ב״ה

Analytic Assessment of Index Formulae Inaccuracies

by

R.B. Yehezkael (formerly Haskell)
Jerusalem College of Technology - Machon Lev,
Havaad Haleumi 21, Jerusalem 91160
E-mail: rafi@mail.jct.ac.il Fax: 02-6422075 Tel: 02-6751111
Revised December 2009 - כסלו תש״ע

Abstract

In an earlier paper (Yehezkael 1991), computational tests and a computer simulation were used to quantitatively assess inaccuracies of several price index formulae. In this paper the analytic approach for assessing these inaccuracies is developed and is found to be in good agreement with the estimates obtained computationally in our earlier paper. The treatment presented here is more general than our earlier approach and covers price and quantity indices. The contribution of a price index formula to price inflation is also discussed.

Keywords

Index, Price, Quantity, Value, Test, Accuracy, Stability, Inflation.

Contents

NOTE: A shortened version of this paper appears in the CDROM Proceedings of the WSES International Conference on Mathematics and Computers in Business and Economics 2001 (MCBE 2001), Skiathos Island, Greece, September 2001. It also appears in "Mathematics and Simulation with Biological, Economical and Musicoacoustical Applications", pp. 260-267, Edited by C. E. D'Attellis, V. V. Kluev, N. E. Mastorakis, WSES Press, 2001.

Introduction

In an earlier paper (Yehezkael 1991), we assessed the inaccuracies of several index numbers from a variety of viewpoints:

- pragmatically, by analyzing United Kingdom price indices for the years 1958 to 1967. Index values from (Fowler 1970) were used; these covered Laspeyres and Paasche indices in chained and unchained forms.
- Qualitatively by mathematical (algebraic) tests of common sense requirements of these formulae (e.g. $I_{bi} = 1/I_{ib}$).
- Quantitatively by computational tests and a computer simulation, this being the main direction of the earlier work.

There are certain limitations of the previous approach:

- We only treated price indices in situations where price level was the independent variable. Situations where quantity is the independent variable were not treated. Similarly the more general case of price and quantity related through a third parameter was not treated.
- Quantity indices were not treated.
- The approach used computer simulation, but expected values were not calculated as it was not practical to generate all possible samples. So one had to be satisfied with averages calculated from samples generated using random number generators.

In this paper while still addressing the issue of quantitative estimates of the inaccuracies in index numbers, the approach we take will be analytical and not computational. We shall treat a situation of price and quantity related through a third parameter (This includes as special cases, price as independent variable, and, quantity as independent variable.)

We shall analyze both price and quantity indices using the concept of convergence in probability to determine the limit of expected values of various indices. The approach of this paper is also limited in that we give results when the number of index constituents $n \to \infty$.

We note that the calculation of the expected value of an expression with n continuous variables, will require evaluating a definite multiple integral with n variables. This will involve calculating the value of an alternating sum with 2^n terms. So for large n, this can not be used in practice in view of both extremely long computational time and accumulated rounding errors. So approximation of one form or another seems inevitable. The "computational approximation" is to use random number generators and not generate all samples. The "analytic approximation" is to give results for $n \to \infty$ (based on convergence in probability). Fortunately we find that the "analytic approximation" and the "computational aproximation" are in good agreement (see later) and this gives us confidence in the results presented both here and in the previous paper.

A related topic we discuss, is how the choice of price index formula affects price inflation.

Conventions

The following conventions are used throughout this paper

1) The values of j, j_1, j_2 are always integers in the range 1 to n where n is a positive integer.
2) Σ_j will always mean $\sum_{j=1}^{n}$ and $\Sigma_{j_1 j_2}$ will always mean $\sum_{j_1=1}^{n} \sum_{j_2=1}^{n}$.
3) $\lim_n$ will alway mean $\lim_{n\to\infty}$.
4) G(X) denotes the geometric expected value based on the geometric mean. More on this in Appendix A.

NOTE: Mathematical preliminaries are described in Appendix A. Price, Quantity, Value Index formulae and other general requirements are described in Appendix B.

The symmetric case - The true geometric expected index value is known

Let us suppose that prices and quantities in month b the base month and month i the current month are identically distributed and let us suppose that $b \neq i$, i.e. they are different months. Let us suppose that prices and quantities in the base month are symmetrically distributed with respect to prices and quantities in the current month. Then for any combination of prices and quantities occuring in the current month, the reverse situation with price and quantity levels in base and current month interchanged is equally likely to occur, since base and current month situations are completely symmetric. It follows that the true geometric expected value of the index should be one in view of property 2 in our general requirements of index formulae that I_{bi} $I_{ib} = 1$. (Actually in the limit, it does not matter whether we use the geometric or usual expected value in view of points 4,6 in Appendix A.) Regarding prices, Fischer's index, the Unit index and the Normalized unit index satisfy property 2. Concerning, quantities, Fischer's index, the Ratio of total quantities, and the Ratio of total normalized quantities satisfy property 2. So all these indices will have geometric expected value one in any such situations, i.e. their geometric expected values are completely accurate in these cases.

To understand what happens to the indices of Laspeyres and Paasche, we shall formalize a slightly restricted form of the previous situation. We shall assume that all prices and quantities are defined by random variables which are bounded, positive, and bounded away from zero. (Our earlier assumptions about symmetry of prices and quantities in base and current month continue to hold).

We shall determine the geometric expected values for Laspeyres' and Paasche's indices using a specific model for prices and quantities. Since the geometric expected value is not analytically tractable when n the number of items is finite, we use the results in Appendix A to determine what happens when $n \to \infty$. We show that even in this case, the geometric expected value can be seriously different from one.

Suppose that $E(p_{ij}) = E(p_{bj}) = \mu^p{}_j$ and that $E(q_{ij}) = E(q_{bj}) = \mu^q{}_j$.

Suppose that the means and covariances of the price quantity products for the same month (b=i) and for different months (b≠i) are defined as follows.

Same month: $\mu_j^{pq} = E(p_{ij}q_{ij}) = E(p_{bj}q_{bj})$ and

$c_{j_1 j_2}{}^{pq} = \mathrm{cov}(p_{ij_1}q_{ij_1}, p_{ij_2}q_{ij_2}) = \mathrm{cov}(p_{bj_1}q_{bj_1}, p_{bj_2}q_{bj_2})$.

Different months: $\mu'_j{}^{pq} = E(p_{ij}q_{bj}) = E(p_{bj}q_{ij})$ and

$c'_{j_1 j_2}{}^{pq} = \mathrm{cov}(p_{ij_1}q_{bj_1}, p_{ij_2}q_{bj_2}) = \mathrm{cov}(p_{bj_1}q_{ij_1}, p_{bj_2}q_{ij_2})$.

So $\Sigma_j\, p_{ij}q_{bj} / \Sigma_j\, \mu'_j{}^{pq} \to_p 1$ and $\Sigma_j\, p_{bj}q_{bj} / \Sigma_j\, \mu_j^{pq} \to_p 1$ by point 8 in Appendix A providing that $\lim_n \Sigma_{j_1 j_2} (c'_{j_1 j_2}{}^{pq}) / (\Sigma_j\, \mu'_j{}^{pq})^2 = \lim_n \Sigma_{j_1 j_2} (c_{j_1 j_2}{}^{pq}) / (\Sigma_j\, \mu_j^{pq})^2 = 0$. (As all the μ's are greater than zero so the denominators are non zero.)

So $(\Sigma_j\, p_{ij}q_{bj} / \Sigma_j\, \mu'_j{}^{pq}) / (\Sigma_j\, p_{bj}q_{bj} / \Sigma_j\, \mu_j^{pq}) \to_p 1$.

In view of our requirements on prices and quantities, division is a well defined and continuous function and so

$\lim_n E((\Sigma_j\, p_{ij}q_{bj} / \Sigma_j\, \mu'_j{}^{pq}) / (\Sigma_j\, p_{bj}q_{bj} / \Sigma_j\, \mu_j^{pq})) = 1$ by point 4 in Appendix A, providing that $(\Sigma_j\, p_{ij}q_{bj} / \Sigma_j\, \mu'_j{}^{pq}) / (\Sigma_j\, p_{bj}q_{bj} / \Sigma_j\, \mu_j^{pq}))$ is bounded.

Rearranging, $\lim_n E(\Sigma_j\, p_{ij}q_{bj} / \Sigma_j\, p_{bj}q_{bj}) = \lim_n \Sigma_j\, \mu'_j{}^{pq} / \Sigma_j\, \mu_j^{pq}$ providing that this limit exists. Also by point 6 in Appendix A, a similar result holds for the limit of the geometric expected value.

So $\lim_n G(L_{bi}) = \lim_n E(L_{bi}) = \lim_n \Sigma_j\, \mu'_j{}^{pq} / \Sigma_j\, \mu_j^{pq}$ for Laspeyres' price index.

Similarly for the Paasche price index it can shown that when similar conditions to the above hold, then $\lim_n G(P_{bi}) = \lim_n E(P_{bi}) = \lim_n \Sigma_j\, \mu_j^{pq} / \Sigma_j\, \mu'_j{}^{pq}$ providing that this limit exists. This we note is the reciprocal of the corresponding limits of Laspeyres price index. For Laspeyres and Paasche's quantity indices we will get the same limits as the corresponding price indices.

In general $\mu_j^{pq} \neq \mu'_j{}^{pq}$ so errors in $G(L_{bi})$ and $G(P_{bi})$ can be expected. We also note that when p_{bj} and q_{bj} are independent, p_{ij} and q_{ij} are independent, p_{ij} and q_{bj} are independent, and p_{bj} and q_{ij} are independent then $\mu_j^{pq} = \mu'_j{}^{pq} = \mu^p{}_j\, \mu^q{}_j$. So in this case $\lim_n G(L_{bi}) = \lim_n G(P_{bi}) = 1$ i.e. in the limit, the geometric expected value is accurate. In our earlier work we observed this empirically (Yehezkael 1991, page 148, start of first paragraph).

Let us now refine the above discussion and consider how elasticity of demand influences the limiting values of Laspeyre's and Paasche's indices in the symmetric case.

Since all prices and quantities are positive, $(\Delta q \Delta p)$, $(\Delta q/\Delta p)$ and the elasticity of demand $(\Delta q/\Delta p) / (q/p)$ all have the same sign. So for example if all elasticities are negative, then $\Delta q \Delta p = (q_{ij}-q_{bj})(p_{ij}-p_{bj}) < 0$

Rearranging, $p_{ij}q_{ij} + p_{bj}q_{bj} < p_{ij}p_{bj} + p_{bj}q_{ij}$

Taking expected values, $\mu_j^{pq}+\mu_j^{pq} < \mu'_j{}^{pq} + \mu'_j{}^{pq}$, i.e. $\mu_j^{pq} < \mu'_j{}^{pq}$.

Summing and rearranging, $\Sigma_j\, \mu'_j{}^{pq} / \Sigma_j\, \mu_j^{pq} > 1$.

So, $\lim_n G(L_{bi}) = \lim_n E(L_{bi}) \geq 1$. Similarly, $\lim_n G(P_{bi}) = \lim_n E(P_{bi}) \leq 1$.

However, these limits should be 1. So when all elasticities are negative, the limits Laspeyre's price index will typically be too high and the limits for Paasche's index will be too low. Similarly, when all elasticities are positive, the limits will typically be too low for Laspeyre's index and too high for Paasche's indices. Also, as elasticity of demand is usually negative, we expect that in practice that Laspeyres' will be too high and Paasche's index too low. We observed similar effects in our earlier paper (Yehezkael 1991, page 148, second and third paragraph). This effect can also be seen from tables 1 and 2 and is discussed directly following these tables.

Tabulations

Let us now tabulate the errors in Laspeyres and Paasches indices for a symmetric situation where there is no correlation between base and current months (as before $b \neq i$). Consider a situation where prices and quantities are related through random variables z_{ij} and the independent varaiations in price and in quantity are provided by random variables x_{ij} and y_{ij} respectively. Specifically, suppose that $p_{ij} = p^*_j (z_{ij})^{e_p} x_{ij}$ and $q_{ij} = q^*_j (z_{ij})^{e_q} y_{ij}$ where e_p , e_q are exponents, p^*_j , q^*_j are positive constants which scale price and quantity levels, and x_{ij}, y_{ij}, z_{ij} are independent random variables having mean 1, which vary uniformly in the ranges $(1 \pm v_p/100)$, $(1 \pm v_q/100)$, $(1 \pm v/100)$ respectively. (We note that when $e_p=1$ and $v_p=0$ i.e. $x_{ij}=1$ then in effect price is the independent variable and quantity the dependent variable. Similarly when $e_q=1$ and $v_q=0$ i.e. $y_{ij}=1$ then quantities are the independent variables and price the dependent variables.)

The bounds put on v, v_p, v_q ensure that all prices, quantities, and the ratio $(\Sigma_j p_{ij}q_{bj} / \Sigma_j \mu'^{pq}_j) / (\Sigma_j p_{bj}q_{bj} / \Sigma_j \mu^{pq}_j)$ are bounded positive and bounded away from zero. The independence criteria will ensure that:

(a) $\mu_j^{pq} = \mu^p_j \, \mu^q_j$

(b) $\lim_n \Sigma_{j_1 j_2} (c'^{pq}_{j_1 j_2}) / (\Sigma_j \mu'^{pq}_j)^2 = \lim_n \Sigma_j (c'^{pq}_{jj}) / (\Sigma_j \mu'^{pq}_j)^2$
 and this is proportional to $\lim_n \Sigma_j (p^*_j q^*_j)^2 / (\Sigma_j p^*_j q^*_j)^2$.

(c) $\lim_n \Sigma_{j_1 j_2} (c^{pq}_{j_1 j_2}) / (\Sigma_j \mu^{pq}_j)^2 = \lim_n \Sigma_j (c^{pq}_{jj}) / (\Sigma_j \mu^{pq}_j)^2$
 and this is proportional to $\lim_n \Sigma_j (p^*_j q^*_j)^2 / (\Sigma_j p^*_j q^*_j)^2$.

So providing that $\lim_n \Sigma_j (p^*_j q^*_j)^2 / (\Sigma_j p^*_j q^*_j)^2 = 0$, the results from the previous section may be applied as follows.

$\lim_n G(L_{bi}) = \lim_n E(L_{bi}) = \lim_n \mu'^{pq}_j / \Sigma_j \mu^{pq}_j = \lim_n \Sigma_j \mu^p_j \mu^q_j / \Sigma_j \mu^{pq}_j$.

$= \lim_n \Sigma_j p^*_j q^*_j E(z_{ij}^{e_p}) E(z_{bj}^{e_p}) / \Sigma_j p^*_j q^*_j E(z_{bj}^{e_p+e_q})$ - in view of independence criteria and since $E(x_{ij}) = E(y_{ij}) = 1$,

$= \Sigma_j p^*_j q^*_j I(e_p, v/100) I(e_q, v/100) / \Sigma_j p^*_j q^*_j I(e_p+e_q, v/100)$

$= I(e_p, v/100) I(e_q, v/100) / I(e_p+e_q, v/100)$.

The requirement above that $\lim_n \Sigma_j (p^*_j q^*_j)^2 / (\Sigma_j p^*_j q^*_j)^2 = 0$, holds in many cases, for example when $(p^*_j q^*_j)$ is constant, or j, or a polynomial in j. It does not hold universally though, for example when $(p^*_j q^*_j) = a^j$.

Similarly $\lim_n G(P_{bi}) = \lim_n E(P_{bi}) = I(e_p+e_q, v/100) / [I(e_p, v/100) I(e_q, v/100)]$.

What is interesting about these limiting values in these situations, is that they do not depend on p^*_j, q^*_j, which scale prices and quantities nor on the independent variations v_p, v_q in prices and quantities. When all the variables z_{ij} have the same variation v they only depend on this variation v. We observed a related effect in our previous paper - (page 147 footnote to table 2).

Since $I(e_p, 0) = I(e_q,0) = I(e_p+e_q, 0) = I(0,v/100) = 1$ by point 9 in Appendix A, the limits of geometric expected values of Laspeyres' and Paasch's index will be exactly 1 (i.e. they are completely accurate) when $v = 0$ or $e_p = 0$ or $e_q = 0$. These cases correspond to either no variation at all of prices and quantities, or prices and quantities being independent of each other.

Tables 1 and 2 give the limiting values in the errors in Laspeyre's and Paasche's indices where we have expressed them as percentage errors from the true geometric expected value of 1. The tables cover e_p and e_q in the range -2 to +2 and v in the range 10% to 90%. (To save space, we do not table values when $v = 0$ or $e_p = 0$ or $e_q = 0$ since here the percentage error is 0 - see above. Also in view of the fact that the error is unchanged if we interchange the values of the exponents e_p and e_q, we show under the same column the errors for two pairs of exponents when $e_p \neq e_q$.)

TABLE 1

Limits of mean percentage error per measurement in Laspeyres' index for various values of v, e_p, e_q

v	(e_p,e_q) (-2, -2)	(e_p,e_q) (-1, -2) (-2, -1)	(e_p,e_q) (-1, -1)	(e_p,e_q) (1, -2) (-2, 1)	(e_p,e_q) (1, -1) (-1, 1)	(e_p,e_q) (1, 1)	(e_p,e_q) (2, -2) (-2, 2)	(e_p,e_q) (2, -1) (-1, 2)	(e_p,e_q) (2, 1) (1, 2)	(e_p,e_q) (2, 2)
10	-1.33	-0.67	-0.33	0.67	0.34	-0.33	1.35	0.67	-0.66	-1.31
20	-5.26	-2.69	-1.36	2.76	1.37	-1.32	5.56	2.72	-2.56	-4.95
30	-11.65	-6.11	-3.13	6.51	3.17	-2.91	13.19	6.27	-5.50	-10.22
40	-20.25	-11.03	-5.77	12.40	5.91	-5.06	25.40	11.56	-9.20	-16.27
50	-30.77	-17.60	-9.48	21.37	9.86	-7.69	44.44	19.02	-13.33	-22.41
60	-42.86	-26.06	-14.59	35.25	15.52	-10.71	75.00	29.39	-17.65	-28.15
70	-56.16	-36.81	-21.71	58.26	23.90	-14.04	128.10	44.14	-21.92	-33.27
80	-70.33	-50.56	-32.11	102.28	37.33	-17.58	237.04	66.62	-26.02	-37.67
90	-85.04	-68.92	-49.16	221.75	63.58	-21.26	568.42	107.75	-29.83	-41.38

TABLE 2

Limits of mean percentage error per measurement in Paasche's index for various values of v, e_p, e_q

v	(e_p,e_q) (-2, -2)	(e_p,e_q) (-1, -2) (-2, -1)	(e_p,e_q) (-1, -1)	(e_p,e_q) (1, -2) (-2, 1)	(e_p,e_q) (1, -1) (-1, 1)	(e_p,e_q) (1, 1)	(e_p,e_q) (2, -2) (-2, 2)	(e_p,e_q) (2, -1) (-1, 2)	(e_p,e_q) (2, 1) (1, 2)	(e_p,e_q) (2, 2)
10	1.35	0.67	0.34	-0.67	-0.33	0.33	-1.33	-0.67	0.66	1.33
20	5.56	2.76	1.38	-2.69	-1.35	1.33	-5.26	-2.65	2.63	5.21
30	13.19	6.51	3.23	-6.11	-3.08	3.00	-11.65	-5.90	5.83	11.38
40	25.40	12.40	6.13	-11.03	-5.58	5.33	-20.25	-10.36	10.13	19.43
50	44.44	21.37	10.47	-17.60	-8.98	8.33	-30.77	-15.98	15.38	28.88
60	75.00	35.25	17.08	-26.06	-13.44	12.00	-42.86	-22.71	21.43	39.18
70	128.10	58.26	27.73	-36.81	-19.29	16.33	-56.16	-30.62	28.08	49.85
80	237.04	102.28	47.30	-50.56	-27.18	21.33	-70.33	-39.98	35.16	60.44
90	568.42	221.75	96.69	-68.92	-38.87	27.00	-85.04	-51.86	42.52	70.58

A crosscheck: Since, $p_{ij} = p^*_j\ (z_{ij})^{e_p}\ x_{ij}$ and $q_{ij} = q^*_j\ (z_{ij})^{e_q}\ y_{ij}$, therefore $q_{ij} = q^*_j\ (p_{ij}/(\ p^*_j\ x_{ij}\ )^{1/e_p})^{e_q}\ y_{ij}$. So the elasticity of demand is $(\partial q/\partial p)/(q/p) = e_q/e_p$. We note from tables 1 and 2, that when the ratio e_q/e_p is negative, then Laspeyres' index is too high and Paasche's index is too low. Similaraly when e_q/e_p is positive we see that in this case Laspeyres' index is too low and Paasche's index is too high. This is consistent with our comments regarding the elasticity of demand in the previous section. However the magnitude of the error is not a function of the elasticity of demand since for example there are many cases when $e_p=e_q$, with different error characteristics and in these cases the elasticity of demand is one.

Notes:

1. When $e_p=e_q$ and $v_p=v_q=0$, there is a proportional relationship between price and quantity.
2. When $e_p+e_q=0$ and $v_p=v_q=0$ price and quantity are inversely proportional to each other i.e. a reciprocal relationship holds between price and quantity.
3. Similarly we can show that if $p_{ij} = p^*_j(1+c_p z_{ij}^{e_p})x_{ij}$ and $q_{ij} = q^*_j(1+c_q z_{ij}^{e_q})y_{ij}$ and providing that the constants c_p, c_q are chosen so that p_{ij}, q_{ij} are always positive and providing that $\lim_n \Sigma_j\ (\ p^*_j\ q^*_j\ )^2 / (\ \Sigma_j\ p^*_j\ q^*_j\ )^2 = 0$, then

$$\lim_n E(L_{bi}) = \lim_n G(L_{bi}) = \frac{1+c_pI(e_p,v)+c_qI(e_q,v)+c_pc_qI(e_p,v)(Ie_q,v)}{1+c_pI(e_p,v)+c_q(I(e_q,v)+c_pc_qI(e_p+e_q,v)}$$

and $\lim_n E(P_{bi}))$ = $\lim_n G(P_{bi})$ = reciprocal of the above fraction.

4. The special case of 3 when $e_p=e_q$ and $v_p=v_q=0$ corresponds to a linear relationship between p_{ij} and q_{ij}, where the straight line does not pass through the origin.
5. While the situations discussed do not constitute an economic model we believe we have included sufficient detail in these situations to appraise the inaccuracies of index formulae. We are able to specify the relationship between price and quantity levels, the independent variations in the common parameter, in prices and in quantities. Indeed, the situations are general and as can be seen from the previous notes, they include linear and reciprocal relationships between prices and quantities. We will draw an analogy from the physical sciences in support of using such a simplification. In order test a scale for weighing human beings, one uses standard test weights; no one would suggest using "standard human beings". Obviously, the more test weights used the greater the confidence in the scale. Similarly to test indices for accuracy, test cases where the true (average) index value is known, seems to us a valid and practical approach for gaining a quantitative indication of the potential errors of index formulae, the use of a realistic economic model not being an essential requirement. Also, an economic model with a large number of different kinds of variables would be difficult to handle analytically.

Previous results from the analytic viewpoint

In our previous paper (Yehezkael 1991) we simulated by computer a restricted form of the symmetric case we just tabulated where there was only one item being sold by many sellers and where there was no correlation between base and current months.

The following substitutions and correspondences should be used in comparing tables 1,2 in this paper, with table 2 in our previous paper.

This paper	Previous paper
e_q	Elasticity
$e_p = 1$	--
v	Independent variation in price
$v_p = 0$	--
v_q	Independent variation in quantity

With this correspondence, there is close agreement between the results of this paper and the previous paper. Let us illustrate this with two example comparisons..

When v=60%, v_q = 80%, e_p = 1 and e_q = -2 the error in Laspeyres index is +35.25% from table 1 in this paper. From table 2 in our previous paper using the above correspondence, we get +35.2873% for this situation. This is very good agreement where in our previous paper, n = 1000 and here $n\rightarrow\infty$.

Similarly when v = 80% and v_q = 0% and e_q = -1 the error in Paasche's index is -27.18% table 2 in this paper. From table 2 in our previous paper using the above correspondence we get a value of -27.1147% again very good agreement.

The single item case - True index formulae are known

We must first ask what can be learnt by using an index formula intended for use with many items, to measure price change of a single item from many sellers. (i.e. when using index formulae, j will denote seller number and p_{bj}, p_{ij} denote price levels of the j^{th} seller and q_{bj}, q_{ij} denote quantity levels sold by the j^{th} seller.). It seems much simpler to handle the single item case than the many item case since here prices of different sellers are likely to be similar whereas in the many item cases there may be orders of magnitude differences between prices of different items. So if an index formula does not handle the single item case well, there is little or no chance of it handling the many item case well.

When we have only one item being sold by many sellers there is only one possibility for the quantity index, namely the Ratio of total quantities sold R_{bj}. There is also only one way to determine average cost per item sold and the ratio of the average cost in the current month with respect to the base month is the Unit index U_{bi} which is the true price index here. We also note that the product $U_{bi}R_{bi}$ = V_{bi} which is the value index and this meets our previous requirments about the product of price and quantity indices. So we can compare other price indices against the Unit index U_{bi} and other quantity indices with the Ratio of total quantities sold R_{bi}. We propose that this comparison be done with real data about prices and quantities of a single item being sold by many sellers. We also propose that this be done computationally using a realistic economic model. Here we shall not carry out such a quantitative analysis but continue to appraise this situation qualitatively.

We note that the Normalized unit index is closest in form to the Unit index (the true price index here) and so may perform better than Fischer's price index which has a rather different form. Similarly the Ratio of total normalized quantities is closest in form to the Ratio of total quantities (the true quantity index here) and so may perform better than Fischer's quantity index which has a rather different form.

We also note that if there are no changes whatsoever in prices but in the current month the cheaper sellers increase their market share with respect to the base month, then the Unit index (the true price index here) will be less than one and Laspeyres', Paasche's and Fischer's index will be exactly one. Also, any weighted price index or (arithmetic and geometric) means of weighted price indices will be exactly one. Similarly if the more expensive sellers increase their market share the Unit index will be greater than one and Laspeyre's, Paasch's and Fischer's pirce index will be exactly one. Also any weighted price index or arithmetic and geometric means of different weighted price indices will be exactly one. Regarding the normalized Unit index, what happens will depend on how m_{bij}, the mean price of the j'th item, is defined. We now consider four possibilities for defining m_{bij} according to the framework given in the section on Notation and formulae.

1) When $m_{bij} = (p_{ij}\, q_{ij} + p_{bj}\, q_{bj}) / (q_{ij} + q_{by})$ then as $p_{ij} = p_{bj}$ i.e. no price change, then $m_{bij} = p_{bj} = p_{ij}$ and so the Normalized unit index will be precisely one. So with this choice of m_{bij} the Normalized unit index will not behave like the Unit index but will behave like the other indices above. Perhaps this indicates that this is a bad coice for m_{bij}.

2) Consider the possibility

$$m_{bij} = \sum_{k=m_1}^{m_2} p_{kj}\, q_{kj} \;\Big/\; \sum_{k=m_1}^{m_2} q_{kj}$$

where m_1, m_2 are the minimum and maximum of b and i respectively. In the chained case i.e. i = b+1 or b = i+1 this is exactly the same as the previous case. However when there is a reasonable gap between b and i, the values of m_{bij} are likely to be similar as j here denotes a seller number and all seller's are selling the same item. So these values will approximately cancel and there is a reasonable chance that the Normalized Unit index will behave like the Unit index. Perhaps then even here this is not a particularly satisfactory situation in view that the quality of the Normalized unit index will depend on the gap between b and i.

3) Consider the possibility for chained indices i.e. i=b+1 or b=i+1 that

$$m_{bij} = \sum_{k=m_1-1}^{m_2+1} p_{kj}\, q_{kj} \;\Big/\; \sum_{k=m_1-1}^{m_2+1} q_{kj}$$

where m_1, m_2 are as in 2). This also seems a good choice since this definition is symmetrical with respect to past and future and uses data from a four month period from a month before and after base and current months. However its use will delay publication of an index by one month which depending on the situation, may or may not be acceptable.

4) Consider the possibility that

$$m_{bij} = \sum_{k=m_1-t}^{m_1-1} p_{kj}\, q_{kj} \;\Big/\; \sum_{k=m_1-t}^{m_1-1} q_{kj}$$

where m_1 is as in 2). In this case, the average is for a t month period before both b and i. Here, regardless of the gap between b and I and providing t is “reasonably long” (4 months? 12 months?), there is, we believe, a good chance of the values of m_{bij} being similar; then the Normalized unit index is likely to behave in similar fashion to the Unit index.

To sum up: (a) Apart for the Normalized unit index none of the price indices we discussed are likely to behave like the Unit index (the true price index in this case). (b) The third and fourth choices for m_{bij} are likely to give good overall performance for the Normalized unit index.

Regarding quantity indices, for similar reasons to those just discussed we think that the third and fourth choices for m_{bij} are also good choices for ensuring that the Ratio of total normalized quantities will be close to the Ratio of total quantities (the true quantity index in this case). Further quantitative analysis will clarify what really are good choices for m_{bij} in these situations.

Regarding other quantity indices we note that as $L_{bi}\, P'_{bi} = P_{bi}\, L'_{bi} = F_{bi}\, F'_{bi} = V_{bi}$ so any error in Laspeyres price index will be accompanied by a reciprocal error in the Paasche quantity index and vice versa. Similarly Fischer's price and quantity indices will also have reciprocal error characteristics.

A case where there is time dependency

In the single item case, true price and quantity indices are known. So we do not need symmetric situations in the base and current month and can assess the inaccuracies of price and quantity indices in cases where there are time dependencies. Let us consider a situation where there are n sellers, "arranged in a circle", all selling the same item. Let us suppose that seller j is influenced only by sellers l(j) and r(j), the sellers to his "left" and "right" respectively, and that:

$l(j) = j-1$ if $j>1$ and $l(j) = n$ if $j=1$,

$r(j) = j+1$ if $j<n$ and $r(j) = 1$ if $j=n$.

Suppose prices and quantities in month b, the base month, are defined by independent positive random variables for different sellers. We assume that in all months, the products of price and quantity of every seller is a constant i.e. $p_{bj}q_{bj}=p_{ij}q_{ij}=c$, etc. A seller determines his new price from his old price and the old prices of his left and right neighbour using a weighted average of 1/2 for his old price and 1/4 for the old price of each of his two neighbors,

i.e. $p_{i+1,j} = p_{ij}/2 + (p_{i,l(j)} + p_{i,r(j)})/4 = (p_{i,l(j)} + 2p_{ij} + p_{i,r(j)})/4$.

We could ask what happens in the long term (large i), but we do not know how to handle this analytically. We could ask what happens in a limiting equilibrium situation (large i) where $p_{i+1,j}=p_{ij}$ and $q_{i+1,j}=q_{ij}$. However, this will give us no information on the inaccuracies of the index formulae since here the $I_{i,i+1}=1$ for all index formulae. So let us determine what happens for various index formulae at the first step. Specifically, we compare a price index formula $I_{b,b+1}$ with the unit index $U_{b,b+1}$ and a quantity index formula $I'_{b,b+1}$ with the ratio of total quantities $R'_{b,b+1}$.

The following can be shown to hold (after simplification).

$U_{b,b+1} = (\Sigma_j\, q_{bj}) / (\Sigma_j\, q_{b+1,j})$

$L_{b,b+1} = \Sigma_j\, p_{b+1,j}q_{bj} / nc$

$P_{b,b+1} = nc / \Sigma_j\, p_{bj}q_{b+1,j}$

$F_{b,b+1} = \sqrt{[(\Sigma_j\, p_{b+1,j}q_{bj}) / (\Sigma_j\, p_{bj}q_{b+1,j})]}$

$N_{b,b+1} = (\Sigma_j\, M_{b,b+1,j}q_{bj}) / (\Sigma_j\, M_{b,b+1,j}q_{b+1,j})$

We see in the above the relevance of our earlier remarks concerning the single item case that the form of the normalized unit index is closest to the Unit index, the true price index here.

In the above situation the value index $V_{b,b+1} = 1$. It therefore follows from the correspondences in table 3 that for quantity indices,

$R'_{b,b+1} = 1/U_{b,b+1}$

$P'_{b,b+1} = 1/L_{b,b+1}$

$L'_{b,b+1} = 1/P_{b,b+1}$

$F'_{b,b+1} = 1/F_{b,b+1}$

$Q'_{b,b+1} = 1/N_{b,b+1}$

So as above with price indices, the ratio of total normalized quantities will be closest in form and likely to have a similar value to the ratio of total quantities, the true quantity index here.

The contribution of a price index formula to price inflation

The symmetric case gives us some indication as to how, on the average, the choice of index formula can affect inflation. We shall now analyze other situations but qualatatively.

Consider a situation where an item sells well in the base month and does not sell at all in the current month because of a drastic price increase and suppose there are no other price changes. Then Laspeyres price index will be greater than Fischer's price index which will be greater than Paasche's price index which will be exactly one. (Paasche's formula will ignore the price of this item in both base and current month in view of the fact that its current quantity level is zero.) It is not clear what will happen to the Unit index and the Normalized unit index but both of them ignore the price of this item in the current month (when the quantity level is zero) and make use of the price of this item in the base month (when the quantity level is not zero).

Consider also the opposite situation where an item does not sell in the base month but sells well in the current month because of a major price reduction and suppose there are no other changes in either price or quantity. Then Paasche's price index will be less than Fischer's price index which will be less than Laspeyres' price index which will be exactly one. It is not clear what will happen to the Unit index and Normalized unit index but both of them will ignore the price of this item in the base month (when the quantity level is zero) and make use of the price of this item in the current month (when the quantity level is not zero).

We believe that the property of both the Unit and Normalized unit index of always multiplying price of an item in a month by its quantity in the same month, will give both these indices good stability characteristics, as unrealistic prices will be typically downweighted by low quantities. (For the Normalized unit index, this is also true for all definitions we gave for m_{bij}.) Laspeyres', Paasche's and Fischer's indices do not have this property and weighted indices and (arithmetic and geometric) means of different weighted indices also do not have this property and they may have poor stability characteristics.

The reader will see from the following two simple examples, how the above stability characteristic, gives the Unit Index and the Normalized Unit the ability to ignore unrealistic prices in both the examples below.

Example of a sudden price increase making a price unrealistic.

Initial purchase -

Fruit basket: 2 kilos apples and 2 kilos bananas.

Prices: apples at 2 coins a kilo and bananas at 2 coins a kilo.

Most recent purchase -

Fruit basket: 4 kilos apples 0 kilo bananas.

Prices: apples at 2 coins a kilo and bananas at 4 coins a kilo.

Here are the values of the indices.

Laspeyres	1.50
Paasche	1.00
Fischer	1.22
Unit	1.00
Normalized Unit	1.00

Example of a sudden price decrease because of an unrealistic price.

Initial purchase -

Fruit basket: 4 kilos apples 0 kilo bananas.

Prices: apples at 2 coins a kilo and bananas at 4 coins a kilo.

Most recent purchase -

Fruit basket: 2 kilos apples and 2 kilos bananas.

Prices: apples at 2 coins a kilo and bananas at 2 coins a kilo.

Here are the values of the indices.

Laspeyres	1.00
Paasche	0.75
Fischer	0.87
Unit	1.00
Normalized Unit	1.00

Note: In calculating the Normalized Unit index we used the first possibility for calculating m_{bi} - see section titled "The single item case - True index formulae are known".

Conclusion

In taking the analytic approach to obtain quantitative estimates in the errors of index numbers, we were able to considerably widen the results obtained using the computational approach in our earlier paper.

In the symmetric case, we handled a situation where prices and quantities of n different items, were related through common parameters a considerably more general situation than handled in our earlier paper of just having one item being sold by n sellers and just having quantities determined from prices. We also handled both price and quantity indices. The results we got in this more general situation have a similar form to the more specific case dealt with previously. There was also close numerical agreement between the analytical and the earlier computational approach, the results in this paper only strengthen and confirm what we wrote in our earlier paper. They confirm our fears of the potential for serious errors of a single measurement in the indices of Laspeyres and Paasche, whether used for prices or quantities and even when large amounts of data are used in the calculation i.e. $n \rightarrow \infty$.

In our qualitative analysis of the single item case, we presented situations where Laspeyres', Paasche's, and Fischer's price indices, and weighted price indices in general, were insensitive to changes in the true price index (the Unit index in this case). We indicated that the

Normalized unit index is likely to behave like the Unit index in this case, providing that m_{bij} (the mean price per item of the j'th item) was appropriately chosen. Similarly the Ratio of total normalized quantities is likely to be close to the Ratio of total quantities - the true quantity index in this case, subject to the same choice for m_{bij}. The other quantity indices are also not likely to handle this case as well.

In our qualitative assessment of the contribution of an index formula to price inflation we suggested that the form of price quantity products used in the Unit and Normalized Unit indices may give them good stability characteristics and the absence of these forms may give poor stability characteristics for example to weighted indices. We illustrated this with simple examples.

Finally on the basis of our work so far we recommend the following indices.

For prices: Unit index in the single item case and Normalized unit index in the many item case.

For quantities: Ratio of total quantities in the single item case and Ratio of total normalized quantities in the many item case.

Further work should concentrate on quantifying the qualitative arguments we presented and confirming or refuting them. The single item case should be further analyzed using real data and using a realistic economic model. A realistic economic model should also be used to analyze how the choice of price index formula affects price inflation.

Acknowledgements

It was J.W. Kwiatkowski of the Department of Business Studies, University of Edinburgh who suggested the general direction for the analytical approach. My sincere thanks are due to him. Sincere thanks are also due to D. Mohr and L. Ritchie for their prompt and efficient typing of the manuscript, and to the referees and to P. Levine for their constructive comments.

References

Allen, R. G. D. 1975 *Index Numbers in Theory and Practice*. Edinburgh: Macmillan

Banerjee, K. S. 1975 *Cost of Living, Index Numbers, Price, Precision and Theory*. New York: Marcel Dekker

Fowler, R. F. 1970 *Some problems of index number construction*. In: *Studies in Official Statistics*, Research serires No. 3, HMSO London, 1970

Parzen E. 1960 *Modern Probability Theory and its Applications*, J.Wiley and Sons.

Yehezkael, R. B. 1991 *Computational tests as a means of assessing the inaccuracies of index numbers*. IMA journal of Mathematics Applied in Business and Industry, Vol. 3 pp 141-152

Yeomans, K. A. 1974 *Statistics for the Social Scientist:1 Introducing Statistics*. Penguin Education

APPENDIX A - Mathematical Preliminaries

We shall assume that the reader is familiar with the following concepts and results some of which appear in (Parzen 1960).

1) Convergence in probability denoted by $\to_p$.

Expected value denoted by E (...). Covariance denoted by cov(... , ...) .

2) If $X_n \to_p X$ and X_n is bounded, then $\lim_n E(X_n) = X$. Recall, $\lim_n$ means $\lim_{n\to\infty}$.

3) If f is a continuous function and $A_n \to_p A$, $B_n \to_p B$, $C_n \to_p C$ ... then $f(A_n, B_n, C_n....) \to_p f(A,B,C...)$.

4) If in addition f is bounded, then $\lim_n E(f(A_n, B_n, C_n...) = f(A,B,C...)$.

5) Providing that the random variable X>0, we shall used G(X) to denote the geometric expected value based on the geometric mean. If X takes only a finite number of values $x_1....x_k$ with probabilities $p_1......p_k$ respectively then $G(X) = x_1^{p_1} x_2^{p_2} ...x_k^{p_k}$ (There is no need for a root here since $p_1 + p_2.....+p_k = 1$. as they are the sum of all the probabilities.)

This can be written in the form $G(X) = e^{(p_1 \ln x_1 + p_2 \ln x_2 + \dots p_k \ln x_k)}$ where "ln" denotes natural logarithm (i.e. $\ln = \log_e$).

However, $E(\ln(X)) = (p_1 \ln x_1 + p_2 \ln x_2 ... + p_k \ln x_k)$.

So $G(X) = e^{E(\ln x)}$, and in this form, G can be determined even for random variables X>0, which vary continuously.

6) A result similar to (4) holds for the geometric expected value G(...) when the function f(...) is continuous, bounded, positive, and bounded away from zero (i.e. for some a, b we have

$0<a\leq f(...)\leq b$). This result follows by applying (4) to the function ln (f(...)) which will be both continuous and bounded and so $\lim_n E(\ln(f(A_n, B_n, C_n...))) = \ln(f(A,B,C...))$. Applying the exponential function e^x to both sides and moving it inside the limit by continuity of e^x, we get: $\lim_n e^{E(\ln(f(A_n,B_n,C_n...))} = e^{\ln(f(A,B,C...)} = f(A,B,C,...)$ which by definition of the geometric expected value this yields $\lim_n G(f(A_n,B_n,C_n...)) = f(A,B,C,....)$.

We thus see that in this case the limits of the usual and geometric expected value are identical and equal to f,(A,B,C,....).

(We noted in our previous paper, pages 147-148, remarks 2, that similar results were obtained when using arithmetic and geometric means in the computations. It is interesting to also see this effect in the analytical setting.)

7) Suppose that X_1, X_2, X_n ... are random variables having means $\mu_1,\mu_2,...\mu_n...$ and standard deviations $\sigma_1,\sigma_2,...\sigma_n...$ and suppose that $\lim_n \mu_n = \mu$ and $\lim_n \sigma_n = 0$. Then $X_n \to_p \mu$.

<u>Proof:</u> Choose N sufficiently large so that $|\mu_N-\mu|<\varepsilon/2$. So when $|X_N-\mu_N|<\varepsilon/2$ we certainly have $|X_N-\mu|\leq|X_N-\mu_N|+|\mu_N-\mu|<\varepsilon$. So clearly $P(|X_N-\mu|<\varepsilon)\geq P(|X_N-\mu_N|<\varepsilon/2)$ or equivalently $P(|X_N-\mu|\geq\varepsilon)\leq P(|X_N-\mu_N|\geq\varepsilon/2)$.

By Chebychev's inequality $P(|X_N-\mu_N|\geq\varepsilon/2)\leq\sigma_N^2/(\varepsilon/2)^2=4\sigma_N^2/\varepsilon$ and so certainly $P(|X_N-\mu|\geq\varepsilon)\leq 4\sigma_N^2/\varepsilon$. But irrespective of the value of ε, the value of $4\sigma_N^2/\varepsilon$ can be made arbitrarily small for sufficiently large N since $\lim_n \sigma_n = 0$. So $X_n\to_p\mu$.

8) Suppose that X_1, X_2, X_n ... are random variables having means $\mu_1,\mu_2,...\mu_n...$ and in addition $\Sigma_j\ \mu_j\neq 0$. Then $\Sigma_j X_j / \Sigma_j \mu_j$ will have expected value 1 (for any n) and variance (i.e. standard deviation squared) $(\Sigma_{j_1 j_2} \mathrm{cov}(X_{j_1}, X_{j_2}))/(\Sigma_j \mu_j)^2$. So by (7) providing that $\lim_n (\Sigma_{j_1 j_2} \mathrm{cov}(X_{j_1}, X_{j_2}))/(\Sigma_j \mu_j)^2 = 0$, then $\Sigma_j X_j / \Sigma_j \mu_j\to_p 1$. (Recall that j, j_1, j_2, and the summations are in the range 1 to n.)

This result reduces to the weak law of large numbers when $\mu_n=\mu\neq 0$, and the random variables are pairwise independent and all have the same variance i.e. $\mathrm{cov}(X_m, X_n)=\sigma^2$ when m=n, and $\mathrm{cov}(X_m, X_n)=0$ when $m\neq n$.

9) Let X be a random varable uniformly distributed from 1-v to 1+v where $0\leq v<1$. Let us define

$I(k,v)=1$ when v=0 and

$$I(k,v)=(1/2v)\int_{1-v}^{1+v} X^k\,dX \qquad \text{when } v\neq 0.$$

We note that

$I(k,v) = [\,(1+v)^{k+1} - (1-v)^{k+1}] / [2v(k+1)]$ when $k\neq 1$ and $v\neq 0$.

$I(k,v) = [\,\ln(1+v) - \ln(1-v)] / [2v]$ when k=1 and $v\neq 0$.

We note here that from these definitions that in all cases $I(0,v) = I(k,0) = 1$ and $E(X^k) = I(k,v)$.

APPENDIX B - Notation and formulae

The following notation is used throughout the paper.

Notation common to all indices:

n number of price and quantity levels i.e. number of items in the many item case and number of sellers in the single item case

i current month

b base month

j item number in the many item case or seller number in the single item case. The value of j is always in the range 1 to n

p_{ij} price in month i of item j

q_{ij} quanity sold in month i of item j

I_{bi} the value of an index formula in current month i with respect to base month b

Notation for the value index:

V_{bi} - Value index

This is the ratio of the change in value sold at month i with respect to month b, and here there is only one possibility.

$$V_{bi} = \frac{\Sigma_j\ p_{ij}\,q_{ij}}{\Sigma_j\ p_{bj}\,q_{bj}}$$

Notation for price indices:

L_{bi} - Laspeyres' price index

P_{bi} - Paasche's price index

F_{bi} - Fischer's price index

U_{bi} - Unit index for prices

N_{bi} - Normalized unit index for prices

$$L_{bi} = \frac{\Sigma_j \; p_{ij} \, q_{bj}}{\Sigma_j \; p_{bj} \, q_{bj}}$$

$$P_{bi} = \frac{\Sigma_j \; p_{ij} \, q_{ij}}{\Sigma_j \; p_{bj} \, q_{ij}}$$

$$F_{bi} = \frac{\sqrt{(\Sigma_j \; p_{ij} \, q_{ij} \;\; x \;\; \Sigma_j \; p_{ij} \, q_{bj})}}{\sqrt{(\Sigma_j \; p_{bj} \, q_{ij} \;\; x \;\; \Sigma_j \; p_{bj} \, q_{bj})}} = \sqrt{(L_{bi} \, P_{bi})}$$

$$U_{bi} = \frac{\Sigma_j \; p_{ij} \, q_{ij} \, / \, \Sigma_j \; q_{ij}}{\Sigma_j \; p_{bj} \, q_{bj} \, / \, \Sigma_j \; q_{bj}}$$

We note that the Unit index is the ratio of the average cost per item in the current month with respect to the average cost per item in the base month.

$$N_{bi} = \frac{\Sigma_j \; p_{ij} \, q_{ij} \, / \, \Sigma_j \; q_{ij} \, m_{bij}}{\Sigma_j \; p_{bj} \, q_{bj} \, / \, \Sigma_j \; q_{bj} \, m_{bij}}$$

Some explanation is in order regarding the Normalized unit index which is based on the Unit index. The Unit index formula can only be used in the case of a single item, being sold for example by various sellers. To extend it to the case of several items we have in a certain sense to equate all items in a natural way. One way of doing this is to say that one "normalized unit" of the j'th item, is the amount of that item that can be purchased for one unit of currency based on its mean price m_{bij}. The cost of one "normalized unit" of item j in month i is therefore $p_{ij} \, / \, m_{bij}$ and the quantity consumed is $q_{ij} \, m_{bij}$. Using the unit index formula with prices and quantitiies based on normalized units gives us the formula above (after some simplification). In words the above formula says the normalized unit index N_{bi} is the change, expressed as a ratio of the average price per "normalized unit", where "normalized unit" was defined above.

Regarding defining m_{bij} , the mean price of the jth item, it is natural to define it as the ratio of the total value of j'th item sold over the desired months to total quantity sold over these months, i.e. $m_{bij} = \Sigma_k \; p_{kj} q_{kj} \; / \; \Sigma_k \; q_{kj}$ where k ranges over the desired months. We shall also require that $m_{bij} = m_{ibj}$ as this will ensure that the Normalized unit index satisfies $N_{bi} = 1/N_{ib}$. Later we will discuss within this framework, four possibilities for m_{bij}.

(We note that if we forget for a moment the interpretation we gave to m_{bij} and allow ourselves to substitute freely on it we can in fact derive the Paasche, Laspeyres and Unit Index formulae for prices from the Normalized Unit Index formula. The substitution $m_{bij} = p_{bj}$ causes that formula to reduce to the Paasche formula. The substitution $m_{bij} = p_{ij}$ causes that formula to reduce to the Laspeyres formula. The substitution $m_{bij} = 1$ causes that formula to reduce to the Unit Index formula.)

Notation for quantity indices:

L'_{bi} - Laspeyres' quantity index

P'_{bi} - Paasche's quantity index

F'_{bi} - Fischer's quantity index

R'_{bi} - Ratio of total quantities

Q'_{bi} - Ratio of total normalized quantities

$$L'_{bi} = \frac{\Sigma_j \; q_{ij} \, p_{bj}}{\Sigma_j \; q_{bj} \, p_{bj}}$$

$$P'_{bi} = \frac{\Sigma_j \; q_{ij} \, p_{ij}}{\Sigma_j \; q_{bj} \, p_{ij}}$$

$$F'_{bi} = \frac{\sqrt{(\Sigma_j \; q_{ij} \, p_{ij} \;\; x \;\; \Sigma_j \; q_{ij} \, p_{bj})}}{\sqrt{(\Sigma_j \; q_{bj} \, p_{ij} \;\; x \;\; \Sigma_j \; q_{bj} \, p_{bj})}} = \sqrt{(L'_{bi} \, P'_{bi})}$$

$$R'_{bi} = \frac{\Sigma_j \; q_{ij}}{\Sigma_j \; q_{bj}}$$

$$Q'_{bi} = \frac{\Sigma_j \; q_{ij} \, m_{bij}}{\Sigma_j \; q_{bj} \, m_{bij}} \quad \text{where } m_{bij} \text{ is defined as before}$$

(Again we note that if we forget for a moment the interpretation we gave to m_{bij} and make the same substitutions we made in our discussion of price indices, we can in fact derive the Laspeyres, Paasche and Ratio of total quantities indices for quantities from the Ratio of total normalized quantities.)

Weighted indices:

Laspeyres' and Paasche's indices are examples of weighted indices which have the forms

$$\frac{\Sigma_j \; p_{ij} \, w_j}{\Sigma_j \; p_{bj} \, w_j} \quad \text{for prices, and} \quad \frac{\Sigma_j \; q_{ij} \, w'_j}{\Sigma_j \; q_{bj} \, w'_j} \quad \text{for quantities.}$$

Laspeyres index is an example of a fixed weight index as the weights do not depend on data from the current month, and Paasche's is an example of a weighted index but not a fixed weight index. Also a mean of weighted indices can be taken, an example being Fischer's index which is the geometric mean of Laspeyres' and Paasche's indices.

Some general requirements:

In similar style to our previous paper let us qualitatively evaluate index formulae with respect to the four properties below (Allen 1975, Banerjee 1975, Yeomans 1970). (The restrictions imposed on properties 2,3 below, make the first three properties independent of each other.)

1) $I_{ii} = 1$.
2) $I_{bi} = 1/ I_{ib}$ for $b \neq i$.
3) $I_{bi} = I_{bk} \, I_{ki}$ for $b < i < k$
4) The value of I_{bi} should be independent of the units in which quantities are expressed.

Regarding these properties and whether we are dealing with prices or quantities, the formulae of Laspeyres and Paasche satisfy properties 1 and 4, the formula of Fischer satisfies properties 1,2, and 4. The Unit index for prices satisfies properties 1,2,3,4 when used to measure the change in price of one item sold by many sellers but it only satisfies properties 1,2,3 if used to measure the change in price of many items. It is therefore not used in the many item case but is the best choice in the single item case. The Ratio of total quantities satisfies properties 1,2,3,4 when used to measure the change in quantity level of one item sold by many sellers but it only satisfies properties 1,2,3 if used to measure the change in price of many items. It is therefore not used in the many item case but is the best choice in the single item case. The Normalized unit index satisfies properties 1,2,4. The Ratio of total normalized quantities satisfies properties 1,2,4. (The Normalized unit index and the Ratio of total normalized quantities both satisfy property 2 because $m_{bij} = m_{ibj}$. They satisfy property 4 because normalized unit is defined in terms of average price of an item and it does not depend on the actual units used to sell an item.)

Regarding the Value index, this satisfies properties 1,2,3,4 but of course it can not be used to measure prices and quantities.

Another sensible requirement regarding the pair of price and quantity indices to use (table 3), is to require that their product equals the Value index which after all has only one possible definition.

TABLE 3

Pairs of price and quantity indices whose product equals the Value index

Price index	Quantity index
Laspeyres	Paasche
Paasche	Laspeyres
Fischer	Fischer
Unit	Ratio of total quantities
Normalized unit	Ratio of total normalized quantities

APPENDIX C - Program in PASCAL for calculating Tables 1, 2

```
PROGRAM tables(output);
TYPE
  real=double;
VAR
  v, ep, eq: integer;
  FUNCTION i(e, v: real): real;
   {Assumption: 0<= v < 1}
  BEGIN
   IF (v = 0) OR (e = 0)
    THEN i:=1
   ELSE IF (e = -1)
    THEN i:=(ln(1+v) - ln(1-v))/(2*v)
    ELSE i:=((1+v)**(e+1) - (1-v)**(e+1))/((e+1)*2*v);
  END;

BEGIN
 FOR ep := -2 TO 2 DO
   FOR eq := -2 TO ep DO
     IF (ep <> 0) AND (eq <> 0)
      THEN write('      ','(',ep:1,', ',eq:1,')');
 writeln;
 FOR ep := -2 TO 2 DO
   FOR eq := -2 TO ep DO
     IF (ep <> 0) AND (eq <> 0)
      THEN IF (ep <> eq)
            THEN write('','(',eq:1,', ',ep:1,')')
            ELSE write(' ');
 writeln;
 FOR v := 1 TO 9 DO
   BEGIN
    write(v*10:2);
    FOR ep := -2 TO 2 DO
      FOR eq := -2 TO ep DO
        IF (ep <> 0) AND (eq <> 0)
         THEN write('   ',((i(ep,v/10)*i(eq,v/10))/i(ep+eq,v/10) - 1)*100 :4:2);
    writeln;
   END;
 writeln;
 FOR ep := -2 TO 2 DO
   FOR eq := -2 TO ep DO
     IF (ep <> 0) AND (eq <> 0)
      THEN write('      ','(',ep:1,', ',eq:1,')');
 writeln;
 FOR ep := -2 TO 2 DO
   FOR eq := -2 TO ep DO
```

```
    IF (ep <> 0) AND (eq <> 0)
     THEN IF (ep <> eq)
           THEN write('','(',eq:1,', ',ep:1,')')
           ELSE write(' ');
writeln;
writeln;
FOR v := 1 TO 9 DO
  BEGIN
   write(v*10:2);
   FOR ep := -2 TO 2 DO
     FOR eq := -2 TO ep DO
       IF (ep <> 0) AND (eq <> 0)
        THEN write('    ',(i(ep+eq,v/10)/(i(ep,v/10)*i(eq,v/10)) - 1)*100 :4:2);
   writeln;
  END;
writeln;
END.
```