respectively (the latter is less common, however). So in theory at least $2^4 = 16$ indices exist due to the four dichotomies:

1. unit value index (UVI) vs. price index (PI) concept (level of aggregation in price data),
2. index describing movement of prices vs. quantities (volumes),
3. Laspeyres vs. Paasche formula and
4. prices of exports vs. those of imports.

German official statistics provides Paasche *unit value* indices in addition to genuine Laspeyres type *price* indices (both of export and import respectively). There are also countries in which use is made of both, prices and unit values in the same (price) index.[37]

Figure A.1: The structure of indices on the basis of unit values*

Indices on the basis of

unit values (UV)		prices	
UV price indices	UV quantity indices	true price indices	true quantity indices
summation over k = 1, ... ,K groups $\sum_k$		summation over n commodities $\sum_k \sum_j$	
formula of Laspeyres		formula of Laspeyres	
$PU_{01}^{L} = \frac{\sum \tilde{p}_1 Q_0}{\sum \tilde{p}_0 Q_0}$	$QU_{01}^{L} = \frac{\sum Q_1 \tilde{p}_0}{\sum Q_0 \tilde{p}_0}$	$P_{01}^{L} = \frac{\sum\sum p_1 q_0}{\sum\sum p_0 q_0}$	$Q_{0t}^{L} = \frac{\sum\sum q_1 p_0}{\sum\sum q_0 p_0}$
formula of Paasche		formula of Paasche	
$PU_{01}^{P} = \frac{\sum \tilde{p}_1 Q_1}{\sum \tilde{p}_0 Q_1}$	$QU_{01}^{P} = \frac{\sum Q_1 \tilde{p}_1}{\sum Q_0 \tilde{p}_1}$	$P_{01}^{P} = \frac{\sum\sum p_1 q_1}{\sum\sum p_0 q_1}$	$Q_{01}^{P} = \frac{\sum\sum q_1 p_1}{\sum\sum q_0 p_1}$

* The universe of n commodities is partitioned into K groups (sub-collections) of related commodities; the subscript k = 1, 2, ..., K denotes the number of the group and the subscript j the j^{th} commodity of the k^{th} group.

[36] The following appendix will present some more details regarding the deficiencies of unit value indices.

[37] As mentioned in footnote 11 according to the Internet the export/import price index (= International Merchandise Trade Price index IMTPI) of Canada seems to be an example for this in that it makes use of both unit values (on the basis of customs data) and when unit values are not accurate (heterogeneous aggregates) or unavailable price data provided by other sources. Moreover both direct index formulas, Laspeyres and Paasche, and both designs, direct as well as chained index formulas are being compiled.

A2. Data basis (survey based price indices vs. customs based unit value indices)

Unit value indices (UVIs) are based on a complete statistics of customs documents rather than on the observation of a sample of carefully specified goods under comparable conditions. Thus UVIs also refrain from using appropriate methods for adjustments of quality changes, temporary (seasonal) unavailability, or outlier detection and deletion. Moreover there are reasons to expect ever more difficulties in the future as regards customs statistics. We observe an increasing proportion of trade in services rather than in goods that physically cross borders. Likewise e-trade and intra-area trade within customs unions without customs documents on which statistics could be based gain importance. In sum unit value indices are less commendable from a theoretical point of view.

Table A.2: Indices of prices in foreign trade (export and import) in Germany

	Price index (PI)	Unit value index (UVI)
Data	Survey based (monthly), sample; more demanding than UVI (empirical weights!)	A by-product of customs statistics, census, in the case of Intrastat* survey
Formula	Laspeyres	Paasche
Prices, aggregates[38]	Prices of specific goods at time of contracting (lead of price index?)	Average value of CNs; time of crossing border (lag of UVI?)
New or dis-appearing goods	Included only with a new base period; vanishing goods replaced by *similar* ones constant selection of goods *	Immediately included; price quotation of disappearing goods is simply discontinued; variable universe of goods
Quality	Quality adjustment are performed	No quality adjustment (not feasible?)

* intra European Community (or Union)

** All price determining characteristics are deliberately kept constant

By contrast to compile a sample survey based PI is more demanding. It requires special surveys addressing exporting and importing establishments as well as compliance with the principle of "pure price comparison". This implies making adjustments (of reported prices) for quality changes in the traded goods or avoiding changes in the collection of goods, reporting firms or in the countries of origin (in the case of imports) or destination involved.

To sum up PIs appears to be theoretically more ambitious and to fit better to the general methodology (and the principle of pure price comparison in particular) of official price statistics whereas UVI might be a low budget "second best" solution and surrogate for PIs as they are more readily available and less demanding as regards data collection.

A3. Hypothesis on the basis of the conceptual differences between P and U indices

The conceptual and methodological differences mentioned give rise to testing empirically some hypotheses. In what follows we refer to an unpublished paper the present author has written in cooperation with Jens Mehrhoff (von der Lippe, Mehrhoff (2008)).[39] We studied altogether six hypotheses (see table A.3 summarizing the main results) using German data (Jan. 2000 through Dec. 2007). The hypotheses were quite obvious given the conceptual differences and most of them proved true. Above all UVIs and PIs of export and import respectively differ with regard to their level and volatility. UVIs tend to display a relative to PIs more moderate rise of prices combined with more accentuating oscillations. An altogether smoother pattern of the time series can also be attributed to the process of quality adjustment of PIs whereas UVIs are habitually not adjusted (which is in no small measure also due to the

[38] King (1993) in particular addressed this problem of a different point in time to which the price recordings refer.
[39] Compared to von der Lippe (2007b) it contains a completely new empirical study (worked out by J. Mehrhoff).

fact that details about the quality of the goods are lacking in customs data). Conspicuously and contrary to our expectations there was no clear evidence for the expected lead of PIs relative to the UVIs.

Table A.3: Summary of tests about differences between unit value indices (U = UVI) and price indices (P = PI) based on empirical calculations of Jens Mehrhoff

Hypothesis	Argument	Method	Result
1) **U < P**, **growing discrepancy**	Laspeyres (P) > Paasche (U) Formula of L. v. Bortkiewicz	Theil's inequality coeff. applied to growth rates of the series	largely confirmed
2) **Volatility** U > P	U no pure price comparison (U reflecting changes in product mix [structural changes])	Dispersion (RMSE) of de-trended (HP Filter) series (of P and U in exports and imports)	confirmed [a)]
3) **Seasonality** U > P	U no adjustment for seasonally non-availability	Standard dev. of seasonal component (Census X-2ARIMA)	similar to hypothesis no. 2
4) **U** suffers from **heterogeneity**	Variable vs. constant selection of goods, CN less homogeneous than specific goods	average correlation (root of mean R^2) of subindices (if small heterogeneity)	U only slightly more heterogeneous [b)]
5) **Lead** of P against U	Prices refer to the earlier moment of contracting (contract-delivery lag; exchange rates)	Correlation between ΔP (shifted forward) against ΔU	no systematic pattern [c)]
6) **Smoothing** in the case of P	Quality adjustment in P results in smoother time series	special data analysis [d)] of the German Stat Office	confirmed

a) Hypothesis largely confirmed, P is integrated, U stationary (depending on the level of (dis)aggregation)
b) more pronounced in the case of imports than of exports
c) in line with Silver's results
d) concerning desktops, notebooks, working storage and hard disks; coefficient of variation was in all cases sizeably smaller after quality adjustment than before.

References

Balk, B. M. (1994), On the First Step in the Calculation of a Consumer Price Index, in: Ducharme, L.M. (Ed.), *Proceedings of the First Meeting of the International Working Group on Price Indices*, Ottawa, Canada: Statistics Canada

Balk, B. M. (1998), On the Use of Unit Value Indices as Consumer Price Subindices, in: Lane, W. (Ed.) *Proceedings of the Fourth Meeting of the International Working Group on Price Indices*, Washington, DC: Bureau of Labour Statistics

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

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

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

Chaffé, A., Lequain, M. and O'Donnell, G., Assessing the Reliability of the CPI Basket Update in Canada Using the Bortkiewicz Decomposition, Statistics Canada, September 2007

Diewert, W. E. (1995), Axiomatic and Economic Approaches to Elementary Price Indexes, *NBER Working Paper* 5104

Diewert, W. E. (2004), Elementary Indices, in: ILO etc. (CPI-Manual)

Diewert, W. E. (2010), Notes on Unit Value Bias Index, Discussion Paper 10-08, Department of Economics, University of British Columbia, Vancouver, Canada, Sept. 22, 2020

Drobisch, M. W. (1871), Ueber die Berechnung der Veränderungen der Waarenpreise und des Geldwerths, Jahrbücher für Nationalökonomie und Statistik, Vol. 16, pp. 143 – 156

ILO, IMF, OECD, UNECE, Eurostat and World Bank (2004), Consumer Price Index Manual Theory and Practice (CPI-Manual), ed. by P. Hill, Geneva

IMF, ILO, OECD, UNECE and World Bank (2008), Export and Import Price Index Manual (XMPI-Manual), ed. by M. Silver, Washington

King, A. (1993), A Note on Export Unit Value Indices in Competitiveness Variables, in:Bulletin of Economic Research 45; 1

Párniczky, G. (1974), Some Problems of Price Measurement in External Trade Statistics, *Acta Oeconomica*, Vol. 12, No. 2, pp. 229 - 240

Silver, M. (2007), Do Unit Value Export, Import, and Terms of Trade Indices Represent or Misrepresent Price Indices? *IMF Working Paper* WP/07/121.

Silver, M. (2008), Chapter 2 "Unit Value Indices" in ILO etc. XMPI-Manual

Timmer, M. (1996), On the Reliability of Unit Value Ratios in International Comparisons, Research Memorandum GD-31, University of Groningen, Dec. 1996

von der Lippe, P. M. (2001), Chain Indices, A Study in Index Theory, Spectrum of Federal Statistics, Ed. Federal Statistical Office Germany, Stuttgart (Metzler-Poeschel)

von der Lippe, P. M. (2007a), Index Theory and Price Statistics, Frankfurt/Main (Peter Lang)

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

von der Lippe, P. M. and Mehrhoff, J. (2008), Unit Value Bias Reconsidered (unpublished paper)