Contingently biased, permanently biased and excellent index numbers for complete micro data

Yrjö Vartia
University of Helsinki, Finland (yrjo.vartia@gmail.com)

Antti Suoperä
Statistics Finland, Helsinki[1]

Sisällys

1. Introduction

While explaining Fisher´s (1922) views of biases in index numbers, Samuelson & Swamy (1974, p. 567) claimed that “exactly what zero bias meant was never thought through”. After almost 100 years, we finally are able to present such a theory by correcting some wrong beliefs of Fisher. We manage to complete our previous papers Vartia (1978) and Vartia & Vartia (1984). Fisher´s main error was to label Laspeyres formula as unbiased and `very good`, which probably explains its popularity in Statistical Offices all over the world. In fact, Laspeyres proves to be *contingently biased* and such formulas should never be used if unbiased and *excellent* formulas can be used. This is the case of *complete data* where in addition to prices also old and new quantities are observed. Our presentation is mostly elementary and it is written non-experts in mind. We hope that the paper can be used even as the first introduction to the subject.

[1] We thank Eugen Koev, Heikki Pursiainen and Timo Koskimäki for their comments and constructive criticism.

This paper considers the construction of CPI in statistical offices, when new complete micro data-sets emerge for commodities. We mean by the complete micro data new data-sets, where both prices and quantities are registered for all homogenous commodities in some market and for all periods. This is often referred as scanner data and is part of Big Data. Complete micro data is different to ´Billion Prices Project´ used in his CPI by Cavallo and Rigobon (2016) of MIT, because their data is based on vast number of web-prices only (no quantity information). Inclusion of complete micro data in CPI's will greatly change and simplify the current practices based on Laspeyres formula and complicated rules for elementary aggregates. When the number of commodities increases 50-fold and also the current quantities are observed, the index number calculations should change accordingly.

The price and quantity indices are the macro equivalents or aggregations of the price and quantity relatives. A price index is roughly speaking some average of the price relatives, $p_i^{t/0} = p_i^t / p_i^0$, relevant to an economic unit, say for example consumer. Normally a price index is some well-known weighted average, like arithmetic A, geometric G or harmonic H average, and the weights w_i are value shares based on either base or observation periods. It is well-known that A > G > H always ("a three-tined fork of index numbers") when price relatives vary and their relative differences are approximately equal to half of a variance of relative changes in prices. This result is based on properties of moment means and derived carefully in Vartia (1978) and it contained the critique of Fisher (1922). The main mistakes of Fisher were assuming that Laspeyres and Paasche indices were almost equal and that variances of price and quantity log-changes are essentially the same, which are true only accidentally. The corrected theory is presented in the paper in a more readable and general form than in Vartia (1978). This is a joint work of the authors and continues the co-work of University of Helsinki and Statistics Finland related more generally to aggregation of behavioral functions, see Vartia (2008) and Suoperä and Vartia (2011).

Unfortunately, the selection of the index formulas is quite rarely analyzed purely in terms of the descriptive index number theory. That is, the concept of bias of an index number formula has not been examined accurately, not even in our previous papers. We correct this drawback in chapter 8.

In this paper, we analyze most popular index number formulas and we show how their numerical values depend on the index formula. We show that index number formulas can be classified to formulas having upward or downward bias for small changes compared to index formulas, that must be classified as 'unbiased' according to the index number theory. This analysis is carried out for small changes both in log prices and log quantities, and it leads to exact and definite concepts there. The concept of bias vs. unbiasedness is developed carefully in chapter 7. Five first chapters consider the index number problem and CPI on a general level. Appendices consider more technical matters.

One may ask, what makes the index number theory important. Our answer is that it tries to solve the Arrow´s Paradox (concerning the incompatibility of intuitive rules of social decisions) in a quantitative way. The quantity index of consumption (calculated by a precision formula) is constructed to measure the welfare based on consumption in a quantitative way, see Vartia-Weymark (1982) and Vartia (1983). Economists agree to a large extend that it succeeds in its goal. Although there does not exist an ideal (or the only correct) way to measure welfare, it can be measured accurately enough in many important situations. If we agree the set of precision formulas, the welfare change is measured with practical accuracy. Not exactly, but almost. The paper Vartia (1980) takes into account also the saving decisions to the consumer behavior.

In our treatment, we are able to correct some misconceptions held since Fisher (1922), which resulted from his faulty generalizations of his price-quantity data. For example, Laspeyres and Paasche cannot be considered generally even roughly equal. However, Laspeyres formula has been taken as a practical solution in CPI´s, because quantity information is available only for the base period (based on consumer surveys). Even more dramatically a new observation is, that so-called factor antithesis formulas of the basic price indices and their rectified versions become useless, if the variance of the changes in log-quantities is many times larger than the similar variance of prices. This is a probable situation in future very detailed "Big Data". This case was ignored in Fisher´s authoritative view, where variations in log-changes in prices and quantities were essentially equal.

On the other hand, time antithesis and its use in rectifying formulas remains an excellent idea of Fisher. To our knowledge, this criticism towards “the bible of index numbers” is a new contribution.

The mathematical methods, that we are going to use, are based on well-known references, Fisher (1922), Törnqvist (1936), Stuvel (1957, 1989), Diewert (1976,1978), Vartia (1976, 1978, 1983) and Vartia & Vartia (1984), Törnqvist-Vartia-Vartia (1985). We base our index number theory on basic algebra or arithmetic. No questionable assumptions of “economic behavior” say from consumer theory are needed nor used in our derivations. Economic theorization has a tendency of increasing the list of desired properties of indices, which is usually even too long in the beginning. It is well-known and proved several times, that there cannot be any index numbers satisfying even certain short and intuitively plausible lists of properties or desiderata. This reminds us of the Arrows paradox in macro decision making. For a comprehensive treatment in index number methods within this approach, see Balk (1995). Hansen-Lucas (1983) documents a quantitative comparison of many important index formulas in the foreign trade of Egypt from 1885-1961.

Our target readers are the index number specialist and constructors in the national statistical offices. We concentrate on those properties of indices, which in our opinion are relevant to these specialists in statistical offices, while they are formulating the new emerging methodology of index construction.

2. Complete micro data and its transformations

We verify our mathematical results by numerical analysis. The data used, is well edited drugs data (i.e. self-care drugs) uphold by Pharmaceutical Information Centre Ltd, Finland. It is *complete micro data* in the sense that it contains all prices and quantities (aggregated for the whole country here) for all homogeneous commodities and periods, which are months in the paper.

Instead of couple of dozens product varieties typically included into a CPI sub-index, the data we use contains complete price and quantity data of all periods (here all months during 2013-2016) for about 5000 drugs that require prescription by a medical professional and about 500 self-care drugs (so called VNR-commodities). These are packages of drugs, that have a constant quality over time. This means that problems of quality change do not appear in this data set. Because the actual prices paid by the consumer for prescription drugs are somewhat complicated because of the sickness insurance schemes applied in Finland, we will use only the data on self-care drugs here. It will form our *test data*, which we assume to simulate well (as basic data in CPI) other similar forthcoming complete micro data-sets. We do not consider the peculiarities of this drug data in the paper. In it the prices are rationed and the same in all shops. We consider only features, which we consider to generalize to other complete micro data.

We concentrate on results and proposals that are applicable for similar CPI data having up to 100 times larger sets of commodities than in the current national CPI practices. New and disappearing commodities can also be treated in a systematic and simple way and quality corrections are unnecessary, because the micro commodities are homogeneous. In the new practices, complete price and quantity data appears to become available of practically all commodities in the retail market. We believe that using these new data sets and improved index number methodology, reliability and accuracy of CPI production will be raised to a new level.

In the index number calculations, we split the data in two by five different ways according to how the values of the VNR-commodities have changed between base and observation periods. In the first splitting in two, one part of data 10S consists of the VNR-commodities, whose values have increased or declined more than 10 times and another part 10N is the complement of 10S. In the second split in two the limit is 5, in third 3, in fourth 2 and in the last one 1.33 times increase or decline of the values.

We comment here only the calculations concerning 5N (N for Normal commodities) and 5S (S for Special). The subgroup 5S includes commodities having large relative changes in values (up by multiple 5 or down by 1/5) and thus in quantity or/and in prices. On the other hand, in 5N the values of the commodities stay relatively constant. We show, what kind of effects the special parts 10S, 5S (reported here), 3S, 2S, 1.333S of the data cause to index number calculations. In fact, we shall show some dramatic effects in them. They are so surprising, that one needs some time to digest their consequences.

3. Basic concepts and notation for index numbers

We represent the basic concepts of index numbers as an easy reference for experts in statistical offices. Our notation for the index number calculations is the following:

Commodities:	$a_1, a_2, \ldots, a_n$ are here self-care drugs and their number n is roughly 500.
Time periods:	*t* = 0, 1, 2, … are the compared situations (only two in binary comparisons).
Prices:	p_i^t is the unit price of a_i in period *t*.
Quantities:	q_i^t is the quantity of a_i in period *t*.
Values:	$v_i^t = p_i^t q_i^t$ is the value of a_i in period *t*.
Total value:	$V^t = \sum v_i^t$ is the total value of all the commodities.
Total value ratio:	$V^{t/0} = V^t / V^0$ is the total value ratio from period 0 to *t*.
Price relatives:	$p_i^{t/0} = p_i^t / p_i^0$ is the price relative of a_i from period 0 to *t*.
Quantity relatives:	$q_i^{t/0} = q_i^t / q_i^0$ is the quantity relative of a_i from period 0 to *t*.
Value relatives:	$v_i^{t/0} = v_i^t / v_i^0$ is the value relative of a_i from period 0 to *t*.
Value shares:	$w_i^t = v_i^t / \sum_i v_i^t$ is the value share of a_i in period *t*.

This will be the data, on which the index number calculations are based. Corresponding n-vects are denoted by the same symbol without commodity sub-index:

(1) $$p^t,\ q^t,\ v^t,\ p^{t/0},\ q^{t/0}, v^{t/0}, w^t$$

This looks much easier: prices, quantities, values. etc.! We assume that all prices and quantities are strictly positive (contain no zeros). This implies that all values, price, quantity and value relatives and values shares are also well-defined and strictly positive. Especially both relatives $p^{t/0} = p_i^t / p_i^0$ and $p^{0/t} = p_i^0 / p_i^t$ are defined (no division by zero). This assumption will be relaxed when we later discuss new and disappearing commodities.

Moment mean of order α (any real number) of strictly positive x-values and strictly positive weights v is defined as follows:

(2) $$M^{(\alpha)}(x,v) = \begin{cases} \left(\sum v_i x_i^\alpha / \sum v_i\right)^{1/\alpha}, if\ \alpha \neq 0 \\ \prod x_i^{v_i/\sum v_j} = G(x,v),\ \ if\ \alpha = 0 \end{cases}$$

This will be a crucial part of our theory, how index numbers deviate from each other. The limit of the moment mean when α approaches zero is the weighted geometric mean $G(x,v)$. Moment mean of non-equal arguments *x* is a strictly increasing continuous function of parameter α from $M^{(-\infty)}(x,v) = \min(x)$ to $M^{(\infty)}(x,v) = \max(x)$. For the important small values of α the relative change of the moment mean satisfies

(3) $$log\left(\frac{M^{(\alpha)}(x,v)}{M^{(0)}(x,v)}\right) = \frac{\alpha}{2} s^2(logx, v) + \frac{\alpha^2}{6} m_3(logx, v) + \cdots$$

where the first two terms of the Taylor series (in respect to α) contains the *v*-weighted variance and third central moment in the $logx$-scale. This expansion is very accurate for small variances in $logx$. The variance term dominates and gives alone for all moderate values of *x* and *v* and α the correct magnitude of the change.

The mathematics is so elegant, that we use this for the price index $P = M^{(0)}(p^{1/0}, w^0) = \exp(\sum w_i^0 \Delta log p_i)$ called log-Laspeyres l later. Define $\dot{p}_i = log(p_i^{1/0}/P) = \Delta log p_i - log P$, the log-deviations from the mean.

$$\begin{aligned}
&log\left(\frac{M^{(\alpha)}(p^{1/0},w^0)}{M^{(0)}(p^{1/0},w^0)}\right) = log M^{(\alpha)}\left(\frac{p^{1/0}}{P}, w^0\right) = log M^{(\alpha)}(e^{\dot{p}}, w^0) = \tfrac{1}{\alpha} log \sum w_i^0\, e^{\alpha \dot{p}_i} = \\
&= \tfrac{1}{\alpha} log \sum w_i^0 (1 + \alpha \dot{p}_i + \tfrac{1}{2!}(\alpha \dot{p}_i)^2 + \tfrac{1}{3!}(\alpha \dot{p}_i)^3 + \cdots) = \\
&= \tfrac{1}{\alpha}\log(\sum w_i^0 + \alpha \sum w_i^0 \dot{p}_i + \tfrac{\alpha^2}{2!} \sum w_i^0 \dot{p}_i{}^2 + \tfrac{\alpha^3}{3!} \sum w_i^0 \dot{p}_i{}^3 + \cdots) \\
&= \tfrac{1}{\alpha}\log(1 + 0 + \tfrac{\alpha^2}{2} a_2(\dot{p}, w^0) + \tfrac{\alpha^3}{6} \sum w_i^0 \dot{p}_i{}^3 + \cdots) \\
&= \tfrac{\alpha}{2} a_2(\dot{p}, w^0) + \tfrac{\alpha^2}{6} a_3(\dot{p}, w^0) + \cdots \qquad \text{where } \dot{p}_i = log\frac{p^{1/0}}{P} = \Delta log p_i - log P \\
&= \tfrac{\alpha}{2} m_2(\ddot{p}, w^0) + \tfrac{\alpha^2}{6} m_3(\ddot{p}, w^0) + \cdots \qquad \text{where } \ddot{p}_i = log p^{1/0} = \Delta log p_i.
\end{aligned} \tag{4}$$

In this derivation, the log-deviations $\dot{p}_i = log\frac{p^{1/0}}{P}$ and their zero mean $\sum w_i^0 \dot{p}_i = w^0 \cdot \dot{p} = 0$ are the essential point. Here $a_2(\dot{p}, w^0)$ is the second origo-moment and $m_2(\ddot{p}, w^0)$ the second central moment. Because trivially (but surprisingly, because of the zero mean of $\dot{p}$) $a_2(\dot{p}, w^0) = m_2(\ddot{p}, w^0) = s^2(\ddot{p}, w^0) = s^2(\dot{p}, w^0) = m_2(\dot{p}, w^0)$, it is possible to name this crucial parameter in several different ways. This is first confusing.

The derivation is closely related to the moment generating function $M(\dot{p}, t) = Ee^{t\dot{p}} = \sum w_i^0\, e^{t\dot{p}_i}$ of the log-deviation variable $\dot{p}$, considered as a discrete random variable with probabilities $p = w^0$. Its logarithm is the cumulant generating function $K(\dot{p}, t) = log M(\dot{p}, t) = log E e^{t\dot{p}} = log \sum w_i^0\, e^{t\dot{p}_i}$, which appears above in the form $\frac{1}{\alpha} K(\dot{p}, \alpha) = log \sum w_i^0\, e^{\,\alpha \dot{p}_i}$ and the coefficients of its Taylor series gives the sequence of cumulants. The first cumulant is the mean, second the variance and third the third central moment, as above. Higher cumulants are combination of central moments.

Arithmetic, geometric and harmonic means are moment means with parameter 1, 0 and -1, respectively. For small changes, their relative differences from the geometric mean equal half of the variance in the log-argument. For simplicity, we ignore here the third moment term in the analysis of index numbers, see however Vartia (1978).

We represent in a condensed form the operations or modifications of any index number formula in the following table. Time Antithesis TA and Fact Antithesis FA are basic concepts in Fisher´s (1922) methodology and we shall review them critically in our paper.

Table 1: Modifications of index numbers. Geometric mean as an example

(1)	(2)	(3)	(4)	(5)
Price Index Number Formula f	f as a Quantity Index Number	TA of f	FA of f	CoF of f
f_p	f_q	$TA(f_p)$	$FA(f_p)$	$CoF(f)$
$f\begin{pmatrix} p^1 & q^1 \\ p^0 & q^0 \end{pmatrix}$	$f\begin{pmatrix} q^1 & p^1 \\ q^0 & p^0 \end{pmatrix}$	$1/f\begin{pmatrix} p^0 & q^0 \\ p^1 & q^1 \end{pmatrix}$	$V^{1/0}/f\begin{pmatrix} q^1 & p^1 \\ q^0 & p^0 \end{pmatrix}$	$V^{1/0}/f\begin{pmatrix} p^1 & q^1 \\ p^0 & q^0 \end{pmatrix}$
$l_p = \prod (\frac{p_i^1}{p_i^0})^{w_i^0}$	$l_q = \prod (\frac{q_i^1}{q_i^0})^{w_i^0}$	$p_p = \prod (\frac{p_i^1}{p_i^0})^{w_i^1}$	$V^{1/0}/\prod (\frac{q_i^1}{q_i^0})^{w_i^0}$	$V^{1/0}/\prod (\frac{p_i^1}{p_i^0})^{w_i^0}$

We have presented so-called *Log-Laspeyres* formula l_p as an example of these modifications. Usually, it is represented in logarithmic form: $logl_p = \sum w_i^0 \log(\frac{p_i^1}{p_i^0})$. We just mention for later use two important modifications.

(5) The TA-rectification of f = $\sqrt{f\begin{pmatrix} p^1 & q^1 \\ p^0 & q^0 \end{pmatrix} / f\begin{pmatrix} p^0 & q^0 \\ p^1 & q^1 \end{pmatrix}}$ = Cross (geometric mean) with its TA and

(6) the FA-rectification of f = $\sqrt{f\begin{pmatrix} p^1 & q^1 \\ p^0 & q^0 \end{pmatrix} V^{1/0} / f\begin{pmatrix} q^1 & p^1 \\ q^0 & p^0 \end{pmatrix}}$ = Cross (geometric mean)) with its FA.

Starting from $logl_p$, we derive as its TA-rectification the *Törnqvist index* $logt_p = \sum \frac{1}{2}(w_i^0 + w_i^1)\log(\frac{p_i^1}{p_i^0})$ and as its FA-rectification the index $logP = \frac{1}{2}(\sum w_i^0 \log(\frac{p_i^1}{p_i^0}) + log\frac{V^1}{V^0} - \sum w_i^0 \log(\frac{q_i^1}{q_i^0}))$. We shall show, that the first is unbiased and excellent in a specific sense, but the latter is biased unless the variances of log-changes in prices and quantities happen to be equal. Therefore, Fisher (1922) was in error while considering FA-rectification as a generally valid and effective operation. This is a new result.

4. Laspeyres, Paasche and Palgrave

These are the main building blocks in in number theory. We show their different representations and also try to correct some common misunderstandings. Etienne Laspeyres and Herman Paasche were German economists measuring price changes in Hamburg around 1865 using slightly different methods. Their proposals are the two most important basic indices. As an introduction, we derive for Laspeyres and Paasche price indices their *representations* as weighted averages. For Laspeyres index this is simple:

(7) $$L = L^{1/0} = \frac{p^1 \cdot q^0}{p^0 \cdot q^0} = \frac{\sum p_i^1 q_i^0}{\sum p_i^0 q_i^0,} = \frac{\sum \frac{p_i^1}{p_i^0} p_i^0 q_i^0}{\sum v_i^0} = \frac{\sum p_i^{1/0} v_i^0}{\sum v_i^0,} = \sum w_i^0 \, p_i^{1/0}.$$

We have applied *the inner product notation* such as $p^1 \cdot q^0 = \sum p_i^1 q_i^0$ and $p^0 \cdot q^0 = \sum p_i^0 q_i^0$, which is easier to read, because it removes unnecessary summation symbols and product sub-indices. By definition, Laspeyres price index is the cost relative ($p^1 \cdot q^0$ divided by $p^0 \cdot q^0$) of the old quantity basket q^0. It can also be calculated as a weighted arithmetic average of the price relatives, where the weights are old values shares. This is called *the practical way of calculating Laspeyres*.

It is applied in national *CPI´s* from the level of its commodity groups by taking $p_i^{1/0}$ as the value of the price index $P_i^{1/0}$ of the commodity group and continuing similarly from these. For commodity groups, physical quantities are not available any more, contrary to *elementary aggregates* where they usually exist. This procedure is based on the *consistent aggregation* property Laspeyres formula. Much of common practice in official statistics depends on it, as Pursiainen (2005) stresses.

For the Paasche price index the derivation is similar but more complicated.

(8) $$Pa = Pa^{1/0} = \frac{p^1 \cdot q^1}{p^0 \cdot q^1} = \frac{\sum p_i^1 q_i^1}{\sum p_i^0 q_i^1} = \frac{\sum p_i^1 q_i^1}{\sum \frac{p_i^0}{p_i^1} p_i^1 q_i^1} = \frac{\sum v_i^1}{\sum p_i^{0/1} v_i^1} = \frac{1}{\sum w_i^1 p_i^{0/1}}$$

By definition, Paasche price index is the cost relative ($p^1 \cdot q^1$ divided by $p^0 \cdot q^1$) of the new quantity basket q^1. It can also be calculated as a weighted *harmonic* average of the price relatives, where the weights are new values shares. In practice, this is the way it is calculated. It can be used only if the new value shares are known, which is a strong condition. In our complete micro data on drugs, also new quantities are known and new values shares can be calculated. This holds for any scanner-type new micro data and it changes strongly the whole methodology of CPI calculation.

Note also that Laspeyres and Paasche are based on the same basket idea of calculation: L is based on the old basket of goods q^0, while Pa uses the new basket q^1. Paasche may be defined actually by Laspeyres calculated in the reverse direction $1 \rightarrow 0$, i.e. it is evidently the inverse of

(9) $$L^{0/1} = \frac{p^0 \cdot q^1}{p^1 \cdot q^1}.$$

This dependence means actually, that Paasche is the *Time Antithesis* of Laspeyres, $Pa = TA(L)$. This concepts is introduced in detail in chapter 8.

An elementary but quite common mistake is to imagine, that the Paasche price index as a weighted *arithmetic* average of the price relatives weighted by new values shares. This is not Paasche, but the index number proposed by an economist Palgrave in 1868. It does not have a quantity basket interpretation, like Laspeyres and Paasche. Palgrave has two doses of upward bias compared to Paasche. Only if Paasche is strongly biased down, Palgrave index as such is applicable. As shown by Vartia-Vartia (1978) an average of Laspeyres and Palgrave is *always* biased upwards compared to Fisher (geometric average of Laspeyres and Paasche) or to other unbiased formulas.

This kind of error was made e.g. in the Bank Finland in the choice of the formula in the Bank of Finland Currency Index, see Vartia-Vartia (1978). Economists in the Bank of Finland evidently thought, that they calculated Edgeworth index. It exaggerated the revaluation in Finnish currency by the amount of the bias of the index. The currency index increased more than was the actual case. The bias and claimed unintentional revaluation of Finnish mark had increased from 1974 to 1984 to 2.4%. Concretely, a corresponding devaluation of 2.4% in 1984 would have been needed to eliminate the discrepancy for 1984. This large devaluation (which would have corrected the actual situation to the officially declared one) would have meant an extra income of roughly FIM 1500 million (250 million euros in 1984 prices) only in 1984 to the export sector. Thus, the export sector of the economy was officially claimed to get systematically more foreign currency than they actually got. Official data did not correctly describe the reality. As this example shows, biases and other weaknesses of index calculations are not unimportant details in the macro-economic analysis. The formulas and weights must be chosen correctly in official indices.

5. Three base and observation period weighted means of price relatives

Irving Fisher (1922) visualized the index numbers as forks containing a certain number of tines. This is an important visualization of differences between various indices and other choices of calculation. In his data, various index numbers happened to fall into five quite separate groups or “tines”, which formed “the five-tined fork”, a kind of quantum theory of index numbers. Much of this was accidental and resulted because *L* and *P* happened to be near each other, which is not generally or even usually true. Fisher regarded these formulas as “very good”, which was a mistake as will be shown below. Fisher´s “Five-tined Fork” with doses of bias (or quanta) = $0, \pm 1, \pm 2$ as a realistic “quantum theory” of index numbers was criticized already in Vartia (1978), but it did not raise much attention. The concepts of bias of an index number was defined for the first time along these lines in Vartia (1978) and Vartia-Vartia (1984) as a small deviation concept, but evidently too loosely. Also, this point was bypassed unnoticed among index number literature. Probably it was not understood, *how* it differed from the Fisher´s concept of bias and unbiasedness. These concepts are now presented here in chapter 7 and in Appendix 1 once again, but more accurately.

Laspeyres has been regarded as the “correct prototype” of all index calculations all over the world both as a price and quantity index because of its simplicity and easy applicability and despite of all sorts of opposing evidence on its upward bias and asymmetry.

Fisher (1922) made some wrong generalizations on these matters. Now almost 100 years later, it is time to build a corrected index number theory on the following foundations: how various choices on averaging and weighting affect index numbers and where the unbiased indices situate. Now the analysis can be based on modern mathematics (moment means, approximations, function theory, knowledge of new "excellent" or "ideal" index numbers), large data sets and computing power. We must, however, appreciate here Fisher´s genius, as his concepts and classifications still after 100 years govern the development of index theory.

Next, we define three price indices based on arithmetic, geometric and harmonic means of price relatives and old value shares. The three-tined fork having base period weights contains the following indices (three weighted averages *A*, *G* and *H*):

$$L = A\left(p^{1/0}, w^0\right) = \sum w_i^0\, p_i^{1/0} \;= \frac{p^1 \cdot q^0}{p^0 \cdot q^0} \qquad \text{"Laspeyres"}$$

(10) $$l = G\left(p^{1/0}, w^0\right) = \prod \left(\frac{p_i^1}{p_i^0}\right)^{w_i^0} = exp\left(\sum w_i^0\; logp_i^{1/0}\right) \qquad \text{"Log-Laspeyres"}$$

$$Lh = H\left(p^{1/0}\,, w^0\right) = 1/\sum w_i^0\, p_i^{0/1} \qquad \text{"Harmonic Laspeyres"}$$

We also use the upper double-dot to denote log-change, as in $\ddot{p}_\iota = \Delta logp_i = \log(p_i^1/p_i^0)$ and similarly for quantities $\ddot{q}_\iota$ and values $\ddot{v}_\iota$. Logarithms are natural, of course, see Törnqvist-Vartia-Vartia (1985). Single upper dot is reserved for log-deviations from the mean as in Vartia (1978), e.g. $\dot{p}_\iota = \Delta logp_i - logP = \ddot{p}_\iota - logP$.

More symmetrically written $logl = \sum w_i^0 \log(p_i^1/p_i^0) = \sum w_i^0\; \ddot{p}_\iota = \ddot{l}$. Using (3-4) for moments means, the relative differences of $L = M^{(1)}\left(p^{1/0}, w^0\right)$ and of $Lh = M^{(-1)}\left(p^{1/0}, w^0\right)$ from of $l = M^{(0)}\left(p^{1/0}, w^0\right)$ are approximately half of the "old" variance $s_0^2(\dot{p}) = s^2(\dot{p}, w^0)$ in the deviations of log-prices[2]. We call $s_0^2(\dot{p})$ shortly the old variance of $\dot{p}$.

We have always $logL \;> logl >\; logLh$ for non-equal price relatives or $\dot{p}$. More accurately, we have $logL = a + \Delta_0 > a > logLh = a - \Delta_0$, where $a = logl \; and \; \Delta_0 = \frac{1}{2} s_0^2(\dot{p})$.

The three-tined fork with the observation period weights contains also three weighted averages:

$$Pl = A\left(p^{1/0}, w^1\right) = \sum w_i^1 p_i^{1/0} \qquad \text{"Palgrave"}$$

(11) $$p = G\left(p^{1/0}, w^1\right) = \prod \left(\frac{p_i^1}{p_i^0}\right)^{w_i^1} = exp\left(\sum w_i^1\; logp_i^{1/0}\right) \qquad \text{"Log-Paasche[3]"}$$

$$Pa = H\left(p^{1/0}, w^1\right) = 1/\sum w_i^1\, p_i^{0/1} = \frac{p^1 \cdot q^1}{p^0 \cdot q^1} \qquad \text{"Paasche"}$$

More symmetrically written $logp = \sum w_i^1 \log(p_i^1/p_i^0) = \sum w_i^1\; \ddot{p}_\iota = \ddot{p}$. The logarithmic differences of Pl and Pa from p are half of the "new" variance $s_1^2(\dot{p}) = s^2(\dot{p}, w^1)$ of the changes of log-prices $\dot{p} = \Delta logp$. We call $s_1^2(\dot{p})$ shortly the new variance of $\dot{p}$.

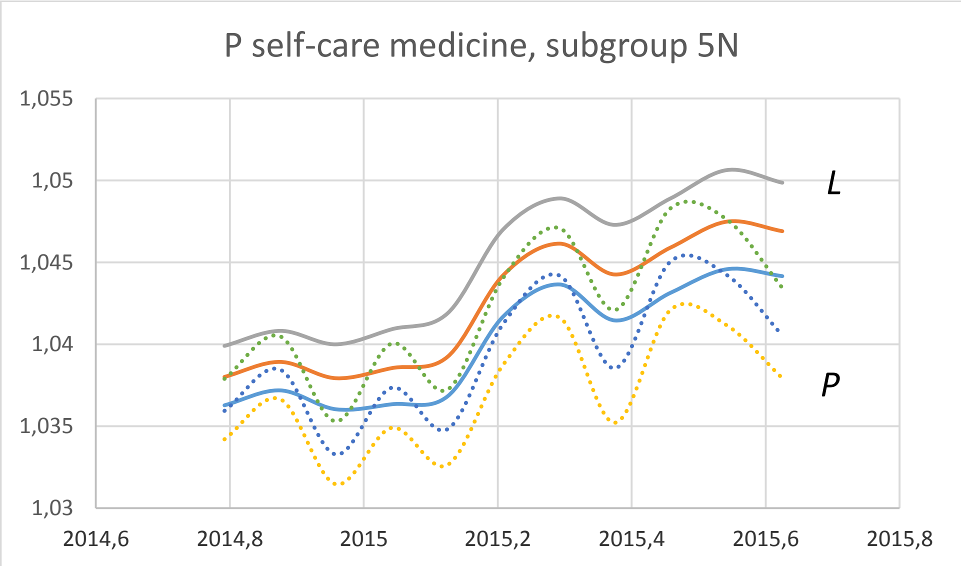
[2] The same variance may be written also $s^2(\ddot{p}, w^0)$ using the log-changes $\ddot{p}_\iota = \Delta logp_i = \log(p_i^1/p_i^0)$.

[3] Log-Paasche and log-Laspeyres are also called Geometric Paasche and Geometric Laspeyres, especially if confusion to the logarithms of Paasche $logPa$ and Laspeyres $logL$ are at stake.

We have always $logPl > logp > logPa$ for non-equal price relatives or $\dot{p}$. More accurately, we have $logPl = b + \Delta_1 > b > logPa = b - \Delta_1$, where $b = logp$ and $\Delta_1 = \frac{1}{2}s_1^2(\dot{p})$.

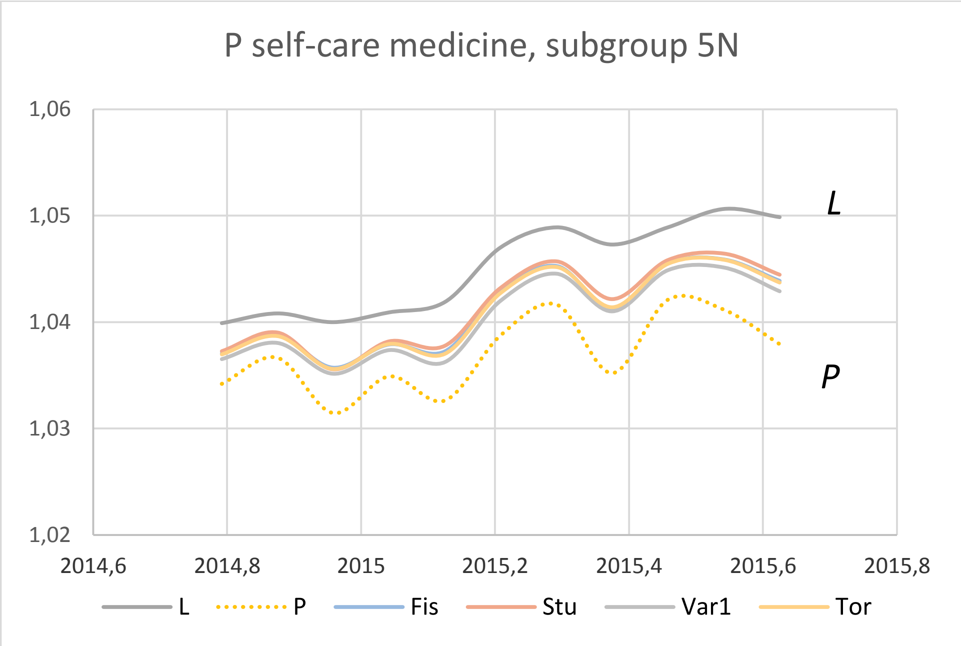
The old and new variances $s_0^2(\dot{p})$ and $s_1^2(\dot{p})$ govern the log-differences of the three indices of the old and new forks.

The problem remains how the three-tined forks of the old and new weights situate in respect to each other. As data shows, there does not exist any firm geometric rules. Especially, Fisher´s ´Five-Tined Fork´ was only an accident, not any rule. In our data for the self-care medicines, Laspeyres and Paasche had for every month large biases up and down, respectively. They did not lie in the center but on the opposite sides of the index fork and they were the most biased of these six formulas all the time. Also, the number of tines is not five, as in Fisher´s data, but varies between the extremes three and six, see figure below.



As the figure shows, the type of the fork that the six indices form, varies from month to month. But here the Laspeyres-type fork with old weights, Laspeyres on the top, is surprisingly always higher than the new weight fork, where Paasche is the lowest index. All other basic indices remain between *L* and *Pa*.

The precision formulas form a much tighter band within the six indices here:

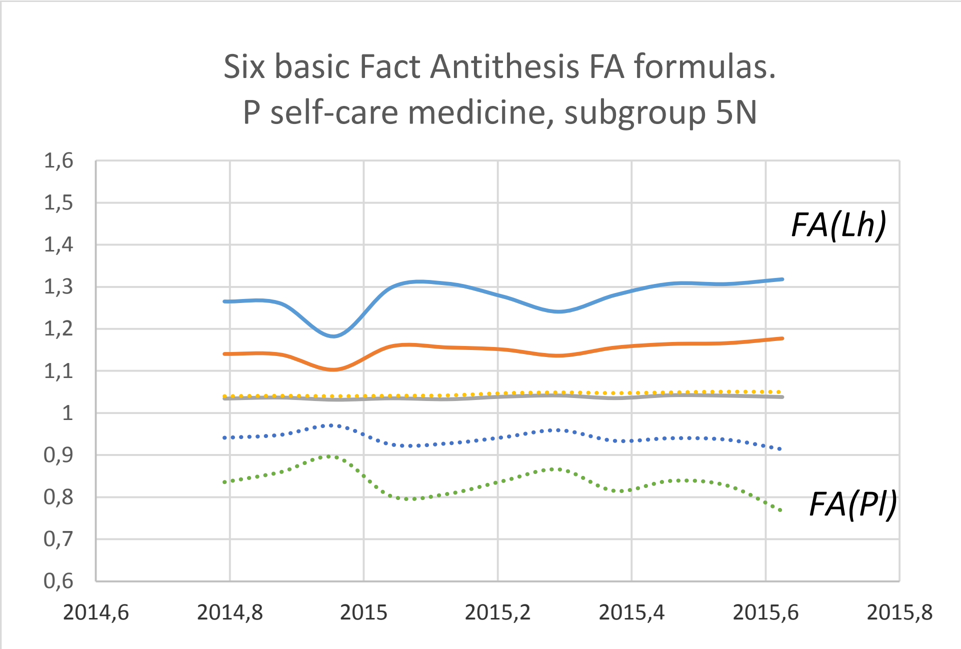


All sensible or minimally accurate index numbers should remain within these limits.

As a warning, we show a picture concerning the six factor antithesis formulae, like $FA(l) = V^{1/0}/l(q)$ and $FA(p) = V^{1/0}/p(q)$. This factor antithesis of l arises by *first* using the same formula for quantities, $l(q)$, instead of prices, and *then* returning back to the price space by calculating its co-factor $CoF\big(l(q)\big) = V^{1/0}/l(q) = FA(l)$. As a technical operation, this is simple, but *conceptually* it requires some expertise in index numbers. Note that the log-difference of two *FA´s* depends only on the difference of the formulas in the quantity side, say: $log\frac{FA(p)}{FA(l)} = log\frac{V^{1/0}}{p(q)} - log\frac{V^{1/0}}{l(q)} = log\frac{l(q)}{p(q)}$. Therefore, the *FA´s* of the six basic indices vary much more than their original price equivalents. The FA’s for the sex indices have represented in Appendix two.

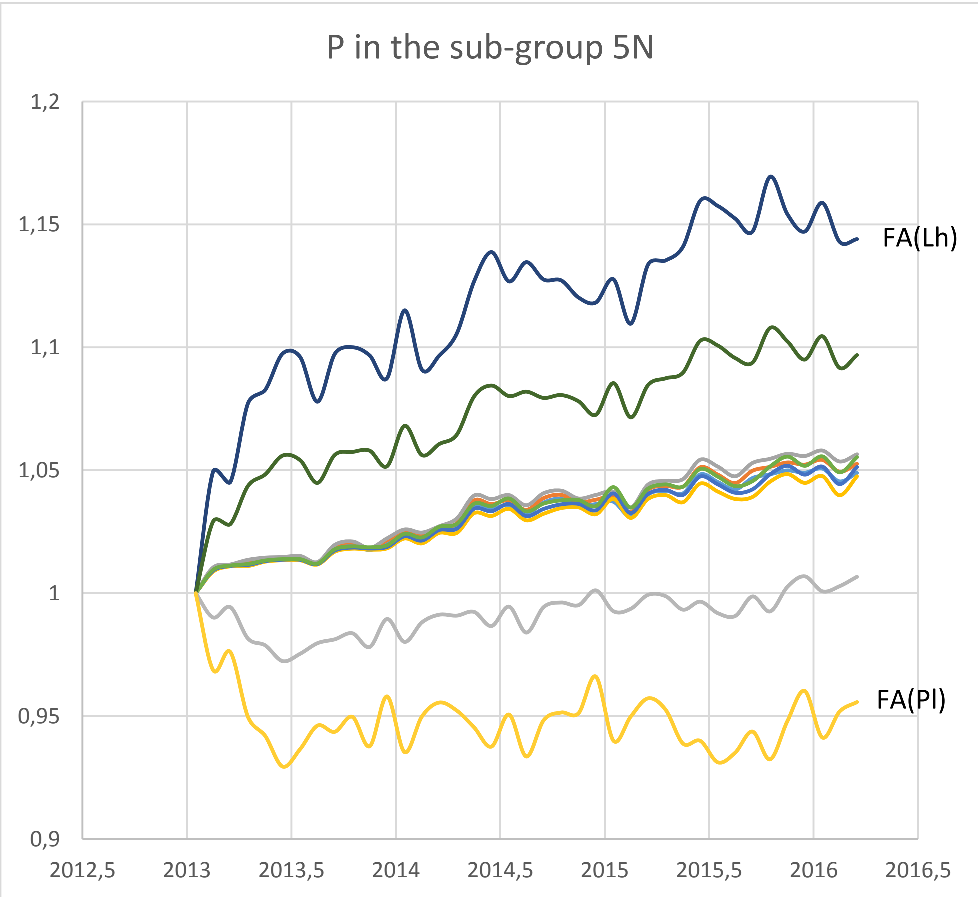
The FA´s of the six basic formulas do not remain within the limits of the picture above and it shows that there is something seriously wrong with them. They are based on applying the six basic indices in the quantity side of the monthly data, where the variance of the quantity changes is roughly 35 times bigger than the similar price variance. Therefore, the FA´s of the six formulas have index forks dispersing roughly 35 times more than the ordinary price index forks in our data. In Fisher´s yearly test data, this was not the case, but instead the variations in price and quantity relatives were exceptionally almost the same. Many of Fisher´s conclusions and suggestions based on the Factor Antitheses FA and on the symmetrical treatment of p´s and q´s are based on this accidental and normally false condition.

We cannot get support from the quantity side of data, because it disperses so heavily!



Note how Laspeyres and Paasche appear in the middle of the picture. They are OK and everything else is out of lines.

In the next picture, we show all the six basic indices and their FA´s in the figure. Note how much tighter the original six price indices move compared to their FA´s.

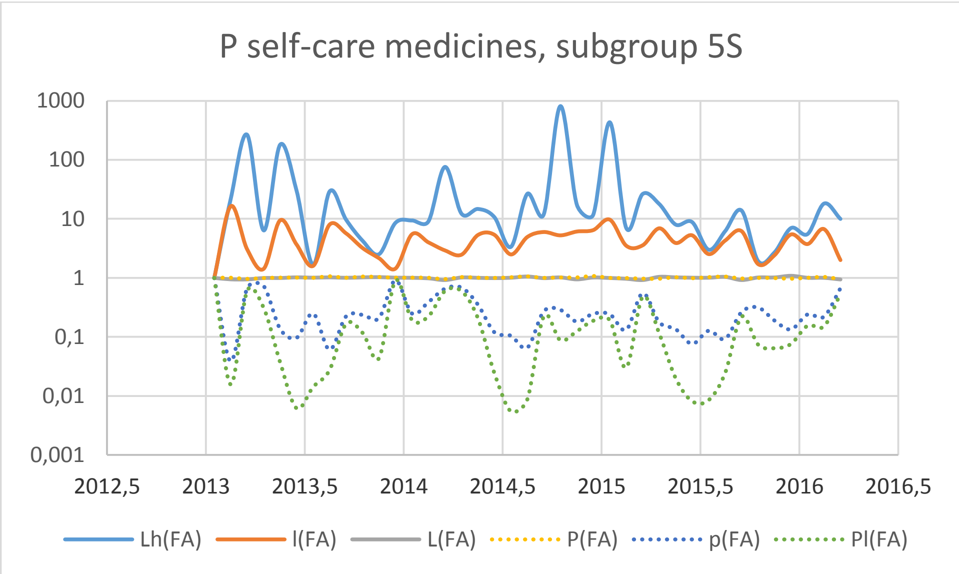


We will comment later some Fisher´s false conclusions which depend on this exceptional feature of his data. In practice, this means that price indices derived from quantity relatives, so called factor antithesis formulas for prices, become less and less useless as the variance of the changes in *quantities* increases. These effects are no minor nuances, because FA-indices in the sub-group of special commodities 5S will be found literally useless (except L and Pa, see Appendix 2).

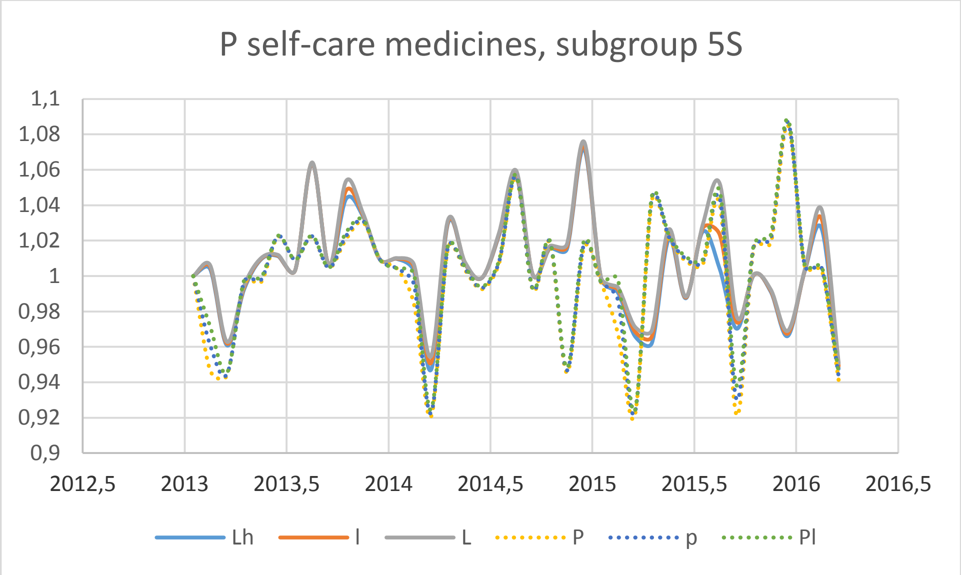
As a concrete conclusion, the factor antithesis formulas of the basic six price index formulas are as such useless formulas because of their large biases (about 35 times larger than for the basic indices). This rejection of FA formulas does not hold for the FA´s of L and P, because their FA´s are the same indices but in the reverse order: $FA(L) = P$ and $FA(P) = L$. Check that these indices satisfy $L * P(q) = P * L(q) = \frac{V^1}{V^0}$ with obvious notation. Prove that for Fisher´s indices $F * F(q) = \frac{V^1}{V^0}$. Hint: multiply the former equations, group their terms and take a square root.

Price indices should not be calculated as functions of the same quantity indices (or cofactors) of the six basic formulas (or their rectifications). Indices calculated via quantity data varied roughly in the same proportion as the variances $s^2(\dot{p}, w)$ and $s^2(\dot{q}, w)$ of price and quantity log-changes!

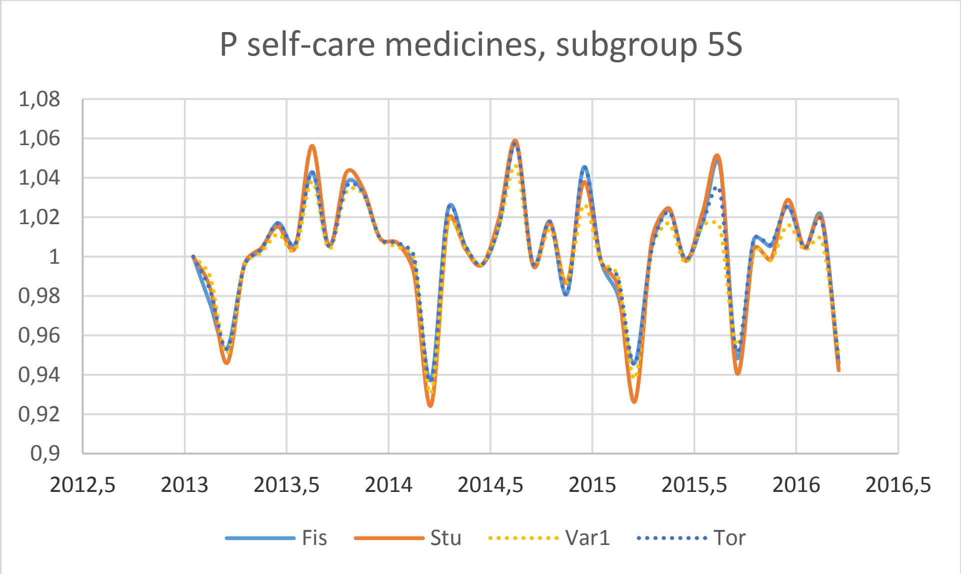
The FA´s of the six basic indices (and their derivations, expect the ones based on L and Pa) become totally useless in for the group 5S of special commodities.



Also, here L and P are OK, and everything else is out of bounds. Factor antithesis indices vary from 0,01 to 1000, as the basic price indices vary within 0,9 - 1,1, as shown below.



The precision formulas are still much more accurate. They do an excellent job also in the set 5S of special commodities having strongly variable values. They vary systematically and their differences are maximally only three percent.



6. Measuring the $Shift$ of an index number P in respect to Log-Laspeyres

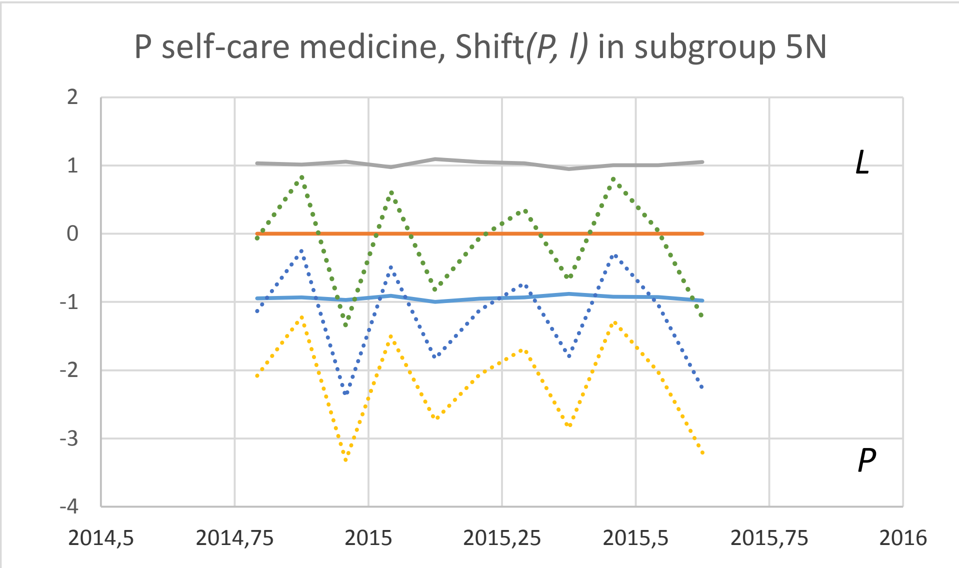
We transform the figure above to a form which better shows the deviations of its all indices from the *Log-Laspeyres l*. We want to compare indices for different periods and for hypothetical data, where changes in prices and quantities approach zero or some proportional values. Therefore, we divide the relative difference of the compared indices by the average relative difference of our six basic indices. We define the *Shift(p,l)* of an index number (say *p*) in respect to *Log-Laspeyres l* as the log-difference of *p* and *l* in bias units[4]. The unit bias equals the average deviation of tines from the central tine:

[4] Compare the definition of Shift with the t-score $t = (\bar{x} - \mu)/(s/\sqrt{n})$ in statistics.

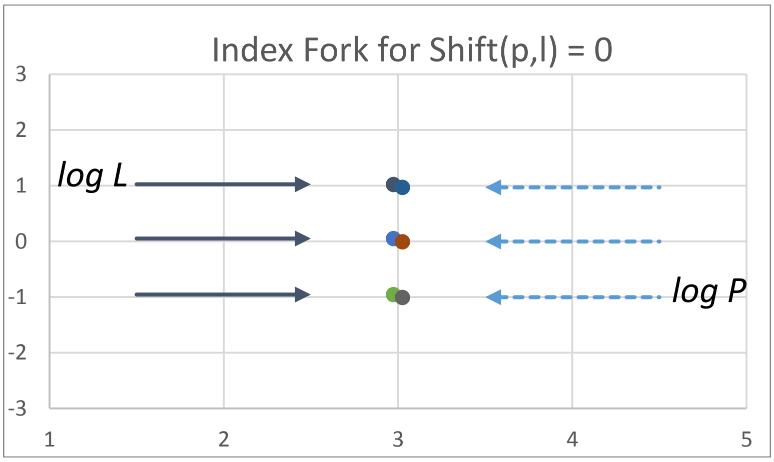
(12) $$S(p) = Shift(p, l) = \frac{\log(\frac{p}{l})}{\frac{1}{4}(\log\left(\frac{L}{Lh}\right)+\log\left(\frac{Pl}{P}\right))} \cong \frac{\log(\frac{p}{l})}{\frac{1}{2}s^2(\dot{p}; \overline{w})}.$$

$S(P) = Shift(P, l)$ is the relative difference of any index P from l expressed in bias units. The denominator, half the variance of $\dot{p}$, acts as a magnifying glass which expanses the picture when the deviations of $\dot{p}$ are small or, especially, when he unit bias $\frac{1}{2}s^2(\dot{p}; \overline{w})$ approaches zero and the 6 basic indices approach the same value. Note that, although $Shift(P, l)$ approaches the form $0/0$, it may have a definite real limit possibly depending on the actual data, see Appendix 1.

For the indices (L, l, Lh) of the old fork the shifts become $\left(S(L), S(l), S(Lh)\right) \cong (1, 0, -1)$ with essential equality at least for small deviations in log-changes of prices and quantities. For the indices (Pl, p, P) of the new fork the shifts $\left(S(Pl), S(p), S(Pa)\right)$ are more interesting. They show, how the new fork deviates (up or down) from the old fork in the normalized scale. For the first period, we have a four-tined fork and for the last a five-tined fork. All the forks differ much from Fisher's data where $L \approx Pa$ as here L and Pa are the largest and lowest figures. We will show how we can get rid of all assumptions related to their sizes and mutual differences.



The small fluctuations in the above figure disappear completely, if we consider the asymptotic case, where the changes in the prices approach zero or some proportional values. These can be calculated by somewhat complicated algebra, see Appendix 1. The case $S(p) = Shift(p, l) \cong 0$ where l and p happen to be almost equal, looks as follows:



The shift parameter S of the 6 indices in the previous figure varies from 1 to -3 and the combined fork has 3, 4 or 5 tines, as the Paasche-type indices (new 3-fork) shift downwards in this data. This data clearly contradicts Fisher´s main conclusion, that L and Pa are unbiased and “very good” index numbers. This conclusion as a general advice was a serious mistake, while they may be strongly biased, as in this data. None of our six indices cannot be classified as unbiased, because in some data their biases are clearly realized. But we shall show in a minute that their rectifications based on their time antithesis are all unbiased and excellent index numbers for *small* price and quantity changes (not necessarily for large changes).

7. Measuring the $Shift$ of an index number P in respect to Fisher

One of the basic features of Fisher´s theory is to relate all other index numbers to it. This makes sense perfectly for small changes but can be questioned for large ones. The *Shift* of an index number formula P, say, $Shift(P;\cdot)$, is naturally defined as a shift S in respect to Fisher, Törnqvist or any other *for small changes unbiased or excellent* index number formula. This helps us to define the *bias* of an index number for small changes. The question of *bias* must be left open for moderate or large changes. Therefore, we call our function symmetrically as a shift between two formulas. To be explicit, we decide to use the traditional Fisher index as the basic point of reference. The expression for $Shift(P;\ F)$ for actual or moderate changes is defined as follows

(13) $$Shift(P,F) = \frac{\log\left(\frac{P}{F}\right)}{\frac{1}{4}\left(\log\left(\frac{L}{Lh}\right)+\log\left(\frac{Pl}{P}\right)\right)} \cong \frac{\log\left(\frac{P}{F}\right)}{\frac{1}{4}\left(s^2(\dot{p},w^0)+s^2(\dot{p},w^1)\right)} \cong \frac{\log\left(\frac{P}{F}\right)}{\frac{1}{2}s^2(\dot{p},\overline{w})}\ .$$

Note that this is an anti-symmetric function: $Shift(F,P) = -Shift(P,F)$. Irrespective of the size of the denominator, some values of $Shift(P,F)$ for our six basic indices P must be outside or at the limits of the interval (-1, +1). Think of the case $logl \approx logp$ shown above, where the two three-forks essentially coincide. $Shift(P,F)$ is well-defined and attains similar values also when its denominator or, more radically, all commodity deviations in log-changes of prices and quantities tend to zero (almost proportional changes APC). These are needed in the definitions of asymptotic bias below.

An index number P is called *unbiased and excellent*[5] *for small changes* only if $Shift(P,F) \to 0$ or its numerator approaches zero quicker than its denominator. This happens for indices, which are quadratic approximations of F for small changes in log-prices and log-quantities. For them the denominator of $Shift(P,F)$ is of third degree smallness in logs and approaches zero quicker than the typical quadratic expression, the variance, in the nominator. Fisher´s theory of *Five-tined Fork* was clearly mistaken, see

[5] Excellent was the second-best group of index numbers after superlative by Fisher (1922).

Vartia (1978) and Vartia-Vartia (1984), and no replacement has been generally presented this far. These concepts aim at developing such a new theory. We call quadratic approximations of F for small changes as *excellent* and not as Pseudo-Superlative as Diewert (1978) does. He motivates this terminology that a *Pseudo*-Superlative formula is a good approximation but is not exact for a flexible functional form. Pursiainen (2016) shows convincedly that Pseudo-Superlative (our excellent formula) approximates the unknown ideal price index as well any superlative index and their separation is not motivated from the point view of accuracy. Our *Shift*-concept this shows nicely without explicit reference to exactness and superlativity.

The function $Shift(P,F)$ takes the role of magnifying glass when its denominator is small or approaches zero. It is essential how it behaves, when its denominator approaches zero. $Shift(P,F)$ is expressed in terms of contracting variables $\left(p(t),q(t)\right) \to (p^0,q^0)$, when $t \to 0$. These are defined in log-scale by $\log p(t) = \log p^0 + t(\log p^1 - \log p^0)$, where $t \in [0,1]$ and similarly for quantities. Thus, the log-changes $\log(\,p(t)/p^0) = t\log(\,p^1/p^0)$ and $\log(\,q(t)/q^0) = t\log(\,q^1/q^0)$ approach zero together with t. Also, $\log(\,v(t)/v^0) = t\log(\,v^1/v^0) \to 0$, $V(t) \to V^0$, and $w(t) \to w^0$, when $t \to 0$. Now calculate the price index P for these prices and quantities approaching their base period values:
$P = P(t) = f\begin{pmatrix} p(t) & q(t) \\ p^0 & q^0 \end{pmatrix}$. The same is done for the Fisher index F and for the variance term $\frac{1}{2}s^2(t\log(\,p^1/p^0)\,, \frac{1}{2}(w(t)+w^0)$. Now we can calculate the shift-function for these contracting prices and quantities as

(13b) $$Shift\left(P(t),F(t);\begin{matrix} p(t) & p^1 & p^0 \\ q(t) & q^1 & q^0 \end{matrix}\right) = \frac{2\log\left(\frac{P(t)}{F(t)}\right)}{s^2\left(t\log(p^1/p^0),\ \frac{1}{2}(w(t)+w^0)\right)}\,.$$

Its limit is defined as the Bias of P:

(13c) $$Bias(P) = \lim_{t\to 0} Shift\left(P(t),F(t);\begin{matrix} p(t) & p^1 & p^0 \\ q(t) & q^1 & q^0 \end{matrix}\right).$$

We classify the index numbers P according to their $Bias(P)$. **Especially important is, whether this function has a unique limit value (independently of data) when price and quantity changes approach zero or is it data-dependent.** This determines whether the price index formula $P$is **unbiased and excellent, permanently biased or contingently biased for small changes.** We formulate first the concept of *bias and unbiasedness* of an index number formula *for small changes*:

An index number P is said to be ***excellent for small changes*** (*for SC*)
if and only if
its $Bias(P)$ exists as a unique real number for all $\begin{pmatrix} p^1 & p^0 \\ q^1 & q^0 \end{pmatrix}$ and equals zero.

An index number P is said to be ***permanently biased for small changes*** (*for SC*)
if and only if
its $Bias(P)$ exists as a unique real number for all $\begin{pmatrix} p^1 & p^0 \\ q^1 & q^0 \end{pmatrix}$ and it is non-zero.
This non-zero value is the value of the permanent bias (a fraction of the norm up or down compared to F) in the index[6].

An index number P is said to be ***contingently biased for small changes*** (*SC*)

[6] In these definitions we used simple the *log-linear curves* $p(t)$ connecting $p^1 = p(1)$ to $p^0 = p(0)$ and similarly for quantities, which implies the use of directional derivatives. In case of definitions of excellent and permanently biased index numbers we could (and perhaps should) use *any curves* $p(t)$ connecting $p^1 = p(1)$ to $p^0 = p(0)$ and similarly for $q(t)$ as is essentially done in Appendix 1. We believe, but have not proved yet, that arbitrary curves and log-linear curves lead to the same sets of excellent and permanently biased index numbers. It may be that the definitions given above are slightly too liberal.

if and only if

its $Bias(P)$ is not a unique real number for all $\begin{pmatrix} p^1 & p^0 \\ q^1 & q^0 \end{pmatrix}$ but depends on $\begin{pmatrix} p^1 & p^0 \\ q^1 & q^0 \end{pmatrix}$. Different real numbers are approached depending on data $\begin{pmatrix} p^1 & p^0 \\ q^1 & q^0 \end{pmatrix}$.

There are only three groups of index numbers:

1 **Excellent** if $Bias(P) = 0$ for all data. These indices P are quadratic approximations of F. Excellent index numbers do not differ from each other for small changes. These differ from F as follows: $LogP = logF + 0 * s^2(\ddot{p}, \overline{w})$ +third and higher order terms in log-changes of prices and quantities. This means that $LogP$ and $logF$ have the same first and second order terms in the Taylor-series expressed in their small log-changes of prices and quantities.

2 **Permanently biased** if $Bias(P) = b \neq 0$ for all data. They are permanently biased up or down by b. These differ from F as follows: $LogP = logF + b * s^2(\ddot{p}, \overline{w})$ +third and higher order terms in log-changes of prices and quantities.

3 **Contingently biased** if $Bias(P)$ is not a unique real number for all data but depends on it. For contingently biased indices $LogP$ is only a linear approximation of $logF$, i.e. it differs from $logF$ by quadratic and higher order terms. Indices P differ from F in a haphazard way even for small values of log-prices and log-quantities, because they depend on the contingent features of the data. For example, $Bias(L) = -cov(\ddot{p}, \ddot{q}; w^0)/var(\ddot{p}; w^0)$ and $Bias(Pa) = cov(\ddot{p}, \ddot{q}; w^0)/var(\ddot{p}; w^0)$.

There is no reason of using permanently biased or contingently biased formulas in case of complete micro data, because there are an infinite number of excellent index numbers, where one can choose of, see Appendix 1. There are no finite steps or quanta by which permanently biased or contingently biased formulas can only differ from *F* or its quadratic approximations. These formulas differ continuously from *F*: there does not exist a valid “quantum theoretic” correction of Fisher´s mistaken Five-tined Fork. The whole idea of Fisher was based on classifying only proposed or other apparent formulas, not all possible formulas. For instance, any weighted mean of *l* and *L* (say $0.1l + 0.9L$) is a valid contingently biased formula. Similarly, any weighted mean of *t* and *F* (say $0.1t + 0.9F$) is a valid excellent formula.

Fisher´s original Five-tined Fork was an excellent selling argument, which succeeded to sell Laspeyres contingently biased and suspicious formula *L* as a standard formula for wide public index number production. It took almost a century to show convincingly that it was good selling argument but against the facts.

This was the idea of definition of unbiasedness in Vartia (1978) and Vartia-Vartia (1984), which was formulated partly verbally and did not raise much attention. In vector notation, the smallness condition means that the (unweighted Euclidean) length of vectors $\|\Delta logp\| = \sqrt{\sum(\Delta logp_i)^2} \to 0$ and $\|\Delta logq\| \to 0$. This implies e.g. $\|\Delta logv\| \to 0$ and $\|\Delta logw\| \to 0$. Also $logP \to 0$, $logQ \to 0$ and $log\frac{V^1}{V^0} \to 0$. All these conditions correspond to “small changes”.

The second definition of *bias and unbiasedness* is a stricter one and is related to almost proportional changes, not to almost zero changes as above. Here an index number *P* must be almost equal (a quadratic approximation) to *F* to be unbiased and excellent for APC already when the log-changes of prices and quantities approach any proportional values. This condition of inputs can be described as “the commodity variation in log-prices and log-quantities vanishes” or the variances $s^2(\dot{p}, \overline{w})$ and $s^2(\dot{q}, \overline{w})$ tend to zero. Now the contracting variables $(p(t), q(t))$ are defined in the log-scale by $\log(p(t)/p^0) = \log l_0^1(p) +$

$t(\log(p^1/p^0) - \log l_0^1(p))$ and $\log(q(t)/q^0) = \log l_0^1(q) + t(\log(q^1/q^0) - \log l_0^1(q))$, where $t \in [0,1]$. Here $(l_0^1(p), l_0^1(q))$ is conveniently chosen as the pair of log-Laspeyres indices. This reduces to the former price and quantity curves by setting $(l_0^1(p), l_0^1(q)) = (1,1)$.

An index number *P* is said to be ***excellent for almost proportional changes*** (for *APC*)
if and only if
its $Shift(P, F)$ in respect[7] to *F* **approaches a unique limit** for all data and it is zero, when all log-changes of prices and quantities approach any proportional values. The same zero limit must be attained by all different curves of prices and quantities. Now the variances $s^2(\dot{p}, \overline{w})$ and $s^2(\dot{q},\ \overline{w})$ of the log-changes in p´s and q´s around their mean values both tend to zero, while $logP$ and $logQ$ may get any constant values.

An index number *P* is said to be ***permanently biased for almost proportional changes*** (*APC*)
if and only if
its $Shift(P, F)$ in respect to *F* **approaches a unique limit** for all data which is a non-zero real number, when all log-changes of prices and quantities approach any proportional values. The same non-zero limit must be attained by all different curves of prices and quantities. This non-zero value is the value of the permanent bias (a fraction of the norm up or down compared to F) in the index.

An index number *P* is said to be ***contingently biased for almost proportional*** (*APC*)
if and only if
its $Shift(P, F)$ in respect to *F* **does not approach the same unique limit** for all data, when all log-changes of prices and quantities approach any proportional values. Different real numbers are approached depending on the situation or data, from which the log-prices and quantities approach constant (commodity independent) values.

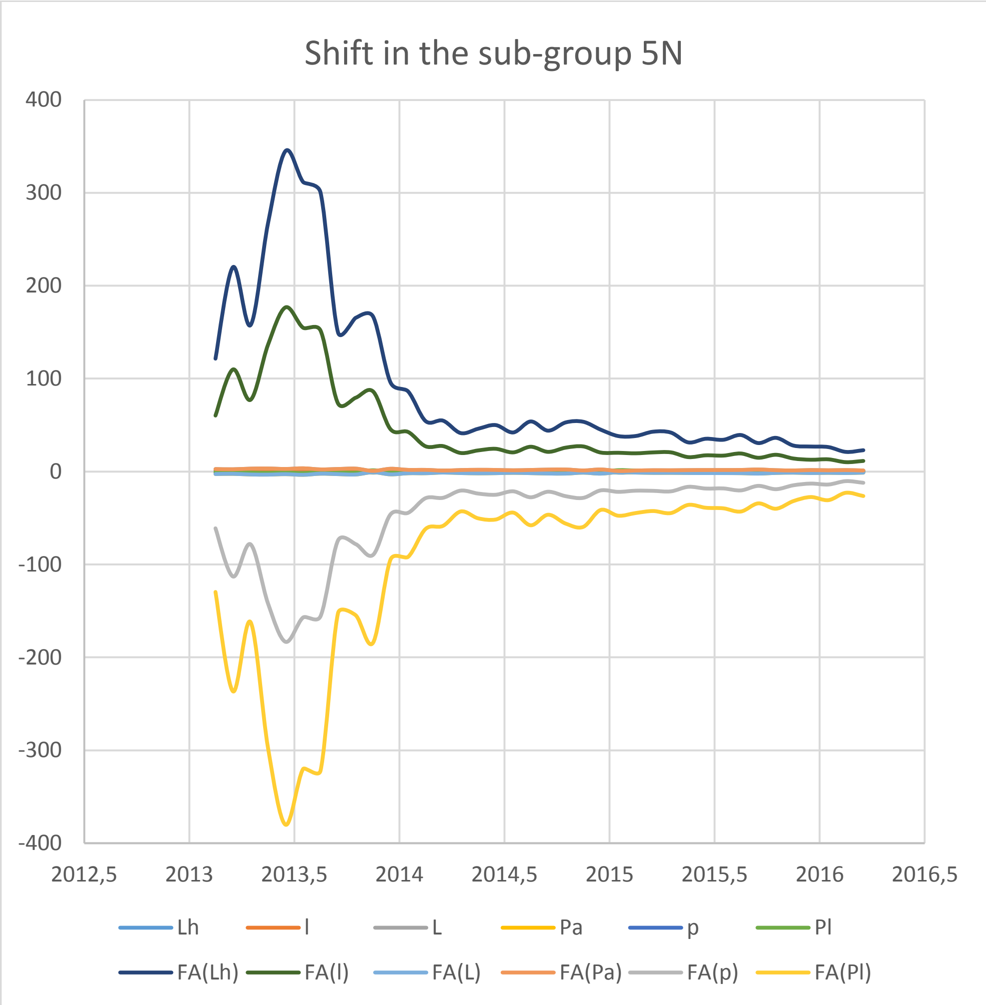
The second definition is stronger and demands more and is presented here only for later reference and in order to give some clarifying comments. If an index is excellent for almost proportional changes, it is necessarily excellent for small changes. It is rather remarkable, that e.g. *F* and t approximate each other equally well irrespective of how much the prices and quantities have changed (i.e. how large their proportional values have been). *MV* is excellent for SC, but it is not excellent for APC.

We use mostly to the first definition, which seems to conform best to the current research in index number theory.

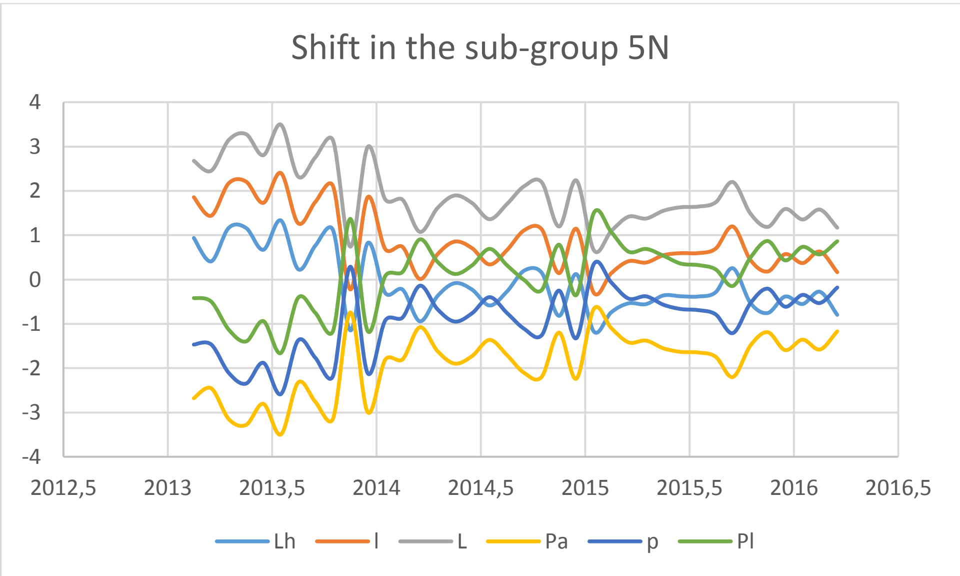
In Appendix 1 one we formulate 14 theorems and their proofs. By these theorems we may classify the index number formulas in to excellent indices, permanently and contingently biased indices. They also verify our classification of index numbers into three groups.

Shifts and biases get large values if also FA´s of our basic formulas are taken into account.

[7] or equivalently in respect to any other asymptotically unbiased or excellent formulas such as *t, D, Dr, SV …*



Especially $FA(Lh)$ and $FA(Pl)$ contain huge biases. If FA´s are dropped, the shifts and biases near the zero-line above become visual:



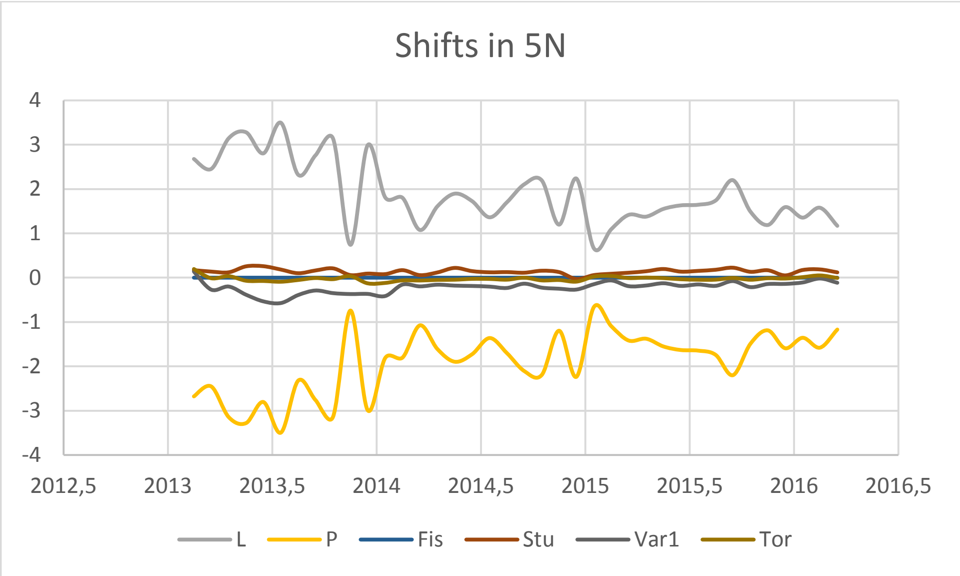
These figures show, that

A. the index forks are not stationary in time and

B. Fisher´s Five-Tined Fork does not appear as the only possibility here.

Fisher´s theory is clearly rejected.

Inside these bounds, we find all the excellent formulas, which have zero biases for SC although their shifts from *F* differ slightly from zero (because changes are not infinitesimal):



We stress that the normed deviation from Fisher, the limit of $Shift(P, F)$, either up or down is a clear fault of formula *P* if the changes in prices and quantities are small.

The shift formulas are related to the same denominator norm, which is calculated from four extreme members of the forks (*l* and *p* are not needed). It is a good practice always to calculate all the six basic indices (not just *L* or perhaps *L* and *P*). This allows calculating the shift and the many excellent indices like Törnqvist (and its derivatives), Fisher, Drobish, Stuvel etc. from them.

The two three-forks (L, l, Lh) and (Pl, p, Pa) provide a natural basis for all *quantitative* comparisons (such as $Shift(p, l)$ and $Shift(P, F)$) of index number formulas. Such data would increase the applications of official data, perhaps to surprising directions.

Fisher thought that biases of formulas lie within limits of (-2, +2). This is false. As shown above, biases of some FA-formulas can be large real numbers, less than -300 or more than 300.

8. Simple derivation of the three basic excellent index numbers

Table 1 in Appendix 3 shows that the TA formulas of the original formulas remain in the same set, but their order is reversed. This is essential: For instance, the highest index of the old fork, i.e. *L*, is mapped to the lowest index of the new fork or *Pa*. It is easier to grasp this algebraic structure in terms of geometry or in the picture, where both forks are presented. In our opinion, this is perhaps the best idea of Fisher. Especially its use in rectification of the formulas gives excellent results!

We will show, that the rectified formulas (in respect to its time antithesis TA) will lead to the same value for small variances of $\dot{p}$ and $\dot{q}$. This point served as the main point in the definition of *asymptotic unbiasedness* of index numbers, which is an exact concept for small *deviations* in $\dot{p}$ and $\dot{q}$, i.e. how much these log-changes *differ from some constants*. Their *magnitude* (i.e. the values of the price and quantity indices or constants) is totally irrelevant! For large deviations of $\dot{p}$ or large deviations of $\dot{q}$ (or both) the question of unbiasedness is left slightly ambiguous. This main message of Vartia (1978) and Vartia-Vartia (1984) was left unnoticed in index number literature. But it still clarifies many problems in index numbers.

The whole idea is the same as in feedback in dynamic systems (see Appendix 2, Table 1): a formula (1) = (2) is used first in the other direction (3), and then returned back to the original comparison (4) = (5). This is used as the correcting information for the original formula (1) = (2). This is the feedback connection used in *rectifying by the time antithesis.* We will show that *TA-rectified formula*

(14) $$\sqrt{f\begin{pmatrix} p^1 & q^1 \\ p^0 & q^0 \end{pmatrix} / f\begin{pmatrix} p^0 & q^0 \\ p^1 & q^1 \end{pmatrix}}$$

gives almost the same results (for small *deviations* in $\dot{p}$ = the log-changes of prices and $\dot{q} =$ the log-changes of quantities) for all our basic indices. This ´constant mean point´ serves as *the reference* for all formulas that are called *asymptotically unbiase*d, as defined in chapter 6.

For small deviations in $\dot{p}$ and $\dot{q}$ or regular moderate deviations, the old and new forks have (approximately) the same dispersions (= deviations or separations of the three tines) in the log-scale, as stated by Theorems 1-3 (see Appendix 1). We describe this asymptotic situation of small changes below and derive simply its main conclusions.

In the base fork, the three log-indices differ approximately by half the old price variance $\Delta^0 = \frac{1}{2}s_0^2(\dot{p})$:

$logL = a^0 + \Delta^0$	$TA(L) = Pa$
$logl = a^0$	$TA(l) = p$
$logLh = a^0 - \Delta^0$	$TA(Lh) = Pl$

In the new fork, the three log-indices differ approximately by half the new price variance $\Delta^1 = \frac{1}{2}s_1^2(\dot{p})$:

$logPl = a^1 + \Delta^1$
$logp = a^1$
$logPa = a^1 - \Delta^1$

For small deviations in $\dot{p}$ and $\dot{q}$ the new and old variances of $\dot{p}$ will approach equality, $\Delta^1 = \Delta^0 = \Delta$, not only absolutely but also relatively. This will normally hold accurately already for moderate deviations, because variances are relatively independent from their weighting (both have the same squared deviations of log-prices). For the variances in $\dot{p}$ and $\dot{q}$ both approaching zero (which makes old and new value shares to approach each other), we have asymptotic equations where Δ clearly indicates 1 dose of bias measured from the central tine of the fork:

$logL = a^0 + \Delta$
$logl = a^0$
$logLh = a^0 - \Delta$

$logPl = a^1 + \Delta$
$logp = a^1$
$logPa = a^1 - \Delta$.

From these equations, we see e.g. that $\frac{1}{2}(logL + logPa) = logF = \frac{1}{2}(a^0 + a^1)$, which equals e.g. $\frac{1}{2}(logl + loggp) = logt = \frac{1}{2}(a^0 + a^1)$. We calculate the geometric averages of between an index and its time antitheses (or equivalently the arithmetic averages of their log´s) before equality $\Delta^1 = \Delta^0 = \Delta$ and get:

$\frac{1}{2}(logL + logPa) = logF = \frac{1}{2}(a^0 + a^1) + \frac{1}{2}(\Delta^0 - \Delta^1) \to \frac{1}{2}(a^0 + a^1)$
$\frac{1}{2}(logl + logp) = logt = \frac{1}{2}(a^0 + a^1)$
$\frac{1}{2}(logLh + logPl) = logD = \frac{1}{2}(a^0 + a^1) - \frac{1}{2}(\Delta^0 - \Delta^1) \to \frac{1}{2}(a^0 + a^1)$

when $\Delta^0 - \Delta^1$ and $\frac{\Delta^0 - \Delta^1}{\Delta}$ approach zero. This happens if both $\frac{1}{2}s^2(\dot{p}\,; w)$ and $\frac{1}{2}s^2(\dot{q}\,; w)$ approach zero, for old and new value shares. This is a differential geometric argument, which becomes exact mathematics by using the shifts between the formulas, as is done in Appendix 1.

The point is that all the three rectified indices attain (for small *deviations* in $\dot{p}$ and $\dot{q}$, the absolute *magnitude* of the log-changes is irrelevant) approximately the same value, and exactly so if the variances of $\dot{p}$ and $\dot{q}$ approach zero. The situation above describes what happens in the limit of almost the same proportional changes. This *common point* $\frac{1}{2}(a^0 + a^1)$ defines the exact point of indices, the average point of these six indices on the log-scale, which can and should be declared as *unbiased*. Or more accurately, as *unbiased for almost proportional changes* of prices *and* of quantities. We have proved the following theorem (see Appendix 1): For all values of $logl = a^0$ and $logp = a^1$, the indices (F, t, D) approach each other and the constant $\frac{1}{2}(a^0 + a^1)$ in the log-scale in the sense that their S*hifts* vanish when the variances of the log-changes in prices $\Delta = \frac{1}{2}s^2(\dot{p}\,; \overline{w})$ and in quantities $\Delta(q) = \frac{1}{2}s^2(\dot{q}\,; \overline{w})$ approach zero.

For large deviations, even these three precision formulas (or other asymptotically unbiased indices) start to scatter and one cannot say any more how the small biases start to distribute among them.

Here are only three TA-rectified new formulas, because e.g. the Törnqvist formula is attained by either rectifying of *l* or *p*. We do not present this as formula (or its simple algebra), because it is much better that the reader looks carefully the table 1 in Appendix 2 and clears to procedure to himself (or to herself).

For small variances in the log-changes of both prices and quantities, the three indices (F, t, D) have approximately the same value, irrespective where the three forks situate (a^0, a^1) and how they disperse Δ. All other averages (than these TA-rectifications) over two members from different three-forks are *always biased* up or down. As proved in Vartia-Vartia (1984), $\frac{1}{2}(logL + logPl) = \frac{1}{2}(a^0 + a^1) + \Delta$ and therefore both $\sqrt{L * PL}$ and $\frac{1}{2}(L + PL)$ have one dose of permanent bias up. Similarly, we get $\frac{1}{2}(logLh + logp) = \frac{1}{2}(a^0 + a^1) - \frac{1}{2}\Delta$ and thus $\sqrt{Lh * p}$ has a half dose of permanent bias down, etc.

Only the three indices (F, t, D) produced by TA-rectification are always unbiased and excellent of the ones we have discussed. Further research is needed to widen this important set of asymptotically unbiased index numbers, but these three indices have served (together with *St* and *MV*) to show where the precision formulas lie and how near each other.

This ends our discussion and correction of the Fisher´s partly experimental discussion of his concept of bias and unbiased index numbers. The main conclusion is that both Laspeyres *L* and Paasche *Pa* price indices are clearly (asymptotically and finitely) biased formulas, whose biases may be small only by accident.

As we have noted, the quantity side of data has much more variability than the price side. Therefore, we should routinely use the price indices as a starting point of the V = P*Q decomposition. For instance, start e.g. from *t* and use the pair $(t,\ CoF(t))$ leading to

(15) $$\frac{V^1}{V^0} = P * Q = t * CoF(t)\,.$$

This would give accurate results, instead of doing it the other way around, $(P, Q) = (\,CoF(t)\,, t)$, where Törnqvist formula is first applied for quantities and then its co-factor $CoF(t)$ is used as a price index. That would give useless results in 5S and poor results in 5N because the quantity indices vary too much.

These FA-indices are totally useless in 5S because of the large variances $s^2(\dot{q}\,;w)$ of the quantity log-changes $\dot{q}$. Their biases are really large real numbers, often outside ± 5000 doses of bias! Concretely, these large figures arise, because the FA-indices disperse so strongly. In many commodities, their quantity relatives approach infinity caused by division of nearly zero, and the index explodes[8].

Look at the figure in page 13.

On the other hand, the original six indices (calculated normally in the price-space) deviate only moderately from each other. Their graphs show that the old and new three-forks change abruptly places, especially in 2015-16.

Look at the figure in page 14.

In the set of normal commodities 5N, the price development is in total control.

Look at the figure in page 10.

9. Index number production for similar complete micro data

If we concentrate only on the precision formulas in 5S, the results are in good control. This will be most probably the case also in other similar detailed homogeneous price-quantity data sets.

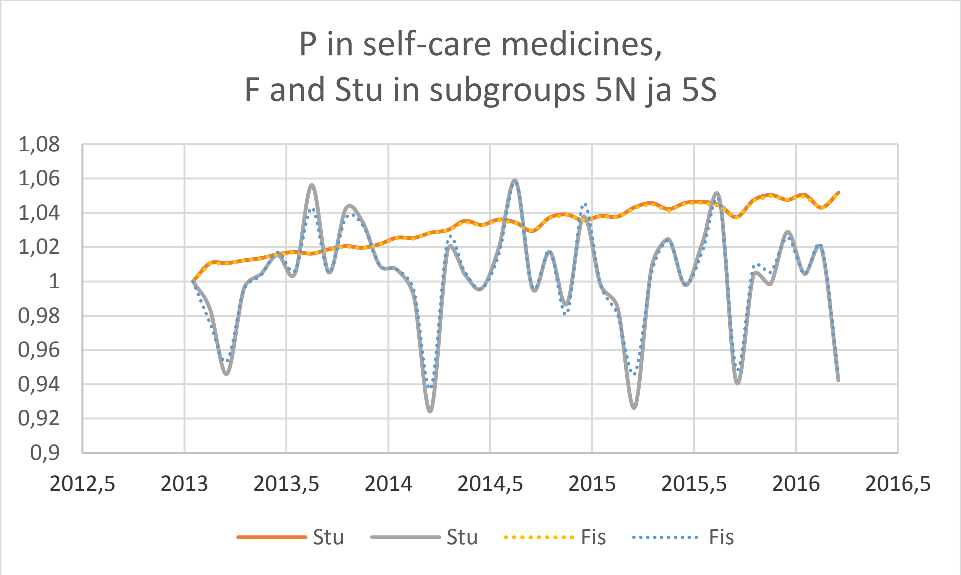
Look at the figure in page 14.

The most important feature in our treatment was the division of data into two, as the union of normal and special commodities, say 5N and 5S. Simple like that! In 5N all the index calculations are quite easy and traditional and it helped us to define the separation between e.g. contingently biased, freakish and excellent formulas. From the excellent formulas, we mentioned only a few: *F, t, St, D, MV*, *VW*. Some others will be introduced in the next paper (VarII = SV, TII, E, Wa, etc.), where also their merits and dis-merits will be discussed. A concrete application of these formulas in measuring welfare change in consumer theory is Vartia (1983).

We shall analyze carefully these formulas and drop away (eliminate) those which do not react correctly to extreme changes (up or down) in prices and quantities, a less analyzed property discussed in Vartia (1976, 1976b). That kind of changes occur all the time in 5S and they become even more pronounced elements in 5S when new and disappearing commodities are added to it.

We hope that the future research will show, that the following picture describes the actual problem very accurately. If so, we can be indifferent between choosing among the indices in $\{F, St\}$ in order to calculate the official price development in 5S and combine that with 5N using the same formula.

[8] Note also that graphs below do not intersect, although they could. Laspeyres and Paasche remain around 1 and even they do not intersect here. The asymmetry of the graphs follow from the third order term $m_3(\dot{p}, w)$ and of still higher order terms.



10. Conclusions

We analyze most typical index number formulas by the method of 'Fisher forks'. We define mathematically the shift measure by which we may classify the index number formulas in to four classes: excellent, permanently biased, contingently biased and freakish. We visualize some time quite difficult mathematics by empirical founding's and the graphs of them. In empirical analyze we use complete micro data including price, quantity and value information in all time periods.

Our data includes large deviations in quantity and/or price changes and naturally large variation in value shares. We show that only few index numbers can *tolerate* them. The situation is complicated in many ways, but logically arranged figures tell the story best. The concept of bias of the index number is very important issue of our study. It is well known, that there does not exist a simple numerical measure for bias or order of quality of an index number, because it depends on so many facts. For that we develop a new quantum theory that arranges the potential deviations of the indices from *logt* and *logF*. We detect that bad indices, like merely base or observation period weighted indices, get values here and there, but good indices go hand in hand. Choosing among the best formulas is more or less indifferent, because their values differ so little.

We analyze the data by classifying the commodities as normal ones, forming here the subset 5N, and special ones with exceptionally large relative changes in value. The latter formed the set of special commodities 5S in this paper, where individual values either increased 5-fold or declined to 1/5 of their old value. The subset 5S revealed simply and concretely, which index number formulas can be used in public information production. We are able to handle also the problems of new and disappearing commodities in a similar fashion.

In the forthcoming future CPI, the number of commodities can be increased by factors 50-100 and also all the quantities will be observed for many subgroups of commodities, based on scanner data on all transactions on the most detailed commodity level. This "Big Data" will remove most of the problems in the former production system, were only some 1500 commodities out of roughly 100 000 actual ones were monitored and quantities were not observed at the same time. Other problems of index number construction possibly remaining are of minor importance.

The accuracy and reliability of CPI will rise to a higher level, without much additional costs.

References

Balk, Bert (1995): Axiomatic index number theory: A survey. *International Statistical Review*, 63, 1, 63-93.
Cavallo, Alberto and Rigobon, Roberto (2016): The billion prices project: Using online prices for measurement and research, *Journal of Economic Perspectives, 30, Nro 2, Spring 2016, 151-178.*
Diewert, Erwin (1976): Exact and superlative index numbers, *Journal of econometrics,* 4, 115-145.
Diewert, Erwin (1978): Superlative index numbers and consistency in aggregation, *Econometrica*, 46, 883-900.
Fisher, Irving (1922): **The making of index numbers,** Houghton Mifflin Company, Boston.
Hansen, Bent and Lucas, Edward F. (1984): On the accuracy of index numbers, *Review of Income and Wealth, 1984,1, 25-38.*
Montgomery, J.K. (1937): **The mathematical problem of the price index,** P.S. King et Son Ltd., London
Pursiainen, Heikki (2005): **Consistent aggregation methods and index number theory,** Research Reports, Kansantaloustieteen laitoksen tutkimuksia (Research of the Department of Economics in Helsinki) No. 106:2005. Dissertationes Oeconomicae
Sato, Kazuo (1976): The ideal log-change index number, *The Review of Economics and Statistics,* 58, 223-8.
Stuvel, G. (1957): A new index number formula, *Econometrica*, 25, 123-131.
Stuvel, G. (1989): **The Index-number problem and its solution,** Basinstoke and London, Macmillan.
Suoperä, Antti and Vartia, Yrjö (2011): Analysis and synthesis of wage determination in heterogeneous gross sections, *HECER Discussion Papers No. 331, June 2011.*
Törnqvist, Leo (1936): The Bank of Finland´s consumption price index, *Bank of Finland Monthly Bulletin*, 10, 1-8.
Törnqvist, Leo, Vartia, Pentti and Vartia, Yrjö (1985): How should the relative changes be calculated, *American Statistician*, 39, 1, 43-46.
Vartia, Yrjö (1976): **Relative changes and index numbers,** Research Institute of the Finnish Economy ETLA, series A4 (dissertation in statistics).
Vartia, Yrjö (1976b): Ideal log-change index numbers, *Scandinavian Journal of Statistics*, 3, 121-126.
Vartia, Yrjö (1978): Fisher´s five-tined fork and other quantum theories of index numbers, *In Theory and applications of economic indices* (Ed. W. Eichhorn, R. Henn, O. Opitz, R. W. Shephard), Physica-Verlag, Wurzburg.
Vartia, Yrjö (1980): Compensated income in consumption-saving economies, Research reports No. 66, Research Institute of the Finnish Economy ETLA, 1-105.
Vartia, Yrjö (1983): Efficient methods of measuring welfare change and compensated income in terms of ordinary demand functions, *Econometrica*, 51,79-98.
Vartia, Yrjö (2008): On the aggregation of quadratic micro equations, *HECER Discussion Papers No. 248, December 2008.*
Vartia, Yrjö (2010): Principles of defining index numbers and constructing index series, *HECER Discussion Papers No. 283, January 2010.*
Vartia, Yrjö and Vartia, Pentti (1984): Descriptive index number theory and the Bank of Finland currency index, *Scandinavian Journal of Economics*, 86, 3, 352-364.
Vartia, Yrjö and Weymark, John A. (1982): Four revealed preference tables, *Scandinavian Journal of Economics,* 83, 408-418.

Appendix 1: Theorems on contingently biased, permanently biased and excellent index numbers

The relative weight bias depends on the difference of means $logp - logl = \sum(w_i^1 - w_i^0)\ddot{p}_i$ between the central tines of the new and old weighted forks. This can be written in several ways as an exact covariance between various variables. This rather surprising idea forms our first theorem.

Theorem 1: Consider the difference $logp - logl = \sum(w_i^1 - w_i^0)\ddot{p}_i = \sum \Delta w_i \Delta logp_i$. This can be expressed as several different (but numerically equal) covariances. Thus, identically for any data:

$$\begin{aligned} logp - logl &= \sum \frac{1}{n} n\Delta w_i \Delta logp_i = cov\left(n\Delta w, \dot{p}; \frac{1}{n}\right) && \text{even weights} \\ &= \sum w_i^0 \frac{\Delta w_i}{w_i^0} \Delta logp_i = cov\left(\frac{\Delta w}{w^0}, \dot{p}; w^0\right) && \text{old weights} \\ &= \sum \frac{1}{2}\left(w_i^0 + w_i^1\right)\left(\frac{\Delta w_i}{\overline{w}_i}\right)\Delta logp_i = cov\left(\frac{\Delta w}{\overline{w}}, \dot{p}; \overline{w}\right) && \text{Törnqvist weights} \\ &= \sum L(w_i^0, w_i^1) \frac{\Delta w_i}{L\left(w_i^0, w_i^1\right)} \Delta logp_i \\ &= S * \sum \frac{L\left(w_i^0, w_i^1\right)}{\sum L\left(w_i^0, w_i^1\right)} \Delta logw_i * \Delta logp_i \\ &= S * cov(\Delta logw, \dot{p}; \widehat{w}), \end{aligned}$$

where $S = \sum L(w_i^0, w_i^1) \leq 1$ and $\widehat{w}_i = \frac{L\left(w_i^0, w_i^1\right)}{\sum L\left(w_i^0, w_i^1\right)}$ = weights of Sato-Vartia index.

Proof: Covariance is defined as $cov(x, y; w) = \sum w_i(x_i - \bar{x})(y_i - \bar{y}) = \sum w_i(x_i - \bar{x})y_i$, where the weights w are non-negative real numbers summing to unity and e.g. $\bar{x} = \sum w_i x_i$ = w-weighted mean of the variable x. Thus, $\sum w_i(x_i - \bar{x}) = 0$ is a characteristic property of any covariance. It implies $\sum w_i(x_i - \bar{x})(y_i - \bar{y}) = \sum w_i(x_i - \bar{x})y_i - \sum w_i(x_i - \bar{x}) * \bar{y} = \sum w_i(x_i - \bar{x})y_i$, i.e. only one of the variables x and y needs to be expressed in the deviation form!

In all the expression above this zero-condition of the first variable applies, as in $\sum \frac{1}{2}\left(w_i^0 + w_i^1\right)\left(\frac{\Delta w_i}{\overline{w}_i}\right) = \sum \Delta w_i = 1 - 1 = 0$. This makes them all covariances of the denoted variables and weights.

Next we prove the following intuitive results concerning the old and new variances $s_0^2(\dot{p}) = s^2(\dot{p}, w^0)$ and $s_1^2(\dot{p}) = s^2(\dot{p}, w^1)$ and their approximations by $s^2(\dot{p}) = s^2(\dot{p}, \overline{w})$ with the Törnqvist weights. They all become equal, or more accurately, their ratios approach unity, when they themselves approach zero while log-changes in prices and quantities approach any proportional values. Their accurate mathematical expression is somewhat tricky:

Theorem 2: For all $(logP, logQ) \in R^2$: $\ddot{p} \to logP * 1_n$ *and* $\ddot{q} \to logQ * 1_n$ implies $\frac{s_1^2(\dot{p}) - s_0^2(\dot{p})}{s^2(\dot{p})} \to 0$.

The condition $\ddot{p} \to logP * 1_n$ means that for all commodities the log-change $\Delta logp_i = \ddot{p}_i$ approaches the same limit: $\ddot{p}_i = \Delta logp_i \to logP$, or the *log-deviation vector* $\ddot{p} - logP * 1_n$ approaches a zero vector. This refers to almost proportional changes APC.

Proof: $s_1^2(\dot{p}) - s_0^2(\dot{p}) = \sum w_i^1(\Delta logp_i - \Delta logp)^2 - \sum w_i^0(\Delta logp_i - \Delta logl)^2 = \sum(w_i^1 - w_i^0)(\Delta logp_i)^2 + (\Delta logp)^2 - (\Delta logl)^2 = \sum(w_i^1 - w_i^0)(\Delta logp_i - \Delta logt)^2$ by a straight calculation left to the reader as an exercise (expand, sum and note definitions of $\Delta logl$ and $\Delta logp$). $\left|\sum(w_i^1 - w_i^0)(\Delta logp_i - \Delta logt)^2\right| = \left|\sum \overline{w}_i\left(\frac{\Delta w_i}{\overline{w}_i}\right)(\Delta logp_i - \Delta logt)^2\right| \leq \sum \overline{w}_i \left|\frac{\Delta w_i}{\overline{w}_i}\right| (\Delta logp_i - \Delta logt)^2 \leq \sum \overline{w}_i max \left|\frac{\Delta w_j}{\overline{w}_j}\right| (\Delta logp_i - \Delta logt)^2 = \sum \overline{w}_i(\Delta logp_i - \Delta logt)^2 * max \left|\frac{\Delta w_j}{\overline{w}_j}\right| = s^2(\dot{p}) * max \left|\frac{\Delta w_j}{\overline{w}_j}\right|$. Thus $\left|\frac{s_1^2(\dot{p}) - s_0^2(\dot{p})}{s^2(\dot{p})}\right| \leq max \left|\frac{\Delta w_j}{\overline{w}_j}\right| = max \left|\frac{L\left(w_j^0, w_j^1\right)}{\overline{w}_j} * \frac{\Delta w_j}{L\left(w_j^0, w_j^1\right)}\right| = max \left|\frac{L\left(w_j^0, w_j^1\right)}{\overline{w}_j} *\right.$

$\Delta logw_j\Big| = max\left|\frac{L(w_j^0,w_j^1)}{\overline{w}_j}(\dot{p}_j+\dot{q}_j)\right| \le max\left|\frac{L(w_j^0,w_j^1)}{\overline{w}_j}\right| * max|(\dot{p}_j+\dot{q}_j)| \to 1*0=0$, when all the deviations $\dot{p}_i=\Delta logp_i-\Delta logt$ and $\dot{q}_i=\Delta logq_i-\Delta logt_q$ approach zero (implying $w_i^1 \to w_i^0$). This proves the theorem.

Theorem 3: For all $(logP, logQ) \in R^2$: $\ddot{p} \to logP*1_n$ and $\ddot{q} \to logQ*1_n$ implies $\frac{s_1^2(\dot{p})+s_0^2(\dot{p})}{2s^2(\dot{p})} \to 1$.

Proof: $\frac{1}{2}(s_1^2(\dot{p})+s_0^2(\dot{p})) = \frac{1}{2}\sum w_i^1(\Delta logp_i-\Delta logp)^2+\frac{1}{2}\sum w_i^0(\Delta logp_i-\Delta logl)^2$ $=\frac{1}{2}\sum(w_i^1+w_i^0)(\Delta logp_i)^2-(\Delta logp)^2-(\Delta logl)^2=\frac{1}{2}\sum(w_i^1+w_i^0)(\Delta logp_i-\Delta logt)^2+(\Delta logt)^2-(\Delta logp)^2-(\Delta logl)^2=s^2(\dot{p})-\frac{1}{4}(logp-logl)^2=s^2(\dot{p})-\frac{1}{4}\left(cov\left(\frac{\Delta w}{\overline{w}},\dot{p};\overline{w}\right)\right)^2$. Thus $\left|\frac{s_1^2(\dot{p})+s_0^2(\dot{p})}{2s^2(\dot{p})}\right| = \left|1-\frac{1}{4s^2(\dot{p})}\left(cov\left(\frac{\Delta w}{\overline{w}},\dot{p};\overline{w}\right)\right)^2\right| \le 1+\left|\frac{1}{4}\left(s^2\left(\frac{\Delta w}{\overline{w}};\overline{w}\right)r^2\left(\frac{\Delta w}{\overline{w}},\dot{p};\overline{w}\right)\right)\right| \le \frac{1}{4}s^2\left(\frac{\Delta w}{\overline{w}};\overline{w}\right) \to 0$ when all the deviations $\dot{p}_i=\Delta logp_i-\Delta logt$ and $\dot{q}_i=\Delta logq_i-\Delta logt_q$ approach zero (implying $w_i^1 \to w_i^0$ for all i). This proves the theorem.

Theorem 4: For all $(logP, logQ) \in R^2$: $\ddot{p} \to logP*1_n$ and $\ddot{q} \to logQ*1_n$ implies

$$\frac{s_1^2(\dot{p})}{s^2(\dot{p})} \to 1 \text{ and } \frac{s_0^2(\dot{p})}{s_1^2(\dot{p})} \to 1.$$

Proof: Insert identity $s_1^2(\dot{p})=\frac{1}{2}\left(s_1^2(\dot{p})+s_0^2(\dot{p})\right)+\frac{1}{2}(s_1^2(\dot{p})-s_0^2(\dot{p}))$ into $\frac{s_1^2(\dot{p})}{s^2(\dot{p})}=\frac{s_1^2(\dot{p})+s_0^2(\dot{p})}{2s^2(\dot{p})}+\frac{s_1^2(\dot{p})-s_0^2(\dot{p})}{2s^2(\dot{p})}$. By theorems 2 and 3, this approaches 1+0 = 1 for APC, which proves $\frac{s_1^2(\dot{p})}{s^2(\dot{p})} \to 1$. Similarly $\frac{s_0^2(\dot{p})}{s^2(\dot{p})} \to 1$.

Now we are ready to state and prove the main result that Törnqvist t and Fisher F approximate each other quadratically for almost proportional changes and thus for all small changes. The proof is short and elegant:

Theorem 5: For all $(logP, logQ) \in R^2$: $\ddot{p} \to logP*1_n$ and $\ddot{q} \to logQ*1_n$ the limes $lim_{APC}Shift(t,F)$ exists and equals zero. Thus, Törnqvist t is excellent for almost proportional changes. Therefore, it is also unbiased for small changes or $Bias(t)=lim_{SC}Shift(t,F)=0$.

Proof: By the properties of moment means

A $\quad logPa-logp=-\frac{1}{2}s_1^2(\dot{p})+o_2$ and $logL-logl=\frac{1}{2}s_0^2(\dot{p})+o_2$,

where $o_2=o_2(\dot{p},\dot{q})$ means *any* expression, which goes to zero faster than the variance (or sum of squares) of $\dot{p}$ when both $(\dot{p},\dot{q})$ approach zero, i.e. $s^2(\dot{p}) \to 0$ and $s^2(\dot{q}) \to 0$. Thus $o_2/s^2(\dot{p}) \to 0$ when the denominator and $s^2(\dot{q})$ approach zero. Summing equations A, and dividing by two gives $logF-logt=-\frac{1}{4}(s_1^2(\dot{p})-s_0^2(\dot{p}))+o_2$. $Shift(t,F)=\frac{2\log\left(\frac{t}{F}\right)}{s^2(\dot{p},\overline{w})}=\frac{1}{2}\frac{(s_1^2(\dot{p})-s_0^2(\dot{p}))}{s^2(\dot{p},\overline{w})}+\frac{o_2}{s^2(\dot{p},\overline{w})}$. Now by Theorem 1, $lim_{APC}Shift(t,F)=0+0=0$.

Theorem 6: The moment means $M^{(\alpha)}(p^{1/0},\overline{w})$ with Törnqvist weights are **permanently biased** for all $\alpha \ne 0$. This is an infinite set (continuum) of permanently biased indices. The permanent bias α is small when it is nearly zero, which equals the excellent Törnqvist index case.

Proof: Apply (4) for Törnqvist weights and get $log\frac{M^{(\alpha)}(p^{1/0},\overline{w})}{t}=\alpha\, s^2(\dot{p},\overline{w})+o_2$ and $Shift\left(M^{(\alpha)}(p^{1/0},\overline{w}),t\right)=Shift\left(M^{(\alpha)}(p^{1/0},\overline{w}),F\right)+Shift(F,t)=\alpha+Shift(F,t)+o_2/s^2(\dot{p},\overline{w}) \to \alpha$, by theorem 5 when both $(\dot{p},\dot{q})$ approach zero. Thus, its limit exists and equals a constant α, which is the value of the permanent bias. This proves the theorem.

Theorem 7: $\log\left(\frac{p}{l}\right) = cov(\dot{v},\dot{p};\overline{w})+o_2 = cov(\dot{p},\dot{q};\overline{w})+s^2(\dot{p},\overline{w})+o_2$ and therefore, e.g.

$Shift(p,t) \rightarrow Shift\left(\frac{p}{F}\right) \rightarrow \frac{cov(\dot{v},\dot{p};\bar{w})}{s^2(\dot{p},\bar{w})}$ when deviations in price and quantity log-changes approach zero or APC.

Proof: Recall the basic properties of the functions cosh and sinh:

$\cosh(z) = \frac{1}{2}(\exp(z) + \exp(-z)) = 1 + \frac{1}{2!}z^2 + \frac{1}{4!}z^4 + \cdots$ and
$\sinh(z) = \frac{1}{2}(\exp(z) - \exp(-z)) = z + \frac{1}{3!}z^3 + \frac{1}{5!}z^5 + \cdots$

First $\bar{w}/\sqrt{w^0 w^1} = \frac{1}{2}\left(\sqrt{\frac{w^1}{w^0}} + \sqrt{\frac{w^0}{w^1}}\right) = \frac{1}{2}\left(\exp\left(\frac{1}{2}\log\left(\frac{w^1}{w^0}\right)\right) + \exp\left(-\frac{1}{2}\log\left(\frac{w^1}{w^0}\right)\right)\right) = \frac{1}{2}(\exp(z) + \exp(-z))$
$= \cosh(z)$ for $z = \frac{1}{2}\log\left(\frac{w^1}{w^0}\right) = \frac{1}{2}\Delta\log w$. Thus $\frac{\bar{w}}{\sqrt{w^0 w^1}} = 1 + \frac{1}{8}log\left(\frac{w^1}{w^0}\right)^2 + \cdots = 1 + \frac{1}{8}\ddot{w}^2 + o_2$.

Similarly $\frac{L(w^0,w^1)}{\sqrt{w^0 w^1}} = L\left(\sqrt{\frac{w^1}{w^0}}, \sqrt{\frac{w^0}{w^1}}\right) = L\left(\exp\left(\frac{1}{2}\log\left(\frac{w^1}{w^0}\right)\right), \exp\left(-\frac{1}{2}\log\left(\frac{w^1}{w^0}\right)\right)\right) = L\big(\exp(z), \exp(-z)\big) = \sinh(z)/z$. Thus
$\frac{L(w^0,w^1)}{\sqrt{w^0 w^1}} = 1 + \frac{1}{3!}z^2 + \cdots = 1 + \frac{1}{24}\ddot{w}^2 + o_2$ and $\frac{L(w^0,w^1)}{\bar{w}} = \frac{L(w^0,w^1)}{\sqrt{w^0 w^1}}\frac{\sqrt{w^0 w^1}}{\bar{w}} = \frac{L(w^0,w^1)}{\sqrt{w^0 w^1}} / \frac{\bar{w}}{\sqrt{w^0 w^1}}$.

Therefore $\frac{L(w^0,w^1)}{\bar{w}} = 1 + \frac{1}{24}\ddot{w}^2 - \frac{1}{8}\ddot{w}^2 + \cdots = 1 - \frac{1}{12}\ddot{w}^2 + \cdots$ and by theorem 1 $logp - logl = \sum\frac{1}{2}\left(w_i^0 + w_i^1\right)\left(\frac{\Delta w_i}{\bar{w}_i}\right)\dot{p}_i = \sum\frac{1}{2}\left(w_i^0 + w_i^1\right)\left(\frac{\Delta w_i}{L(w_i^0,w_i^1)}\right)\frac{L(w_i^0,w_i^1)}{\bar{w}_i}\dot{p}_i = \sum\frac{1}{2}\left(w_i^0 + w_i^1\right)\Delta log w_i(1 + \frac{1}{12}\ddot{w}_i^2)\dot{p}_i + o_2 = \sum\frac{1}{2}\left(w_i^0 + w_i^1\right)\Delta log v_i\,\dot{p}_i + o_2 = cov(\dot{v},\dot{p};\bar{w}) + o_2$. Insert this in $Shift(p,t) = \frac{\log\left(\frac{p}{t}\right)}{\frac{1}{2}s^2(\dot{p},\bar{w})} = \frac{\frac{1}{2}\log\left(\frac{p}{l}\right)}{\frac{1}{2}s^2(\dot{p},\bar{w})} = \frac{cov(\dot{v},\dot{p};\bar{w})}{s^2(\dot{p},\bar{w})} + \frac{o_2}{s^2(\dot{p},\bar{w})}$. Now let deviations in price and quantity log-changes approach zero without altering their essential mutual dependencies, such as correlation and the ratio of deviations, see p. 16. Then $Shift(p,t) \rightarrow \frac{cov(\dot{v},\dot{p};\bar{w})}{s^2(\dot{p},\bar{w})}$, which is not a unique real number, but a contingent data-dependent function. This proves that a unique limit does not exist and *p* is contingently biased in respect to *t* or equivalently in respect to Fisher. This proves the theorem for APC and thus for SC.

.

Theorem 8: $\log\left(\frac{Pa}{L}\right) = cov(\dot{p},\dot{q};\bar{w}) + o_2$ and therefore, e.g. $Shift(Pa,F) \rightarrow \frac{cov(\dot{p},\dot{q};\bar{w})}{s^2(\dot{p},\bar{w})} = \frac{cov(\dot{p},\dot{q};w^0)}{s^2(\dot{p},w^0)}$ when deviations in price and quantity log-changes approach zero or APC.

Proof: Subtracting equations A gives by theorems 2 and 6: $logPa - logL = logp - logl - \frac{1}{2}(s_1^2(\dot{p}) + s_0^2(\dot{p})) + o_2 = cov(\dot{v},\dot{p};\bar{w}) - s^2(\dot{p},\bar{w}) + o_2 = cov(\dot{p},\dot{q};\bar{w}) + o_2$. The $Shift(Pa,F)$ and its limit behavior follows just like in theorem 7.

Analogically we prove the similar result for Laspeyres:

Theorem 9: $\log\left(\frac{Pa}{L}\right) = cov(\dot{p},\dot{q};\bar{w}) + o_2$ and therefore, e.g. $Shift(L,F) \rightarrow -\frac{cov(\dot{p},\dot{q};\bar{w})}{s^2(\dot{p},\bar{w})} = \frac{cov(\dot{p},\dot{q};w^0)}{s^2(\dot{p},w^0)}$ when deviations in price and quantity log-changes approach zero or APC.

Theorem 10: Drobish $Dr = \frac{1}{2}(L + Pa)$ is excellent for APC and thus for SC. It has a very appealing basket interpretation $Dr = \frac{p^1\cdot(q^0+q^1/L(q))}{p^0\cdot(q^0+q^1/L(q))}$ in contrast to Fisher. It is an example of a excellent index fulfilling neither Time Reversal nor Quantity Reversal Tests.

Proof: $\frac{Dr}{F} = \frac{1}{2}\left(\frac{L}{F} + \frac{Pa}{F}\right) = \frac{1}{2}\left(\sqrt{\frac{L}{Pa}} + \sqrt{\frac{Pa}{L}}\right) = \frac{1}{2}\left(\exp\left(\frac{1}{2}log\frac{L}{Pa}\right) + \exp\left(-\frac{1}{2}log\frac{L}{Pa}\right)\right) = \frac{1}{2}(\exp(z) + \exp(-z))$
$= \cosh(z) = 1 + \frac{1}{2!}z^2 + \frac{1}{4!}z^4 + \cdots = 1 + \frac{1}{8}(log\frac{L}{Pa})^2 + o_2$ as in theorem 7.
$\log\left(\frac{Dr}{F}\right) = \log\left(1 + \frac{1}{8}(log\frac{L}{Pa})^2 + o_2\right) = \frac{1}{8}(log\frac{L}{Pa})^2 + o_2 = \frac{1}{8}cov(\dot{p},\dot{q})^2 + o_2$ by theorem 8.

$Shift(Dr,F) = \frac{2\log\left(\frac{Dr}{F}\right)}{s^2(\dot{p},\bar{w})} = \frac{1}{4}\frac{cov(\dot{p},\dot{q},\bar{w})^2+o_2}{s^2(\dot{p},\bar{w})} = \frac{1}{4}\frac{(s(\dot{p})s(\dot{q})r(\dot{p},\dot{q}))^2+o_2}{s^2(\dot{p})} = \frac{1}{4}s^2(\dot{q})r(\dot{p},\dot{q})^2 + \frac{o_2}{s^2(\dot{p})}$.

This approaches zero when for all $(logP, logQ) \in R^2$: $\ddot{p} \to logP * 1_n\ and\ \ddot{q} \to logQ * 1_n$ or when $s^2(\dot{p})$ and $s^2(\dot{q})$ both go to zero. We have proved that $lim_{APC}Bias(Dr\ ,F)$ exists and equals zero. This means, that Drobish is unbiased for APC and thus for SC.

Theorem 11: Similarly constructed logarithmic mean $L(L, Pa)$ of Laspeyres and Paasche, which lies between Dr and F (and is thus even nearer to F than Dr) is excellent for APC and thus for SC. It does not have a basket interpretation and does not fulfill Time Reversal or Quantity Reversal Tests.

Theorem 12: Any evenly weighted moment mean $M^{(\alpha)}(L, Pa)$ or $M^{(\alpha)}(l, p)$ with parameter $\alpha \in [0,1]$ is excellent for APC and thus for SC. For the parameter α smaller than 1 (= the arithmetic mean), the moment mean is nearer to F and t. This produces an infinite number of excellent indices.

Theorem 13: "Vartia-Walsh" $logVW = \sum \sqrt{w_i^0\, w_i^1}\ \Delta logp_i$ is excellent for SC but not for APC.

Proof: From theorem 6

$\frac{\sqrt{w^0 w^1}}{\bar{w}} = 1 - \frac{1}{8}\ddot{w}^2 + o_2$ and $\frac{1}{8}\ddot{w}^2 = 1 - \frac{\sqrt{w^0 w^1}}{\bar{w}} + o_2$. For simplicity, we drop the commodity sub-indices.

$$\log(t) - log\mathrm{VW} = \sum \bar{w}\Delta logp - \sum \sqrt{w^0 w^1}\ \Delta logp$$

$$= \sum \left(\bar{w} - \sqrt{w^0 w^1}\right)\Delta logp = \sum \bar{w}\left(1 - \frac{\sqrt{w^0 w^1}}{\bar{w}}\right)\Delta logp = -\frac{1}{8}\sum \bar{w}\ddot{w}^2\Delta logp + o_2$$

Here the dotted variables are log-deviations from mean, not mere log-changes. This difference causes first difficulties. $Shift(\mathrm{VW}, t) = = \frac{-\frac{1}{8}\sum \bar{w}\ \ddot{w}^2\Delta logp + o_2}{\frac{1}{2}s^2(\dot{p})}$ and $-4Shift(\mathrm{VW}, t) = \frac{\sum \bar{w}\ \ddot{w}^2\Delta logp + o_2}{s^2(\dot{p})}$.

$\sum \bar{w}\ \ddot{w}^2\ = \sum \bar{w}(\ddot{v} - log\frac{V^1}{V^0})^2\ = \sum \bar{w}\ \dot{v}^2\ + o_2 = var(\ddot{v}; \bar{w}) + o_2 = \sum \bar{w}\ (\ddot{p}^2 + 2\ddot{p}\ddot{q} + \ddot{q}^2) + o_2 = var(\ddot{p}) + 2cov(\ddot{p}, \ddot{q}) + var(\ddot{q}) + o_2$. Here we could equally well have single dots instead of double dots, because e.g. $var(\ddot{p}) = var(\dot{p})$. By triangle inequality $\left|\frac{\sum \bar{w}\ \ddot{w}^2\Delta logp + o_2}{s^2(\dot{p})}\right| \le \frac{\sum \bar{w}\ \ddot{w}^2|\Delta logp| + |o_2|}{s^2(\dot{p})} \le \frac{(\sum \bar{w}\ \ddot{w}^2)max|\Delta logp| + o_2}{s^2(\dot{p})} \le \left(1 + 2\frac{s(\dot{q})}{s(\dot{p})} + \frac{s^2(\dot{q})}{s^2(\dot{p})}\right)max|\Delta logp| + \frac{o_2}{s^2(\dot{p})}$. Suppose that that the log-changes approach zero so that $\frac{s^2(\dot{q})}{s^2(\dot{p})} \le K$, say ≤ 400. Then $\left|\frac{\sum \bar{w}\ \dot{w}^2\Delta logp + o_2}{s^2(\dot{p})}\right| \le (1 + 2\sqrt{K} + K)max|\Delta logp| + \frac{o_2}{s^2(\dot{p})}$, which approaches zero when all changes of log-prices and log-quantities approach zero. Because this holds for all choices of K, it does not restrict the limit processes in any way and $Lim_{SC}Shift(\mathrm{VW}, t) = Lim_{SC}Shift(\mathrm{VW}, F) = 0$.

Corollary: "Vartia-Walsh" $logVW = \sum \sqrt{w_i^0\, w_i^1}\ \Delta logp_i$ is probably the simplest excellent index for SC which is consistent in aggregation CA. Its weights sum at most to unity.

Theorem 14: Montgomery-Vartia index $logMV = \sum \widehat{w}_i\ \Delta logp_i, \widehat{w}_i = L(v_i^0, v_i^1)/L(V^o, V^1)$ is excellent for SC but not for APC.

Proof: By similar argumentation as above $Shift(\mathrm{MV}, t) = = \frac{-\frac{1}{24}\sum \bar{w}\ \dot{w}^2\Delta logp + o_2}{\frac{1}{2}s^2(\dot{p})}$ from which the claim follows similarly. It is nearer to Törnqvist than "Vartia-Walsh".

Appendix 2: Fisher´s five-tined fork finally corrected

All these 12 – 2 = 10 indices (L and Pa appear twice in 12) can be estimated (with increasing accuracy for progressively smaller changes) using only four parameters:

$m = logt,\ d = \log\left(\frac{p}{l}\right) \cong cov(\dot{v},\ \dot{p};\ \overline{w}), \Delta(p) = cov(\dot{p},\ \dot{p};\ \overline{w})/2, \Delta(q) = cov(\dot{q},\ \dot{q};\ \overline{w})/2.$

We have recursively $logp = m + d/2$, $logl = m - d/2$, $logPl = logp + \Delta(p)$, $logPa = logp - \Delta(p)$, $logL = logl + \Delta(p)$, $logLh = logl - \Delta(p)$.

FA´s differ from each other by $\Delta(q)$: $logFA(L) = logPa$, $logFA(Pa) = logL$, $logFA(l) = logPa + \Delta(q)$, $logFA(p) = logL - \Delta(q)$, $logFA(Lh) = logPa + 2\Delta(q)$, , $logFA(Pl) = logL - 2\Delta(q)$.

All these indices are clearly biased, more or less depending on the situation. They all can be corrected for bias by TA-rectification, which implies that none of them should be ever used (in case of complete micro data). This is the starting point of our new "quantum theory of index numbers" for six base indices and their FA´s, which replaces the false "Fisher´s Five-Tined Fork". It must be completed by including "the center of gravity" $m = logt$ and other new points corresponding to derivative indices. It is a rather complicated set of points and gets different forms for different values of its four parameters $(m, d, \Delta(p), \Delta(q))$. It reduces to Fisher´s Five-Tined Fork for the two parametric case $(m, d = 2\Delta(p) = 2\Delta(q))$.

Appendix 3: TA and FA for the basic six formulas

Table 1: Time Antithesis formulas *TA* for the basic six formulas.

(1)	(2)	(3)	(4)	(5)
	$P^{1/0}$	$P^{0/1}$	$1/\ P^{0/1}$	
Symbol of the formula	$f\begin{pmatrix}p^1 & q^1\\ p^0 & q^0\end{pmatrix}$	$f\begin{pmatrix}p^0 & q^0\\ p^1 & q^1\end{pmatrix}$	$1/f\begin{pmatrix}p^0 & q^0\\ p^1 & q^1\end{pmatrix}$	Symbol of the TA formula
L	$\frac{p^1\cdot q^0}{p^0\cdot q^0}$	$\frac{p^0\cdot q^1}{p^1\cdot q^1}$	$\frac{p^1\cdot q^1}{p^0\cdot q^1}$	$Pa = TA(L)$
l	$\prod(\frac{p_i^1}{p_i^0})^{w_i^0}$	$\prod(\frac{p_i^0}{p_i^1})^{w_i^1}$	$\prod(\frac{p_i^1}{p_i^0})^{w_i^1}$	$p = TA(l)$
Lh	$1/\sum w_i^0 p_i^{0/1}$	$1/\sum w_i^1 p_i^{1/0}$	$\sum w_i^1 p_i^{1/0}$	$Pl = TA(Lh)$
Pl	$\sum w_i^1 p_i^{1/0}$	$\sum w_i^0 p_i^{0/1}$	$1/\sum w_i^0 p_i^{0/1}$	$Lh = TA(Pl)$
p	$\prod(\frac{p_i^1}{p_i^0})^{w_i^1}$	$\prod(\frac{p_i^0}{p_i^1})^{w_i^0}$	$\prod(\frac{p_i^1}{p_i^0})^{w_i^0}$	$l = TA(p)$
Pa	$\frac{p^1\cdot q^1}{p^0\cdot q^1}$	$\frac{p^0\cdot q^0}{p^1\cdot q^0}$	$\frac{p^1\cdot q^0}{p^0\cdot q^0}$	$L = TA(Pa)$

Table 2: Factor Antithesis formulas *FA* for the basic six formulas.

(1)	(2)	(3)	(4)	(5)
	$P^{1/0}$	$P^{1/0}(q)$	$CoF(\ P^{1/0}(q))$	
Symbol of the formula	$f\begin{pmatrix}p^1 & q^1\\ p^0 & q^0\end{pmatrix}$	$f\begin{pmatrix}q^1 & p^1\\ q^0 & p^0\end{pmatrix}$	$V^{1/0}/f\begin{pmatrix}q^1 & p^1\\ q^0 & p^0\end{pmatrix}$	Symbol of the FA formula
L	$\frac{p^1 \cdot q^0}{p^0 \cdot q^0}$	$\frac{q^1 \cdot p^0}{q^0 \cdot p^0}$	$\frac{p^1 \cdot q^1}{p^0 \cdot q^1}$	$Pa = FA(L)$
l	$\prod (\frac{p_i^1}{p_i^0})^{w_i^0}$	$\prod (\frac{q_i^1}{q_i^0})^{w_i^0}$	$V^{1/0}/\prod (\frac{q_i^1}{q_i^0})^{w_i^0}$	$FA(l)$
Lh	$1/\sum w_i^0 p_i^{0/1}$	$1/\sum w_i^0 q_i^{0/1}$	$V^{1/0}/\sum w_i^0 q_i^{0/1}$	$FA(Lh)$
Pl	$\sum w_i^1 p_i^{1/0}$	$\sum w_i^1 q_i^{1/0}$	$V^{1/0}/\sum w_i^1 q_i^{1/0}$	$FA(Pl)$
p	$\prod (\frac{p_i^1}{p_i^0})^{w_i^1}$	$\prod (\frac{q_i^1}{q_i^0})^{w_i^1}$	$V^{1/0}/\prod (\frac{q_i^1}{q_i^0})^{w_i^1}$	$FA(p)$
Pa	$\frac{p^1 \cdot q^1}{p^0 \cdot q^1}$	$\frac{q^1 \cdot p^1}{q^0 \cdot p^1}$	$\frac{p^1 \cdot q^0}{p^0 \cdot q^0}$	$L = FA(Pa)$