Comparing Basic Averages, Index Numbers and Hedonic Methods as Price Change Statistic

Suoperä, Antti, Statistics Finland
Nieminen, Kristiina, Statistics Finland
Montonen, Satu, Statistics Finland
Markkanen, Hannele, Statistics Finland
2020

Abstract

Measurement of price change and hedonic price index numbers are based on averages. As NSIs obtain scanner data having varying content, question arises: what is the most correct average to use?

We investigate differences of two price change estimates. We show that the differences can be expressed by covariances between weight differences of given two 'averages' and changes of log-prices. All analysis is done in logarithmic scale. This way these 'averages' can be compared pair wisely in order to show their relation.

First, we show that excellent index numbers are practically independent of weighting. This means that any of the excellent index number formula may be selected as benchmark. Then we calculate differences of any given average with respect to our benchmark estimate. If the covariances and especially their differences are always close to zero, then the analyzed average is appropriate for estimating the price change. Our previous tests have revealed that quantities consumed may change even more than tenfolds between compared periods; thus causing serious bias for price change estimate based on some subgroup averages. We show that when quantities change substantially and not proportionally, it introduces a significant bias in the index calculations. Therefore finally, we test a hedonic method in order to remove quantity changes from the actual log-price change separately for the basic averages. The statistical analysis of our data is based on graphical presentations and mathematical analysis is presented mostly in our earlier papers.

1 Introduction

Current state

Starting from year 2017, Statistics Finland has introduced new data sources to the production of CPI and HICP. These new data sources are based on either transaction (consumer purchases of daily products and pharmaceutical products recorded at point of sales) or on operative data covering service sales, such as charges of mobile phone subscriptions and internet broadband. Introduction of the data regarding daily products have been quite straightforward since one of our main principles, matching-pair-comparison works well. All products are identified with the unique identifier, GTIN-code or VNR-code. Thus, one may follow identified product for longer periods, for example several years and quality of product remains static.

Challenges and aim of the study

As we have already acquired quite much practice with processing of scanner datasets and now we want to proceed to new problems. Hence our aim is to replace some of the existing parts of traditional field collection with new groups of commodities like durables and consumer goods. How to perform this as we face new problems in the obtained scanner datasets? Below is a list of common challenges in datasets:

- *scanner datasets may have differing set of information.* Some scanner datasets may be called as complete datasets containing all needed information (prices, values, quantities, features, register information, categories) for various types of index calculations. Others contain limited information having either sales value with quantities but without characteristics, or only prices without any other information.
- *commodities have differing characteristics*, meaning that in some cases product quality remains quite stable as time goes on, while some products and single items “age” faster that others due to the technical development or fashion. Examples of these are eggs (stable), women’s clothes (fashion) and mobile phones (technical development). New products having more advanced technical features replace older models from markets even though the older products are still usable.

We also have *various methods available for the calculation of price change.* Selection of method depends on the content of scanner-data and characteristics of commodities. If obtained scanner-datasets contain all needed information, we may follow the traditional matching-pairs principle and calculate price-ratios by uniquely identified products. When we recognize higher churn of product items, faster aging of products or clear improvement of product quality, it means pair-matching is not possible anymore. In these situations, we have to select another method like hedonic model for quality control or averages for product groups.

In this paper, our aim is to focus on these alternative methods that could be used for calculating pure price-ratios. In the following tests, we use so called complete dataset containing monthly observations of alcoholic beverages: their unit prices, sold quantities, some product features and retail-specific categories. By using complete dataset in test calculations, it enables adoption of alternative methods.

The design of study

All index number formulas and price change statistics have their most popular representations. Unfortunately, comparison between them is quite hard task for most statisticians. Normally the relationships cannot be seen directly (for example index numbers compared to statistics based on basic averages i.e. elementary aggregates). How could we compare different price change estimates and select the best one? Experts suggest different methods: Balk (1990) suggest to use axiomatic index number theory, Diewert (1976, 1978) superlative index numbers, Vartia (1976a, 1976b, 1978) ideal log-change index numbers, Pursiainen (2005) consistent aggregation and index numbers, Vartia and Suoperä (2017, 2018) excellent index numbers. It is time to simplify confusing issue by generalizing the problem.

So, we present most known index numbers and other price change statistics by their logarithmic representations of averages. Logarithmic arithmetic averages were presented in paper Suoperä (2006), Vartia and Suoperä and Vuorio (2019), Laspeyres and Paasche in paper Vartia (1976) and even Fisher may be presented in logarithmic form. This means that as these logarithmic representations are available, we can then calculate for example 'Törnqvist averages'[1], 'Montgomery-Vartia averages', weighted and unweighted arithmetic and geometric averages in logarithmic form. Finally, we calculate also pair-wise differences between them to get information how they deviate and are related with each other.

In this study, we show empirically different averages in logarithmic form, their standard deviations and especially their covariances. We also show that the hedonic quality correction for quantities consumed do not help to remove serious bias for price change estimate based on some subgroup averages.

The implementation of study

There is no rule how to weight price observations for base and observation periods, so we trust in index number theory. We divide our problem into four steps and proceed as follows:

1. we compare six excellent index numbers and select a benchmark index
2. we calculate basic index numbers and compare with benchmark
3. we calculate basic averages, weighted or not, and compare with the benchmark index
4. we calculate quality adjusted averages and compare with benchmark.

The structure of the study is as follows: In chapter two we present basic concepts, notations, description of the data and show logarithmic representations for index number formulas and for basic averages. In chapter three we compare different index number formulas and basic averages or 'quality adjusted averages' derived with hedonic method. This study is based on some important equations but most often their graphical presentations.

2 Notations, Test Data, Index Number Formulas and Price Aggregation of Averages

2.1 Notations

Our notation in this study uses familiar notations used in index number calculations, that is

Commodities: $a_1, a_2, \dots, a_{n_t}$ and the number n_t in this study is about 3300 - 4100.
Time periods: $t = 0, 1, 2, \dots$ are the compared situations and are months (years 2016 – 2019).
Quantity: q_i^t is the quantity of a_i in period *t*.
Value: v_i^t is the value (expenditure) of a_i in period *t*.
Unit value: $p_i^t = v_i^t / q_i^t$ is the unit price of a_i in period *t*.
Total value: $V^t = \sum_i v_i^t$ is the total value (expenditure) of all the commodities in period *t*.
Total quantity: $Q^t = \sum_i q_i^t$ is the total quantity of all the commodities in period *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 $\mathrm{a_i}$ in period t.
Price index number: $P_f^{t/0}$ is the price index number for price-link 0 → *t* and formula *f*.
Quantity index number: $Q_f^{t/0}$ is the quantity index number for price-link 0 → *t* and formula *f*.

[1] With the term "average" we highlight that also index numbers belongs to the category of averages.

2.2 Description of Test Data

The following Table 1 demonstrates the content of typical *complete micro data*. It presents one subgroup *k* including *n* commodities, alcoholic beverages, that are comparable in quality in time. The base period is the previous year normalized as an average month. An observation in scanner-data contains unit prices (p_i^t, euro/liters), quantities (q_i^t, liters) and expenditures (v_i^t, euro) for the base period (i.e. $t = 0$) and for the observation month of a current year(i.e. $t = 1$) .

Commodities are arranged based on the information content: disappearing commodities are put to category A, new to C and all other that are constantly in sale to B. Commodities in group A and C are called as '*one sided null values*'. Most often "complete" scanner-datasets that NSI's obtain have one-sided null-values – the life cycle of the commodities decides what is relative share of expenditures for commodities belonging into categories A and C. As a warning, categories A and C have nothing to do with quality change but are 'vanishing commodities forward or backward in time'. Bilateral price-links can be calculated for group B and hedonic method is applied for data where bilateral price-links cannot be formed (groups A and C).

Table 1: Description of data containing price observations of alcoholic beverages

				Base period, *t=0*			Observation period, *t=1*		
Subgroup *k*	**ID-code**	**Retail specific classification**	**coicop 5**	**p**	**q**	**v**	**p**	**q**	**v**
A	912727	F110	Non-alcoholic wines	18	2	36	.	.	.
A	004696	C130	Wine	10	32	320	.	.	.
…	…								
A	007123	C130	Wine	12	61	732	.	.	.
B	429674	F110	Non-alcoholic wines	13	445	5 785	12	1 192	14 304
B	429677	F110	Non-alcoholic wines	11	265	2 915	11	523	5 753
…	…								
B	007131	C130	Wine	13	922	11 986	14	833	11 662
C	417217	F110	Non-alcoholic wines				8	43	344
C	466367	C112	Wine				17	32	544
…	…								
C	421937	C130	Wine				13	116	1 508

We use the following economic facts and methods for the data above:

1. Disappearing commodities simply means that consumption, production, selling or buying has come to end. New commodities mean that economic activity starts first time in the observation period.

   What kind of effects these commodities belonging into categories A and C should have for a subgroup *k* index number? Interpreting Pursiainen for vanishing commodities (time forwards): '*This condition states that the expenditure on a commodity tends to zero, then its effect on the index should vanish*' (Pursiainen, 2005, pp. 32-33).

2. For index number formulas satisfying the time reversal test (i.e. TRT) the condition may be reversed in time such that it should hold also for new commodities. In empirical analysis we impute missing quantities by small amount of quantities, say numbers 0.01 liters or even smaller, to prevent their effect on index number independently how missing prices have been imputed. We show that the excellent index numbers yield almost identical estimates and the basic index numbers are contingently biased.
3. We derive logarithmic presentations for all averages (including elementary aggregates) and show their inferiority in estimating price changes compared to benchmark statistic (i.e. excellent index number). We point out that the order of the calculation matters: First calculate relative change and then aggregate. Changing this order leads to serious bias.
4. It is sometimes thought that seriously biased price change estimates derived from some relative change of averages can be corrected by a hedonic method. This means simply that we remove the quality correction form actual price change of averages (see Vartia, Suoperä and Vuorio, 2019, Suoperä and Vuorio, 2019). We show that for our data the price models based on heterogeneously behaving cross-sections or time series econometric models applied for panel data are not practicable – price change estimates based on averages remains biased even after the quality correction of quantities is applied.

All analysis is done in logarithmic scale which enables the comparison of different methods.

2.3 Index Number Formulas

In this study all index number formulas are based on their logarithmic representation, that is

(1) $$log\left(P_f^{1/0}\right) = \sum_i w_{i,f} \cdot log\left(p_i^1\right) - \sum_i w_{i,f} \cdot log\left(p_i^0\right) = \sum_i w_{i,f} \cdot log\left(p_i^1/p_i^0\right)$$

As we see, the index number is based on 'averages' calculated for base and observation periods. In Table 2 we give weights ($w_{i,f}$) for most known index number formulas (sub index *f*).

The logarithmic presentation of formulas enables the comparison of different formulas by calculating deviations between them. The deviations between them is validated by simple algebra leading to covariances (or very close approximation of covariances) between, generally saying, weight differences and log-change of prices. Table 2 shows us that all 10 formulas looks similar and differ only by weighting. Six of them (*T, l, p* and *W* excluded), even Fisher, have been defined by logarithmic mean (see Vartia, 1976). The logarithmic representation of Stuvel have developed by Pursiainen (2005, pp. 88). Like Pursiainen noted, Stuvel and Montgomery-Vartia are closely connected - the weights of the Stuvel decomposition are based on the arithmetic averages of quantities and the weights of Montgomery-Vartia are based on the actual quantities instead. (Pursiainen, 2005, pp. 88).

Because log-changes are additive and symmetric, all index numbers *f* can be expressed by their corresponding 'averages'. This form is used when 'optimal weighting for index number', say Törnqvist weighting, is compared for example with different basic averages (i.e. arithmetic and geometric averages of prices) calculated for base and observation periods.

Table 2: Weights for the most known index number formulas in logarithmic form (see Vartia & Suoperä, 2017, 2018).

Basic formula	
Symbol and name of formula	Weights of the formula
Laspeyres, f = La	$w_{i,f} = w_{i,La}^0 = L(p^1 q^0, p^0 q^0)$[2]
log- Laspeyres, f = l	$w_{i,f} = w_{i,l}^0 = v_i^0 / V^0$
log-Paasche, f = p	$w_{i,f} = w_{i,p}^1 = v_i^1 / V^1$
Paasche, f = P	$w_{i,f} = w_{i,P}^1 = L(p^1 q^1, p^0 q^1)$
Excellent formula	
Törnqvist, f = T	$w_{i,f} = \overline{w}_{i,T} = 0.5 \cdot (w_i^0 + w_i^1)$
Sato-Vartia, f = SV	$w_{i,f} = \overline{w}_{i,SV} = \frac{L(w_i^1, w_i^0)}{\sum L(w_i^1, w_i^0)}$
Montgomery-Vartia, f = MV	$w_{i,f} = \overline{w}_{i,MV} = L(p^1 q^1, p^0 q^0)$
Fisher, f = F	$w_{i,f} = \overline{w}_{i,F} = 0.5 \cdot (L(p^1 q^0, p^0 q^0) + L(p^1 q^1, p^0 q^1))$
Walsh, f = W	$w_{i,f} = \overline{w}_{i,W} = \left(w_i^0 \cdot w_i^1\right)^{1/2}$
Stuvel, f = S	$w_{i,f} = \overline{w}_{i,S} = L(p^1 \overline{q}, p^0 \overline{q})$

2.4 Aggregation of Basic Averages

We call unweighted and weighted arithmetic and geometric means as basic averages. In this study, we use their logarithmic representations. The analysis of price aggregation in logarithmic form is presented in Suoperä (2006, Appendix) and Vartia, Suoperä and Vuorio (2019, Appendix). Aggregation of arithmetic averages is based again to the logarithmic mean. Different averages for prices may be presented as

$$(2) \qquad log\left(A_m(p^t, w_m)\right) = \sum_i w_{i,m} \cdot log\left(p_i^t\right)$$

where sub index *m* marks different averages (i.e. *m* = *a*, *aw*, *g* and *gw*, see Table 3 below).

Table 3: Weights for basic averages in logarithmic form (i.e. equation (2)).

Basic averages, see Suoperä (2005) and Vartia, Suoperä and Vuorio (2019)	
Name and symbol of formula	Weights of the formula
Unweighted Arithmetic Average, m = a	$w_{i,m} = w_{i,a}^t = L(p^t, 1)$
Weighted Arithmetic Average, m = aw	$w_{i,m} = w_{i,aw}^t = L(p^t q^t, q^t)$
Unweighted Geometric Average, m = g	$w_{i,m} = w_{i,g}^t = 1/n$
Weighted Geometric Average, m = gw	$w_{i,m} = w_{i,gw}^t = q_i^t / Q^t$

The representation of averages in equation (2) makes possible to compare them with any 'average' presented in equation (1).

[2] L=logarithmic average (Vartia, 1976, p. 128)

3 Empirical Analysis of Differences Between Index Number Formulas

Our data contains about 3300 – 4100 price observations on alcoholic beverages that are comparable in quality. We partition these commodities into 41 separate subgroups based on retailer's categories and analyze empirically pair-wise differences between Stuvel (*S*), Törnqvist (*T*), Sato-Vartia (*SV*), Montgomery-Vartia (*MV*), Walsh (*W*) and Fisher (*F*). We also show empirically that the basic index numbers – Laspeyres (*L*), Log-Laspeyres (*l*), Paasche (*P*) and Log-Paasche (*p*) – are also contingently biased also for this data (Vartia & Suoperä, 2018). Our empirical analysis is a simple one – calculate differences between any two index numbers, say *r* and *k* using equation (1),

(3) $$log\left(P_r^{1/0}\right) - log\left(P_k^{1/0}\right) = \sum_i (w_{i,r} - w_{i,k}) \cdot log\left(p_i^1/p_i^0\right)$$

The above relation between any two index numbers *r* and *k* is interesting - when $\sum_i w_{i,r} = \sum_i w_{i,k} = 1$ (i.e. $E\left(w_f\right) = 1/n$, for all *f* = *r*, *k*), then the equation (3) reduces to[3]

(4) $$\sum_i (w_{i,r} - w_{i,k}) \cdot log\left(p_i^t/p_i^0\right) = n \cdot \Delta cov\left\{\left(w_{i,r} - w_{i,k}, log\left(p_i^t\right)\right)\right\}$$

The algebra for the equations 3 and 4 is based on well-known property of the covariance, that is [4], $cov(x, y) = E\{(x - E(x))(y - E(y))\} = E\{x - E(x)\} \cdot y = E\{y - E(y)\} \cdot x$ (see Vartia 1979). As was noted in Table 2, the weights for six index number formulas, or more precisely their logarithmic forms, are based on the logarithmic mean meaning that $\sum_i w_{i,r} \leq \sum_i w_{i,k} \leq 1$. In practice the sum of weights is normally very close to unity and clearly, when the number of commodities increases, the above equation (4) is closely approximated also for these index numbers (i.e. $E(w_f) \approx 1/n$). What will the equation (4) tell us? If for any two index numbers, the $cov\left\{\left(w_{i,r} - w_{i,k}, log\left(p_i^1/p_i^0\right)\right)\right\}$ approach zero, the difference of weights and logarithmic price changes are uncorrelated and very likely independently distributed. Practically this means that these two index numbers go 'hand-in-hand' and either of them can be selected. In this study of alcoholic beverages and their 41 subgroups, we show that the $cov\left\{\left(w_{i,r} - w_{i,k}, log\left(p_i^1/p_i^0\right)\right)\right\}$ approach zero for almost any two excellent index numbers like as a rule. We also show that this is not true for any basic index number – they are also here contingently biased. Order of our empirical analysis is as follows:

1. We calculate pair wise differences between excellent index numbers and because eq. (4) is almost always very close to zero any of them can be selected as benchmark statistics.
2. We calculate pair wise differences between benchmark statistics and basic index numbers. We show empirically that basic index numbers are data contingently biased.
3. We analyze how the basic averages (arithmetic and geometric) estimate the price changes. The comparations are again done pair-wisely between benchmark statistics and these basic averages. We show that basic averages are poor statistics for estimating price change for our data.
4. We study how suitable hedonic quality corrections for basic averages are for estimating price change.

Analysis of differences is based on graphical presentations of index series and their differences in the logarithmic scale.

---

[3] see Suoperä, 2006, pp. 3-6
[4] basic clause of aggregation

3.1 Difference Between Excellent Index Numbers

Our strategy for estimating price changes and constructing index series follows our earlier studies (see Vartia, Suoperä, Nieminen and Montonen, 2018a, 2018b). With this strategy the index series are free from chain error. We show here graphically the index series for some commodity groups and their differences in log-scale, that is eq. (3) and (4).

Most of 41 commodity subgroups behave similarly as group 'Mild wines=C110'[5] in Figures 1 and 2 – differences in log-scale are always very close to zero. The most extreme difference for index series constructed for excellent index numbers is seen in group 'Mild wines=C130'. In this group only Sato-Vartia deviates 'moderately' but others are closely related.

Figure 1: Index series for excellent index numbers commodity group 'Mild wines=C110' from 2016.0 to 2019.12.

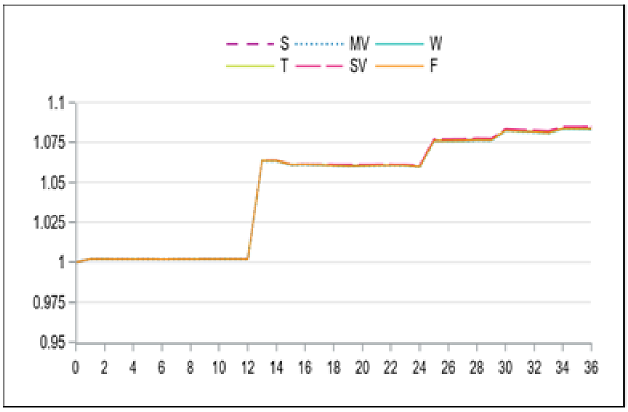


Figure 2: Corresponding differences between excellent index numbers and Törnqvist in log scale. a=*S-T*, b=*MV-T*, c=*SV-T*, d=*W-T* and e=*F-T*

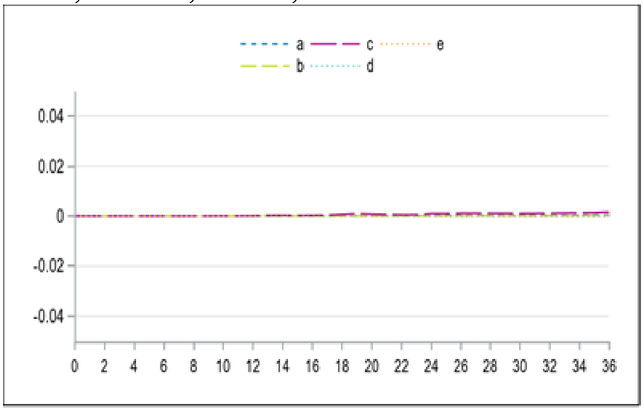


Figure 3: Index series for excellent index numbers commodity group 'Mild wines=C130' from 2016.0 to 2019.12.

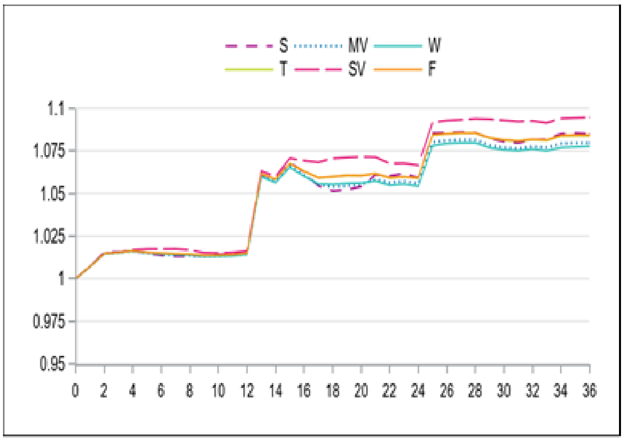
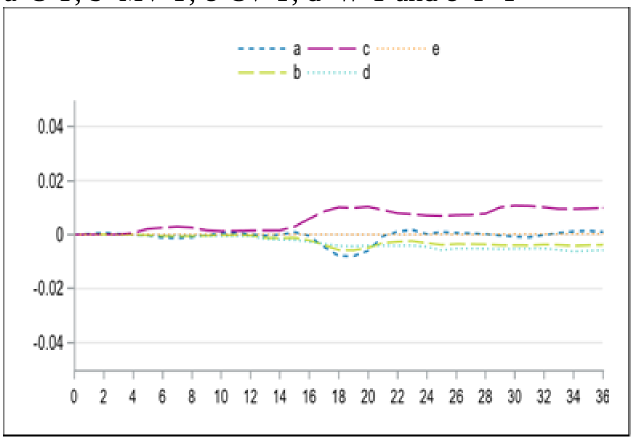


Figure 4: Corresponding differences between excellent index numbers and Törnqvist in log scale. a=*S-T*, b=*MV-T*, c=*SV-T*, d=*W-T* and e=*F-T*



The graphs tell more than 30 axioms[6] about what index number formula we should select for official statistics or at least for this data. We have analyzed differences for 41 subgroups and almost all groups give similar results as in Figure 2. Only Sato-Vartia deviates slightly from the other index formulas.

[5] C110=Champagne, C130=Rosé

[6] CPI manual p.7 Axiomatic and stochastic approaches to index numbers

Practically this means that for all 41 subgroups the $cov\left\{\left(w_{i,r} - w_{i,T}, log\left(p_i^1/p_i^0\right)\right)\right\} \cong 0$, *r* = *S*, *MV*, *W*, *F*, that is, pair-wise differences between weights $(w_{i,r} - w_{i,T})$ and $log\left(p_i^1/p_i^0\right)$ are almost surely independently distributed and are uncorrelated for these index numbers. The independence may be simply tested by regressing log price change on pair wise weight differences.

3.2 Difference Between Excellent and Basic Index Numbers

The following Figures show what data contingently biased index numbers means in practice. Figures 5 and 6 regarding 'Mild wines=C110' show that basic index numbers practically unbiased. For group 'Mild wines=C130', that is Figures 7 and 8, basic index numbers are already quite biased compared to excellent ones (here compared with Törnqvist).

Figure 5: Index series for basic index numbers[7] commodity group 'Mild wines=C110' from 2016.0 to 2019.12.

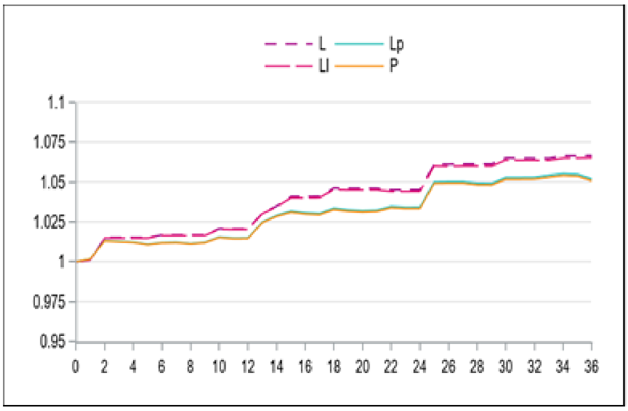


Figure 6: Corresponding differences between basic index numbers and Törnqvist in log scale. a=*L*-*T*, b=*Ll*-*T*, c=*Lp*-*T* and d=*P*-*T*.

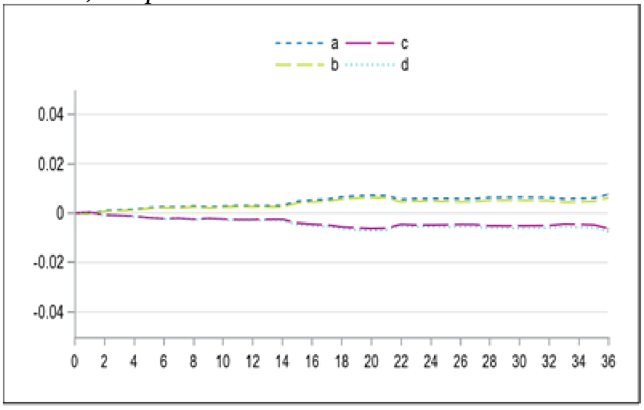


Figure 7: Index series for basic index numbers commodity group 'Mild wines=C130' from 2016.0 to 2019.12.

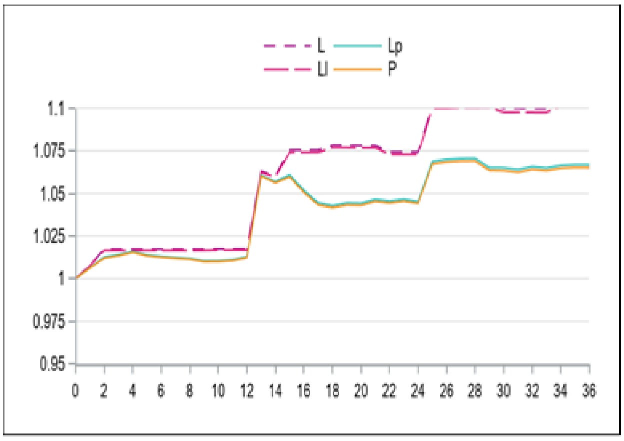
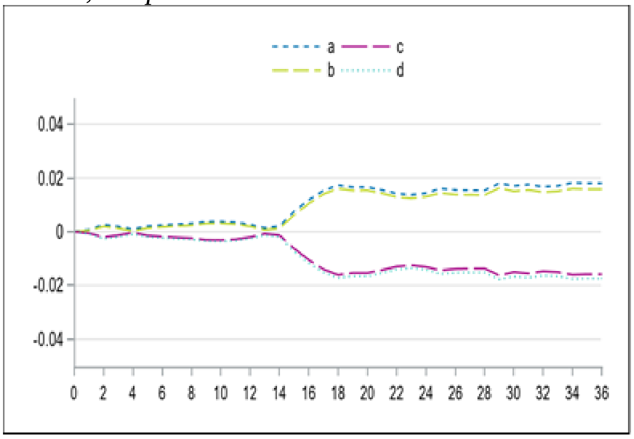


Figure 8: Corresponding differences between basic index numbers and Törnqvist in log scale. a=*L*-*T*, b=*Ll*-*T*, c=*Lp*-*T* and d=*P*-*T*.



In the other words, the data contingently biased index numbers behave quite nicely for some subgroups but are for some other seriously biased. The excellent index numbers mainly are closely related and their deviations from the *T* in log scale are always similar like in Figures 2 - never like in Figure 8.

[7] L=Laspeyres, Ll=log-Laspeyres, P=Paasche, Lp=log-Paasche

3.3 Difference Between Excellent Index Numbers and Basic Averages

The basic averages – weighted and unweighted arithmetic and geometric averages (i.e. eq. (2) and Table 3) - are useful for statistics but not universally for officially produced price indices. These averages are represented also here in logarithmic form so that their comparison with different index numbers may be easily done. The analysis of price aggregation is derived in Suoperä (2006, pp. 3-6, and Appendix) and in Vartia, Suoperä and Vuorio (2019, Appendix). See analogy also with Vartia (1979) and Suoperä & Vartia (2011, 2017, 2018). The basic idea is to compare different averages to 'optimally weighted averages for index numbers', say, averages of excellent index numbers, that is (see eq (1) and (2) and Tables 2 and 3)

$$(5) \qquad \sum_i w_{i,m} \cdot log(p_i^t) - \sum_i w_{i,f} \cdot log(p_i^t) = n \cdot cov\left\{\left(w_{i,m} - w_{i,f}, log(p_i^t)\right)\right\}$$

where n ·is number of observations, sub index *m* = *a*, *aw*, *g*, *gw* [8](see Table 3) and *f* = *S*, *T*, *MV*, *SV*, *W*, *F* (see Table 2).

Now we select for example *f* = *T* and perform numerical analysis for eq. (5) to get estimates for covariances, that is c_k^t, for all periods and subgroups *k* in question. There is two possible solutions: the estimates of covariances, c_k^t, will 'freeze' into a constant c_k for a given pair (*m*, *T*), for all time periods *t* and for subgroups *k*, or they do not. First possibility means practically that the differences of the eq. (5) between base and observation period, i.e. for price-link 0 → *t*, will be close to zero. In other words, the consumption pattern of quantities consumed in base and observation periods are closely related with the weights of the Törnqvist – logarithmic change of average *m* in question estimates closely the Törnqvist index number. If these differences of covariances deviates significantly from zero, the average *m* in question is not useful as price change statistic.

The equation (5) have another logarithmic representation, which more explicitly tells how the difference of (5) can be explained. This can be presented most simply for the weighted arithmetic average for commodity subgroup *k* (i.e. unit values for subgroup *k*), that is

$$(6) \qquad log\left(\frac{V^t}{Q^t}\Big/\frac{V^0}{Q^0}\right) - log\left(P_f^{t/0}\right) = log(V^t/V^0) + log(Q^0/Q^t) - log\left(P_f^{t/0}\right)$$

In the left side we have difference between the logarithmic change of weighted arithmetic averages and the logarithmic change of some excellent price index number. When the index number *f* decomposes precisely the total value change into price and quantity changes (here in log scale, i.e. $log(V^{t/0}) = log(P^{t/0}) + log(Q^{t/0})$), the equation (6) gives simple explanation to the left-side difference of (5) and (6).
For example, Fisher[9] satisfies this property and so we select *f* = *F* and get

$$(7) \qquad log\left(\frac{V^t}{Q^t}\Big/\frac{V^0}{Q^0}\right) - log\left(P_F^{t/0}\right) = log\left(P_F^{t/0}\right) + log\left(Q_F^{t/0}\right) + log(Q^0/Q^t) - log\left(P_F^{t/0}\right)$$

$$= log\left(Q_F^{t/0}\right) + log(Q^0/Q^t)$$

The eq. (7) reveals that price change based on the weighted arithmetic averages or unit values differs from the Fisher price index (or from *S*, *MV*, *SV* index numbers) by two quantity components – first the quantity index for formula *F* (or *S*, *MV*, *SV*) and second the pure change of total quantities. Two important situation emerges: If the consumption pattern of quantities consumed is not changed or change proportionally, then these two clauses, $log\left(Q_F^{t/0}\right)$ and $log(Q^0/Q^t)$, are opposite numbers and the weighted arithmetic averages may be used as price change estimate (i.e. equals with $log\left(P_F^{t/0}\right)$).

[8] a= unweighted arithmetic average, aw=weighted arithmetic average, g= unweighted geometric average, gw=weighted geometric average

[9] The similar mathematical characteristic is satisfied exactly also for Montomery-Vartia, Stuvel, and Sato-Vartia, but not for Törnqvist and Walsh (i.e. only close approximation for small changes).

For other situations, this occurs only by accident and the weighted arithmetic average may not be used as official price change estimate. Practically this means: When consumption patterns are linearly related for all pairs (0, *t*) then logarithmic change of the weighted arithmetic averages is a proper price change estimate, otherwise not. Based on this evidence, consumers do not behave as ideal demand theory suggests they would (see Vartia, Suoperä, Nieminen and Montonen (2018a)).

Most index series for excellent index numbers are similar step-type functions like in Figures 9 and 11 and they are very closely related. The Törnqvist in the Figure 9 and 11 show how prices have increased three times, each time at the beginning of the year. Between these increases, prices are held almost constant (i.e. no price change). This is because prices are not determined by markets but by administrative decision. All the excellent index numbers identify this phenomenon. Also, from the pictures we can see that the consumers' consumption pattern (i.e. quantities consumed) during any year changes even if prices have not practically changed.

Figure 9: 'Index series' for averages and Fisher, commodity group 'Mild wines=C130' from 2016.0 to 2019.12.

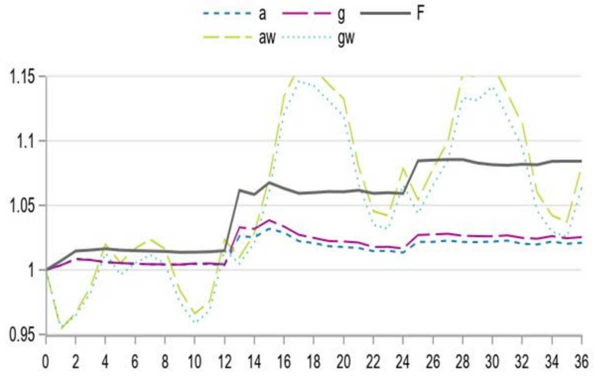


Figure 10: Corresponding differences between the for basic averages and Fisher in log scale. *a* = *a-F*, *b*=*aw-F*, c=*g-F* and *d*=*gw-F*.

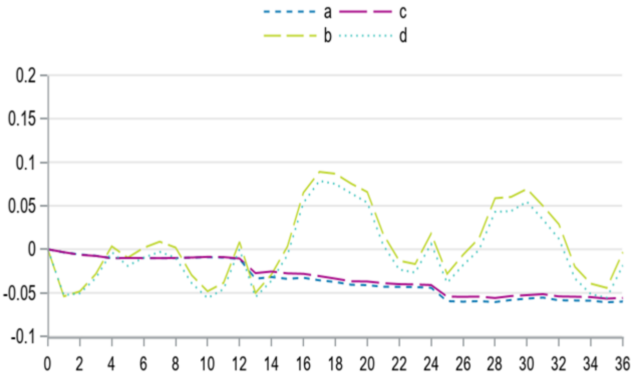


Figure 11: 'Index series' for averages and Fisher commodity group 'Other strong spirits=A151' from 2016.0 to 2019.12.

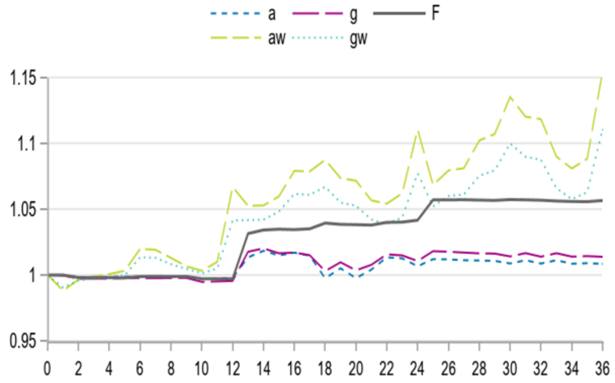
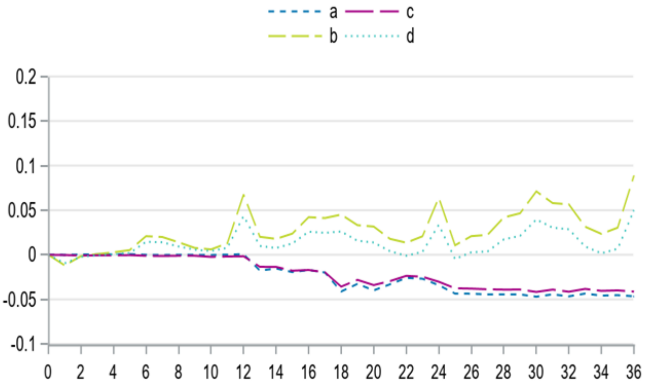


Figure 12: Corresponding differences between the for basic averages and Fisher in log scale. *a* = *a-F*, *b*=*aw-F*, c=*g-F* and *d*=*gw-F*..



It is sometimes believed that perhaps Jevons (i.e. unweighted geometric average $m = g$ in Figures) is a solution but we see that Jevons is seriously downward biased compared to the Fisher and other excellent index numbers. In other words, Jevons is data contingently biased compared to excellent index number – here and there it is quite properly behaving statistic but somewhere else seriously biased compared to *F* or any other excellent index number. As a conclusion, the basic averages are not suitable for official statistics for data described in Table 1.

Next we ask: is it possible to correct the bias of averages by a hedonic method? We use the technique described in Suoperä (2006) and Vartia, Suoperä and Vuorio (2019, Chapter 2.2.1) and we specify two specifications for our semilogarithmic price model – heterogeneously behaving cross-sections and time series econometric model (analysis of panel data, see Greene, 1993 pp. 621-623). We show that a hedonic method will not help us – basic averages are still out of use as official statistic even quality correction of quantities is performed.

3.4 A Hedonic Method – Alternative Method for Bilateral Analysis?

In this chapter, we investigate the possibility of using hedonic quality adjustment in averages although previous tests suggested otherwise. The mathematical and statistical analysis of price aggregation and a hedonic quality adjustment is presented in papers of Suoperä (2006, Chapter 3 and Appendix 5), Koev (2003), Vartia, Suoperä and Vuorio (2019) and Suoperä and Vuorio (2019). These papers are based on data where bilateral price-links cannot be formed - in the other words perhaps nothing but a hedonic method may be applied to get estimates for price changes. Our data contains bilateral price-links for which we apply a hedonic method. We use a so-called Oaxaca decomposition (1973) by which we divide the actual price change of a given average (i.e. *m* = *a*, *aw*, *g*, *gw*, see Table 3) into two parts – to the quality correction and the quality adjusted price changes. The results are presented for two different specifications of the price model.

Now we already know what distinguishes the basic averages from the best price change estimates, from the excellent index numbers. The explicit expression for the differences is derived in the equations (5), (6) and (7). Now we ask: How closely the hedonic quality corrections estimate the differences presented in (5), (6) and (7)? With empirical analysis, it can be shown that hedonic quality corrections do not correct the estimate in any way.

3.4.1 A Hedonic Method Assuming Heterogeneously Behaving Cross-Sections

We specify a simple semilogarithmic price model for 41 subgroups simply regressing logarithmic prices on quantities consumed for all subgroups *k* separately in time *t* (see Suoperä, 2006, Suoperä and Vartia, 2011, Vartia, Suoperä and Vuorio 2019), that is

$$\log(p_{ikt}) = \alpha_{kt} + x_{ikt}\beta_{kt} + \varepsilon_{ikt}$$

where $\log(p_{ikt})$ represents the logarithmic unit price for commodity *i* in period *t* and in subgroup *k*. The explanatory variable x_{ikt} represents quantities consumed for commodity *i* in period *t* and in subgroup *k*. The parameter β_{kt} in the regression model may vary according to grouping and time. Parameters α_{rkt} represent subgroup effects in period *t*. The term ε_{irt} is random error term, which does not contain systematic information about the data generating process. It is assumed, that $E\left(\varepsilon_{ikt}\middle|x'_{ikt}\right) = 0$ and $Var\left(\varepsilon_{ikt}\middle|x'_{ikt}\right) = \sigma_{kt}^2 < \infty$. In our model specification, the error covariance matrix is diagonal – a most natural situation for cross-sectional data. We apply the Oaxaca decomposition for averages presented in table 3. The method decomposes the logarithmic price change for our average *m* into quality correction and quality adjusted price change separately for each average (i.e. $log\left(p_A^{t/0}\right) \equiv log\left(p_{QC}^{t/0}\right) + log\left(p_{QA}^{t/0}\right)$, see Suoperä, 2006; Vartia, Suoperä and Vuorio, 2019, pp.13). The empirical analysis show that the price model based on heterogeneously behaving cross-sections estimates very poorly the quality corrections for quantities consumed. The estimation results presented in Figure 13 reveals this. For example ‘Mild wines= C130’ includes commodities that are seasonal meaning that quantities for (0, *t*) are not linearly related. In this case the hedonic method estimates poorly the quality corrections of quantities consumed in base and observation periods. The quality corrections should be approximately equal to eq. (5) – (7), but they are not. The Figures 13 – 14 tell the story.

Figure 13: Quality adjusted 'index series' for averages and Törnqvist for commodity group 'Mild wines=C130' from 2016.0 to 2019.12.

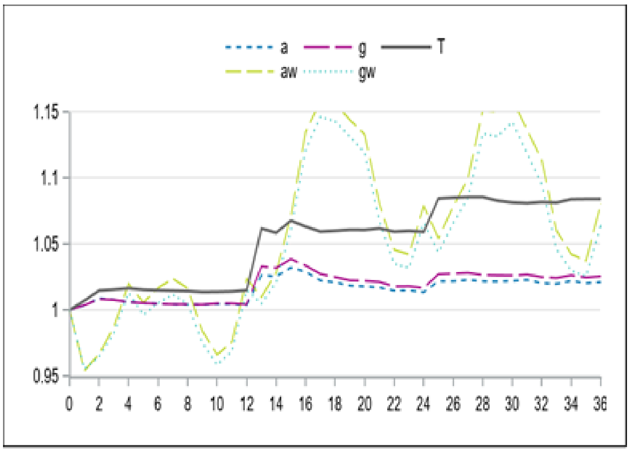
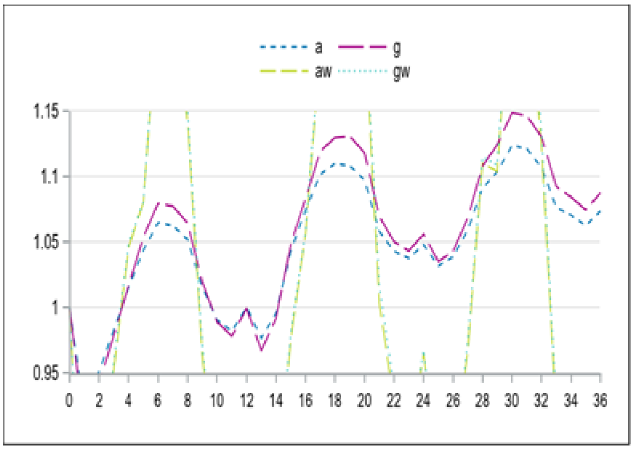


Figure 14: Corresponding differences between quality adjusted price change and Törnqvist in log-scale. *a* = *a-T*, *b*=*aw-T*, c=*g-T* and *d*=*gw-T*..



In Figure 14 we see how quality adjusted price changes for averages deviate from Törnqvist in log-scale. The deviations (i.e. difference of covariances) behave very systematically year-to-year and month-to-month. We have shown that consumption pattern changes all the time and is clearly seasonal, even when prices have not changed or have changed almost proportionally. These results confirm earlier conclusions; basic averages should not be used for estimating price changes when quantities vary a lot in time – even a hedonic method cannot help. Simply saying, do not use the basic averages as price change estimate for the CPI, if quantities consumed in base and observation periods are not linearly related all the time.

3.4.2 A Hedonic Method: Price Model Including Fixed Time and Group Effects

The price models belonging into this category are called as panel data models (or econometric time series models). The price model (i.e. regress logarithmic prices on quantities consumed) is completely based on bilateral price-links. The price model is equal to the multilateral unweighted time-product dummy model (TPD, Diewert and Fox, 2018, pp.15), that is

$$\log(p_{ikt}) = \alpha + \gamma_i + \rho_t + x_{ikt}\beta_k + \varepsilon_{ikt}$$

The dependent and explanatory variables are equal with the model specified as heterogeneously behaving cross-section. The unweighted estimation of unknown parameters is explained in Greene (1993, pp 621 - 623) and price aggregation and hedonic method in Suoperä (2006) and Vartia, Suoperä and Vuorio (2019). The estimation results are presented graphically in Figures 15-18.

Figure 15: Quality adjusted ‘index series’ for averages and Törnqvist for commodity group ‘Mild wines=C130’ from 2016.0 to 2019.12.

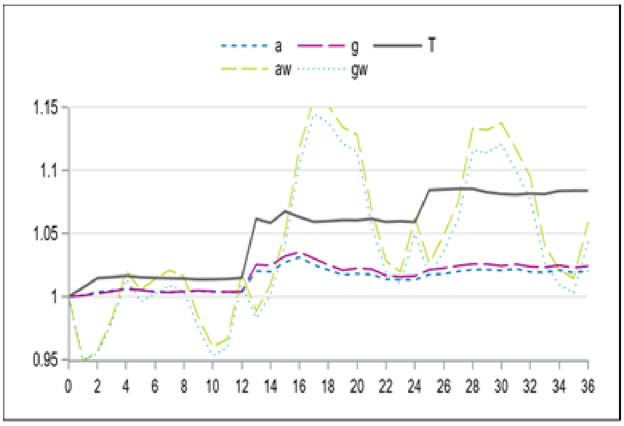


Figure 16: Corresponding Log-differences between quality adjusted index series for averages and Törnqvist. *a* = *a-T*, *b*=*aw-T*, c=*g-T* and *d*=*gw-T*.

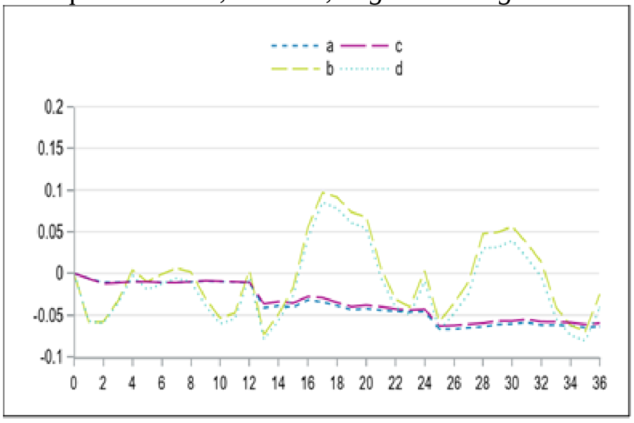


Figure 17: Quality adjusted ‘index series’ for averages and Törnqvist for commodity group ‘Other strong spirits=A151[10]’ from 2016.0 to 2019.12.

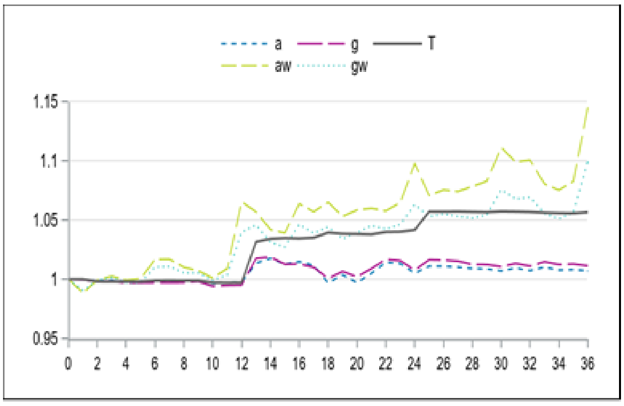
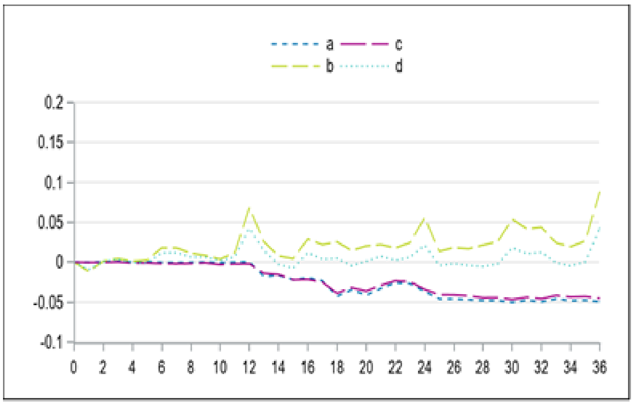


Figure 18: Corresponding Log-differences between quality adjusted index series for averages and Törnqvist. *a* = *a-T*, *b*=*aw-T*, c=*g-T* and *d*=*gw-T*.



The hedonic method based on the semilogarithmic price model including fixed time and commodity effects do not help – the bias compared to excellent Törnqvist remains severe. We have shown in this study for data having enough bilateral price-links that we should prefer excellent index numbers together with our base strategy instead.

[10] A151=Dark rums

4 Conclusion

Statistics Finland has examined alternative index calculation methods for several years in order to find best solution for calculating indices using various kinds of datasets having different composition of information. As result to this, we have introduced several scanner-datasets to the production of CPI and HICP during the recent years.

Now the aim is to replace traditional field collection of durables and consumer goods with similar commodities in scanner-data. One question is whether our current strategy and Törnqvist formula is sufficient or not, other is whether there are other issues that need to be studied before new commodities may be extracted from scanner-data and introduced to CPI? We noticed that

- scanner-datasets may have differing set of information and most often this is dependent of the data provider.
- commodities have different features, meaning that in some cases product quality remains stable in time and in some cases not
- we have various methods for calculating the price change. Selection of method depends on the content of scanner-data and product features.

Thus, it is necessary to study these challenges more carefully in order to find applicable solution for e.g. garments, shoes, mobile phones and home electronics. Design of the study is based on the fact that we must be a way to compare alternative methods in pairs because only this way we may recognize relation between the methods.

Our data is a typical complete micro data. It contains price observations of alcoholic beverages that are identified with GTIN-code; hence all products are comparable in quality in time. An observation contains unit prices, quantities and expenditures for all months since 2016. When investigating the observations (table 1, page 4), we notice that there are disappearing commodities, new commodities and commodities that are consumed steadily in time. In this study, we have 41 subgroups for alcoholic beverages, that we divide into three categories according to the information content:

- A, vanishing commodities,
- B, new commodities
- C, commodities that include information of prices, quantities and values for both base and observation periods and are comparable in quality.

The groups A and C contains missing values: in these groups data contains ‘one-side null values’, where the expenditure on a commodity tends to zero for base or observation period. According to Pursiainen ‘*This condition states that, its effect on the index numbers should vanish*’ (2005, pp. 32-33). We use this condition and impute vanishing expenditures (i.e. quantities) and missing prices such that this condition is satisfied very closely.

We use the following methods:

1. Bilateral price index numbers based on the base strategy that is free of drift error. We analyze the set of excellent and basic index numbers.
2. We compare basic averages (i.e. arithmetic and geometric) pair wisely with benchmark statistic (here excellent index number)
3. We try to quality adjust the quantities consumed by a hedonic method.

All analysis are done in logarithmic scale. This means that all index number formulas and basic averages are presented by their logarithmic representations. This makes it possible to compare these statistics pair wisely and understand how they are related. The logarithmic form of these statistics makes it easier to calculate differences between them.

We tested following index number formulas : Stuvel (*S*), Törnqvist (*T*), Sato-Vartia (*SV*), Montgomery-Vartia (*MV*), Walsh (*W*) and Fisher (*F*) and basic index number formulas: Laspeyres (*L*), Log-Laspeyres (*l*), Paasche (*P*) and Log-Paasche (*p*). Empirical analysis was simple, because we calculated differences between any two index numbers. If covariance for any two index numbers approach zero, the difference of weights and logarithmic price changes are uncorrelated and very likely independently distributed.
Practically this means that these two index numbers go 'hand-in-hand' and either of them may be selected. In this study we showed that the covariance approach to zero for almost any two excellent index numbers like as a rule, so any of the excellent index number formulas may selected for the production of official CPI. We also showed that this is not true for any basic index number – they are more or less contingently biased.

The basic averages – weighted and unweighted arithmetic and geometric averages are useful for statistics but not universally for official price change estimates. We compared different averages to 'optimally weighted averages for index numbers', say averages of excellent index numbers such as *MV* or *T*.
We selected Törnqvist formula as benchmark *(f = T)* and performed numerical analysis to get estimates for covariances for all periods and subgroups. Analysis revealed that there were two possible results: First, the estimates of covariances, 'freeze' into a constant for a given pair (*m*, *T*), for all time periods *t* and for subgroups *k*, or second, they vary all the time. First possibility means that the consumption pattern of quantities consumed in base and observation periods are closely related with the weights of the Törnqvist; logarithmic change of average *m* in question estimates closely the Törnqvist index number. If these differences of covariances deviates significantly from zero, the average is not applicable for price change statistic.
When consumption patterns are linearly related for all pairs (0, *t*) then logarithmic change of the weighted arithmetic averages is a proper price change estimate, otherwise not.

Selection of the method is crucial as consumption of the products differs even within 7-digit coicop sub-class. We have shown examples of how much basic averages, basic index number formulas and hedonic methods differ from the benchmark index, Törnqvist.

As a conclusion these tests have shown that

1. Any index number formula is suitable if products are sold evenly from month-to-month. It is strategy that matters more in these cases.
2. The basic averages are good price change estimates only if quantities have not changed or have changed proportionally.
3. A hedonic method is useful only if quantities consumed are linearly related.
4. The solution for alcoholic beverages where churn of products is not that high:
    a. Use the base strategy, where base period is previous year normalized as average month (covers also seasonal products) and use one of the excellent index number formula for production of official price statistics.
5. The solution for commodity groups where churn of products is clearly high:
    a. use bilateral comparison when matching-pairs are feasible and combine it with another method that is applied for one-sided nulls, e.g. a hedonic method[11].

[11] A hedonic method is a method for data that have no bilateral price-links. Examples of this kind of analysis is for example new or old dwellings (see Koev and Suoperä, 2002, Koev, 2003, Vartia, Suoperä and Vuorio, 2019).

References

Balk, B.M. ‘Axiomatic Price Index Theory: A Survey’, International Statistical Review 63: 69-93, 1990.

Diewert, W. E. ‘Exact and Superlative Index Numbers’, Journal of Econometrics, 4, 114-45, 1976.

Diewert, W. E. ‘Superlative Index Numbers and Consistency in Aggregation’, Econometrica, 46, 883-900, 1978.

Greene, W. “Econometric Analysis”, Prentice-Hall International, Inc. (third ed.), 1993.

Koev, E. ‘Combining Classification and Hedonic Quality Adjustment in Constructing a House Price Index’, Licentiate thesis, Helsinki, 2003.

Koev. E. & Suoperä. A. (2002), ’Omakotitalojen ja Omakotitonttien hintaindeksit 1985=100 (http://www.tilastokeskus.fi/tk/hp/kiinteistot_menetelma.pdf).

Oaxaca, R. ‘Male-Female Wage Differentials in Urban Labour Markets’, International Economic Review, 14, pp. 693 . 709, 1973.

Pursiainen, H. ‘Consistent Aggregation Methods and Index Number Theory’, Doctorial thesis in University of Helsinki, 2005.

Suoperä, A. ’Some new perspectives on price aggregation and hedonic index methods: Empirical application to rents of office and shop premises’, 2006 (unpublished, Statistics Finland).

Suoperä, A. & Vartia, Y. ‘Analysis and Synthesis of Wage Determination in Heterogeneous Cross-sections’, Discussion Paper No. 331, 2011.

Suoperä, A. and Vuorio, J. ’A Hedonic Price Index Number for New One-Family Detached Houses in Finland’ 2019

Vartia, Y. ‘Relative Changes and Index Numbers’, Ser. A4, Helsinki, Research Institute of Finnish Economy, 1976.

Vartia, Y. ‘Ideal Log-Change Index Numbers’, Scandinavian Journal of Statistics., 3, pp. 121 - 126, 1976.

Vartia, Y. ’Kvadraattisten mikroyhtälöiden aggregoinnista’, ETLA, Discussion Papers no. 25,1979.

Vartia, Y. & Suoperä, A. “Index number theory and construction of CPI for complete micro data”, 2017. (http://www.stat.fi/meta/menetelmakehitystyo/index_en.html).

Vartia, Y. & Suoperä, A. “Contingently biased, permanently biased and excellent index numbers for complete micro data”, 2018. (http://www.stat.fi/static/media/uploads/meta_en/menetelmakehitystyo/contingently_biased_vartia_suopera_updated.pdf)

Vartia, Y., Suoperä, A., Nieminen, K. & Montonen, S. “Circular Error in Price Index Numbers Based on Scanner Data. Preliminary Interpretations”, 2018a. (http://www.stat.fi/meta/menetelmakehitystyo/index_en.html)

Vartia, Y., Suoperä, A., Nieminen, K. & Montonen, S. “The Algebra of GEKS and Its Chain Error”, 2018b. (http://www.stat.fi/meta/menetelmakehitystyo/index_en.html)

Vartia, Y., Suoperä, A. and Vuorio, J. 'Hedonic Price Index Number for New Blocks of Flats and Terraced Houses in Finland', 2019.