INDEX NUMBER THEORY AND MEASUREMENT ECONOMICS

By W.E. Diewert. April, 2004.

CHAPTER 11: Alternative Index Number Formulae using an Artificial Data Set

1. Introduction

In order to give the reader some idea of how much the various index numbers might differ using a "real" data set, we compute all of the major indexes defined in the previous chapters using an artificial data set consisting of prices and quantities for 6 commodities over 5 periods. The period can be thought of as somewhere between a year and 5 years. The trends in the data are generally more pronounced than one would see in the course of a year. The price and quantity data are listed in Tables 11.1 and 11.2 below. For convenience, we have also listed the period t nominal expenditures, $p^t \bullet q^t \equiv \sum_{i=1}^N p_i^t q_i^t$, along with the corresponding period t expenditure shares, $s_i^t \equiv p_i^t q_i^t / p^t \bullet q^t$, in Table 11.3.

Table 11.1 Prices for Six Commodities

Period t	p_1^t	p_2^t	p_3^t	p_4^t	p_5^t	p_6^t
1	1.0	1.0	1.0	1.0	1.0	1.0
2	1.2	3.0	1.3	0.7	1.4	0.8
3	1.0	1.0	1.5	0.5	1.7	0.6
4	0.8	0.5	1.6	0.3	1.9	0.4
5	1.0	1.0	1.6	0.1	2.0	0.2

Table 11.2 Quantities for Six Commodities

Period t	q_1^t	q_2^t	q_3^t	q_4^t	q_5^t	q_6^t
1	1.0	1.0	2.0	1.0	4.5	0.5
2	0.8	0.9	1.9	1.3	4.7	0.6
3	1.0	1.1	1.8	3.0	5.0	0.8
4	1.2	1.2	1.9	6.0	5.6	1.3
5	0.9	1.2	2.0	12.0	6.5	2.5

Table 11.3 Expenditures and Expenditure Shares for Six Commodities

Period t	$p^t \bullet q^t$	s_1^t	s_2^t	s_3^t	s_4^t	s_5^t	s_6^t
1	10.00	0.1000	0.1000	0.2000	0.1000	0.4500	0.0500
2	14.10	0.0681	0.1915	0.1752	0.0645	0.4667	0.0340
3	15.28	0.0654	0.0720	0.1767	0.0982	0.5563	0.0314
4	17.56	0.0547	0.0342	0.1731	0.1025	0.6059	0.0296
5	20.00	0.0450	0.0600	0.1600	0.0600	0.6500	0.0250

We will explain the trends that are built into the above tables. Think of the first 4 commodities as the consumption of various classes of *goods* in some economy while the last two commodities are the consumption of two classes of *services*. Think of the first good as *agricultural consumption*, which fluctuates around 1 and its price also fluctuates around 1. The quantity of the second good is *energy consumption* which trends up gently during the five periods with some minor fluctuations. However, note that the price of energy fluctuates wildly from period to period.[1] The third good is *traditional manufactures*. We have built in rather high inflation rates for this commodity for periods 2 and 3 which diminishes to a very low inflation rate by the end of our sample period.[2] The consumption of traditional manufactured goods is more or less static in our data set. The fourth commodity is *high technology manufactured goods*; e.g., computers, video cameras, compact disks, etc. We have the demand for these high tech commodities growing twelve times over our sample period while the final period price is only one tenth of the first period price. The fifth commodity is *traditional services*. The price trends for this commodity are similar to traditional manufactures, except that the period to period inflation rates are a bit higher. However, we have the demand for traditional services growing much more strongly than for traditional manufactures. Our final commodity is *high technology services*; e.g., telecommunications, wireless phones, internet services, stock market trading, etc. For this final commodity, we have the price trending downward very strongly to end up at 20% of the starting level while demand increases fivefold. The movements of prices and quantities in this artificial data set are more pronounced than the year to year movements that would be encountered in a typical country but they do illustrate the problem that is facing compilers of the Consumer Price Index; namely, *year to year price and quantity movements are far from being proportional across commodities so the choice of index number formula will matter*.

Every price statistician is familiar with the *Laspeyres index* P_L and the *Paasche index* P_P defined in chapter 1. We list these indexes in Table 11.4 along with the two unweighted indexes that we considered in the previous chapter: the *Carli index* defined by (2) and the *Jevons index* defined by (3) in chapter 10. The indexes in Table 11.4 compare the prices in period t with the prices in period 1; i.e., they are *fixed base indexes*. Thus the period t entry for the Carli index, P_C, is simply the arithmetic mean of the 6 price relatives, $\sum_{i=1}^{6} (1/6)(p_i^t/p_i^1)$, while the period t entry for the Jevons index, P_J, is the geometric mean of the 6 price relatives, $\prod_{i=1}^{6} (p_i^t/p_i^1)^{1/6}$.

Table 11.4 The Fixed Base Laspeyres, Paasche, Carli and Jevons Indexes

Period t	P_L	P_P	P_C	P_J
1	1.0000	1.0000	1.0000	1.0000
2	1.4200	1.3823	1.4000	1.2419
3	1.3450	1.2031	1.0500	0.9563
4	1.3550	1.0209	0.9167	0.7256
5	1.4400	0.7968	0.9833	0.6324

Note that by period 5, the spread between the fixed base Laspeyres and Paasche price indexes is enormous: P_L is equal to 1.4400 while P_P is 0.7968*, a spread of about 81%*. Since both of these indexes have exactly the same *theoretical* justification, it can be seen that the choice of index number formula matters a lot. The period 5 entry for the

[1] This is an example of the price bouncing phenomenon noted by Szulc (1983). Note that the fluctuations in the price of energy that we have built into our data set are not that unrealistic: in the past 3 years, the price of a barrel of crude oil has fluctuated in the range $10 to $37 U.S.

[2] This corresponds roughly to the experience of most industrialized countries over the period starting in 1973 to the mid 1990's. Thus we are compressing roughly 5 years of price movement into one of our periods.

Carli index, 0.98333, falls between the corresponding Paasche and Laspeyres indexes but the period 5 Jevons index, 0.63246, does not. Note that the Jevons index is always considerably below the corresponding Carli index. This will always be the case (unless prices are proportional in the two periods under consideration) because a geometric mean is always equal to or less than the corresponding arithmetic mean.[3]

It is of interest to recalculate the 4 indexes listed in Table 11.4 above using *the chain principle* rather than the *fixed base principle*. Our expectation is that the spread between the Paasche and Laspeyres indexes will be reduced by using the chain principle. These chain indexes are listed in Table 11.5.

Table 11.5 Chain Laspeyres, Paasche, Carli and Jevons Indexes

Period t	P_L	P_P	P_C	P_J
1	1.0000	1.0000	1.0000	1.0000
2	1.4200	1.3823	1.4000	1.2419
3	1.3646	1.2740	1.1664	0.9563
4	1.3351	1.2060	0.9236	0.7256
5	1.3306	1.1234	0.9446	0.6325

It can be seen comparing Tables 11.4 and 11.5 that chaining eliminated about 2/3 of the spread between the Paasche and Laspeyres indexes. However, even the chained Paasche and Laspeyres indexes differ by about 18% in period 5 so the choice of index number formula still matters. Note that chaining did not affect the Jevons index. This is an advantage of the index but the lack of weighting is a fatal flaw.[4] We would expect the "truth" to lie between the Paasche and Laspeyres indexes and from Table 11.5, we see that the unweighted Jevons index is far below this acceptable range. Note that chaining did not affect the Carli index in a systematic way for our particular data set: in periods 3 and 4, the chained Carli is above the corresponding fixed base Carli but in period 5, the chained Carli is below the fixed base Carli.[5]

2. Asymmetrically Weighted Indexes

We turn now to a systematic comparison of all of *the asymmetrically weighted price indexes* (with the exception of the Lloyd Moulton index which we will consider later). The *fixed base indexes* are listed in Table 11.6. The fixed base *Laspeyres* and *Paasche indexes*, P_L and P_P, are the same as those indexes listed in Table 11.4 above.

---

[3] This is the Theorem of the Arithmetic and Geometric Mean; see Hardy, Littlewood and Polyá (1934).

[4] The problem with the evenly weighted geometric mean is that the price declines in high technology goods and services are given the same weighting as the price changes in the other 4 commodities (which have rising or stationary price changes) but the expenditure shares of the high technology commodities remain rather small throughout the 5 periods. Thus weighted price indices do not show the rate of overall price decrease that the unweighted Jevons index shows. These somewhat negative comments on the use of the unweighted geometric mean as an index number formula at higher levels of aggregation do not preclude its use at the very lowest level of aggregation where a strong axiomatic justification for the use of this formula can be given. If probability sampling is used at the lowest level of aggregation, then the unweighted geometric mean essentially becomes the logarithmic Laspeyres index.

[5] For many data sets, we would expect the chained Carli to be above the corresponding fixed base Carli; see Szulc (1983).

The *Palgrave index*, P_{PAL}, was defined by equation (53) in Chapter 1.[6] The indexes denoted by P_{GL} and P_{GP} are *the geometric Laspeyres and geometric Paasche indexes*[7] which are special cases of the fixed weight geometric indexes defined by Konüs and Byushgens. For *the geometric Laspeyres index*, P_{GL}, we let the weights α_i be the *base period expenditure shares*, s_i^1. This index should be considered an alternative to the fixed base Laspeyres index since each of these indexes makes use of the same information set. For *the geometric Paasche index*, P_{GP}, we let the weights α_i be the *current period expenditure shares*, s_i^t. Finally, the index P_{HL} *is the harmonic Laspeyres index* that was defined by (57) in Chapter 1.

Table 11.6 Asymmetrically Weighted Fixed Base Indexes

Period t	P_{PAL}	P_L	P_{GP}	P_{GL}	P_P	P_{HL}
1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	1.6096	1.4200	1.4846	1.3300	1.3824	1.2542
3	1.4161	1.3450	1.3268	1.2523	1.2031	1.1346
4	1.5317	1.3550	1.3282	1.1331	1.0209	0.8732
5	1.6720	1.4400	1.4153	1.0999	0.7968	0.5556

By looking at the period 5 entries in Table 11.6, it can be seen that the spread between all of these fixed base asymmetrically weighted indexes has increased to be even larger than our earlier spread of 81% between the fixed base Paasche and Laspeyres indexes. In Table 11.6, the period 5 Palgrave index is about 3 times as big as the period 5 harmonic Laspeyres index, P_{HL}! Again, *this illustrates the point that due to the nonproportional growth of prices and quantities in most economies today, the choice of index number formula is very important.*

It is possible to explain why certain of the indexes in Table 11.6 are bigger than others. It can be shown that a *weighted arithmetic mean* of N numbers is equal to or greater than the corresponding *weighted geometric mean* of the same N numbers which in turn is equal to or greater than the corresponding *weighted harmonic mean* of the same N numbers.[8] It can be seen that the three indexes P_{PAL}, P_{GP} and P_P all use the current period expenditure shares s_i^t to weight the price relatives (p_i^t/p_i^1) but P_{PAL} is a weighted *arithmetic* mean of these price relatives, P_{GP} is a weighted *geometric* mean of these price relatives and P_P is a weighted *harmonic* mean of these price relatives. Thus by Schlömilch's inequality, we must have:[9]

$$(1)\quad P_{PAL} \geq P_{GP} \geq P_P .$$

Viewing Table 11.6, it can be seen that the inequalities (1) hold for each period. It can also be verified that the three indexes P_L, P_{GL} and P_{HL} all use the base period expenditure shares s_i^1 to weight the price relatives (p_i^t/p_i^1) but P_L is a weighted *arithmetic* mean of these price relatives, P_{GL} is a weighted *geometric* mean of these

[6] The Palgrave index is a share weighted arithmetic average of the price relatives but using the weights of period t.
[7] Vartia (1978; 272) uses the terms *logarithmic Laspeyres* and *logarithmic Paasche* respectively.
[8] This follows from Schlömilch's (1858) inequality; see Hardy, Littlewood and Polyá (1934).
[9] These inequalities were noted by Fisher (1922; 92) and Vartia (1978; 278).

price relatives and P_{HL} is a weighted *harmonic* mean of these price relatives. Thus by Schlömilch's inequality, we must have:[10]

(2) $P_L \geq P_{GL} \geq P_{HL}$.

Viewing Table 11.6, it can be seen that the inequalities (2) hold for each period.

We continue with our systematic comparison of all of *the asymmetrically weighted price indexes*. These indexes using the *chain principle* are listed in Table 11.7.

Table 11.7 Asymmetrically Weighted Indexes Using the Chain Principle

Period t	P_{PAL}	P_L	P_{GP}	P_{GL}	P_P	P_{HL}
1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	1.6096	1.4200	1.4846	1.3300	1.3824	1.2542
3	1.6927	1.3646	1.4849	1.1578	1.2740	0.9444
4	1.6993	1.3351	1.4531	1.0968	1.2060	0.8586
5	1.7893	1.3306	1.4556	1.0266	1.1234	0.7299

Viewing Table 11.7, it can be seen that although the use of the chain principle dramatically reduced the spread between the Paasche and Laspeyres indexes P_P and P_L compared to the corresponding fixed base entries in Table 11.6, the spread between the highest and lowest asymmetrically weighted indexes in period 5 (the Palgrave index P_{PAL} and P_{HL}) did not fall as much: the fixed base spread was 1.6720/0.5556 = 3.01 while the corresponding chain spread was 1.7893/0.7299 = 2.45. *Thus in this particular case, the use of the chain principle combined with the use of an index number formula that uses the weights of only one of the two periods being compared did not lead to a significant narrowing of the huge differences that these formulae generate using the fixed base principle. However, with respect to the Paasche and Laspeyres formulae, we find that chaining does significantly reduce the spread between these two indexes.*

Is there an explanation for the results reported in the previous paragraph? It can be shown that all 6 of the indexes that are found in the inequalities (1) and (2) approximate each other to the first order around an equal prices and quantities point. Thus with smooth trends in the data, we would expect all of the chain indexes to more closely approximate each other than the fixed base indexes because the changes in the individual prices and quantities would be smaller using the chain principle. This expectation is realized in the case of the Paasche and Laspeyres indexes but not with the others. However, for some of the commodities in our data set, the trends in the prices and quantities are not smooth. In particular, the prices for our first two commodities (agricultural products and oil) bounce up and down. As noted by Szulc (1983), this will tend to cause the chain indexes to have a wider dispersion than their fixed base counterparts. In order to determine if it is the bouncing prices problem that is causing some of the chained indexes in Table 11.7 to diverge from their fixed base counterparts, we recomputed all of the indexes in Tables 11.6 and 11.7 but excluding commodities 1 and 2 from the computations. The results of excluding these bouncing commodities may be found in Tables 11.8 and 11.9.

[10] These inequalities were also noted by Fisher (1922; 92) and Vartia (1978; 278).

Table 11.8 Asymmetrically Weighted Fixed Base Indexes for Commodities 3-6

Period t	P_{PAL}	P_L	P_{GP}	P_{GL}	P_P	P_{HL}
1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	1.2877	1.2500	1.2621	1.2169	1.2282	1.1754
3	1.4824	1.4313	1.3879	1.3248	1.2434	1.1741
4	1.6143	1.5312	1.4204	1.3110	1.0811	0.9754
5	1.7508	1.5500	1.4742	1.1264	0.7783	0.5000

Table 11.9 Asymmetrically Weighted Chained Indexes for Commodities 3-6

Period t	P_{PAL}	P_L	P_{GP}	P_{GL}	P_P	P_{HL}
1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	1.2877	1.2500	1.2621	1.2169	1.2282	1.1754
3	1.4527	1.4188	1.4029	1.3634	1.3401	1.2953
4	1.5036	1.4640	1.4249	1.3799	1.3276	1.2782
5	1.4729	1.3817	1.3477	1.2337	1.1794	1.0440

It can be seen that excluding the bouncing price commodities does cause the chain indexes to have a much narrower spread than their fixed base counterparts. *Thus our conclusion is that if the underlying price and quantity data is subject to reasonably smooth trends over time, then the use of chain indexes will narrow considerably the dispersion in the asymmetrically weighted indexes.* We now turn our attention to index number formulae that use weights from both periods in a symmetric or even handed manner.

3. Symmetrically Weighted Indexes

Symmetrically weighted indexes can be decomposed into two classes: *superlative indexes* and *other symmetrically weighted indexes*. Superlative indexes have a close connection to economic theory; i.e., as we saw in Chapter 4, a superlative index is exact for a representation of the consumer's preference function or the dual unit cost function that can provide a second order approximation to arbitrary (homothetic) preferences. We considered the following 4 superlative indexes in Chapter 4:

- the *Fisher ideal price index* P_F;
- the *Walsh price index* P_W defined by (20) in chapter 3 (this price index also corresponds to the quantity index Q^1 defined by (44) in chapter 4);
- the *Törnqvist-Theil price index* P_T defined by (45) in chapter 1 and
- the *implicit Walsh price index* P_{IW} defined by (49) in chapter 4.

These 4 symmetrically weighted superlative price indexes are listed in Table 11.8 using the fixed base principle. We also list in this table two symmetrically weighted (but not superlative) price indexes:[11]

[11] Diewert (1978; 897) showed that the Drobisch (1871; 423-425) Sidgwick (1883; 68) Bowley (1901; 227) price index approximates any superlative index to the second order around an equal price and quantity point; i.e., P_{SB} is a *pseudo-superlative index*. Straightforward computations show that the Marshall Edgeworth index P_{ME} is also pseudo-superlative.

- the Marshall Edgeworth price index P_{ME} defined by (19) in chapter 3 and
- the Drobisch Sidgwick Bowley price index P_D (the arithmetic average of the Paasche and Laspeyres indexes) defined above (25) in chapter 1.

Table 11.10 Symmetrically Weighted Fixed Base Indexes

Period t	P_T	P_{IW}	P_W	P_F	P_D	P_{ME}
1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	1.4052	1.4015	1.4017	1.4011	1.4012	1.4010
3	1.2890	1.2854	1.2850	1.2721	1.2741	1.2656
4	1.2268	1.2174	1.2193	1.1762	1.1880	1.1438
5	1.2477	1.2206	1.1850	1.0712	1.1184	0.9801

Note that the Drobisch index P_D is always equal to or greater than the corresponding Fisher index P_F. This follows from the facts that the Fisher index is the geometric mean of the Paasche and Laspeyres indexes while the Drobisch index is the arithmetic mean of the Paasche and Laspeyres indexes and an arithmetic mean is always equal to or greater than the corresponding geometric mean. Comparing the fixed base asymmetrically weighted indexes, Table 11.6, with the symmetrically weighted indexes, Table 11.10, *it can be seen that the spread between the lowest and highest index in period 5 is much less for the symmetrically weighted indexes*. The spread was 1.6720/0.5556 = 3.01 for the asymmetrically weighted indexes but only 1.2477/0.9801 = 1.27 for the symmetrically weighted indexes. If we restrict ourselves to the superlative indexes listed for period 5 in Table 11.10, then this spread is further reduced to 1.2477/1.0712 = 1.16; i.e., the spread between the fixed base superlative indices is "only" 16% compared to the fixed base spread between the Paasche and Laspeyres indexes of 81% (1.4400/0.7968 = 1.81). We expect to further reduce the spread between the superlative indexes by using the chain principle.

We recompute the symmetrically weighted indexes using the chain principle. The results may be found in Table 11.11.

Table 11.11 Symmetrically Weighted Indexes Using the Chain Principle

Period t	P_T	P_{IW}	P_W	P_F	P_D	P_{ME}
1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	1.4052	1.4015	1.4017	1.4011	1.4012	1.4010
3	1.3112	1.3203	1.3207	1.3185	1.3193	1.3165
4	1.2624	1.2723	1.2731	1.2689	1.2706	1.2651
5	1.2224	1.2333	1.2304	1.2226	1.2270	1.2155

A quick glance at Table 11.11 shows that *the combined effect of using both the chain principle as well as symmetrically weighted indices is to dramatically reduce the spread between all indexes constructed using these two principles*. The spread between all of the symmetrically weighted indexes in period 5 is only 1.2333/1.2155 = 1.015 or 1.5% and the spread between the 4 superlative indexes in period 5 is an even smaller 1.2333/1.2224 = 1.009 or about 0.1%. The spread in period 5 between the two most commonly used superlative indexes, the Fisher P_F and the Törnqvist P_T, is truly tiny: 1.2226/1.2224 = 0.0002.[12]

The results listed in Table 11.11 reinforce the numerical results tabled in Hill (2004) and Diewert (1978; 894): *the most commonly used chained superlative indexes will generally give approximately the same numerical results*.[13] In particular, the chained Fisher, Törnqvist and Walsh indexes will generally approximate each other very closely.

4. Two Stage Aggregation

We now turn our attention to the differences between superlative indexes and their counterparts that are constructed in two stages of aggregation; see chapter 9 above for a discussion of the issues and a listing of the formulae used. In our artificial data set, we will first aggregate the first 4 commodities into a *goods aggregate* and the last 2 commodities into a *services aggregate*. In the second stage of aggregation, the goods and services components will be aggregated into an all items index.

We report the results in Table 11.12 for our two stage aggregation procedure using period 1 as the *fixed base* for the Fisher index P_F, the Törnqvist index P_T and the Walsh and implicit Walsh indexes, P_W and P_{IW}.

Table 11.12 Fixed Base Superlative Single Stage and Two Stage Indexes

Period t	P_F	P_{F2S}	P_T	P_{T2S}	P_W	P_{W2S}	P_{IW}	P_{IW2S}
1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	1.4011	1.4004	1.4052	1.4052	1.4017	1.4015	1.4015	1.4022
3	1.2721	1.2789	1.2890	1.2872	1.2850	1.2868	1.2854	1.2862
4	1.1762	1.2019	1.2268	1.2243	1.2193	1.2253	1.2174	1.2209
5	1.0712	1.1286	1.2477	1.2441	1.1850	1.2075	1.2206	1.2240

Viewing Table 11.12, it can be seen that the fixed base single stage superlative indexes generally approximate their fixed base two stage counterparts fairly closely with the exception of the Fisher formula. The divergence between the single stage Fisher index P_F and its two stage counterpart P_{F2S} in period 5 is 1.1286/1.0712 = 1.05 or 5%. The other divergences are 2% or less.

Using *chain indexes*, we report the results in Table 11.13 for our two stage aggregation procedure. Again, the single stage and their two stage counterparts are listed for the Fisher index P_F, the Törnqvist index P_T and the Walsh and implicit Walsh indexes, P_W and P_{IW}.

Table 11.13 Chained Superlative Single Stage and Two Stage Indexes

Period t	P_F	P_{F2S}	P_T	P_{T2S}	P_W	P_{W2S}	P_{IW}	P_{IW2S}

[12] However, in other periods the differences were larger. On average over the last 4 periods, the chain Fisher and the chain Törnqvist indexes differed by 0.0025 percentage points.

[13] More precisely, the superlative quadratic mean of order r price indexes P^r defined and the implicit quadratic mean of order r price indexes P^{r*} defined in chapter 4 will generally closely approximate each other provided that r is in the interval $0 \leq r \leq 2$.

1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	1.4011	1.4004	1.4052	1.4052	1.4017	1.4015	1.4015	1.4022
3	1.3185	1.3200	1.3112	1.3168	1.3207	1.3202	1.3203	1.3201
4	1.2689	1.2716	1.2624	1.2683	1.2731	1.2728	1.2723	1.2720
5	1.2226	1.2267	1.2224	1.2300	1.2304	1.2313	1.2333	1.2330

Viewing Table 11.13, it can be seen that *the chained single stage superlative indexes generally approximate their fixed base two stage counterparts very closely indeed.* The divergence between the chained single stage Törnqvist index P_T and its two stage counterpart P_{T2S} in period 5 is 1.2300/1.2224 = 1.006 or 0.6%. The other divergences are all less than this. Given the large dispersion in period to period price movements, these two stage aggregation errors are not large.

5. Lloyd Moulton Indexes

The next formula that we illustrate using our artificial data set is the Lloyd Moulton index P_{LM} defined by (68) in chapter 4. Recall that this formula requires an estimate for the parameter σ, *the elasticity of substitution* between all commodities being aggregated. Recall also that if σ equals 0, then the Lloyd Moulton index collapses down to the ordinary Laspeyres index, P_L. When σ equals 1, the Lloyd Moulton index is not defined but it can be shown that the limit of P_{LM_σ} as σ approaches 1 is P_{GL}, the geometric Laspeyres index or the logarithmic Laspeyres index with base period shares as weights. This index uses the same basic information as the fixed base Laspeyres index P_L and so it is a possible alternative index for CPI compilers to use. As was shown by Shapiro and Wilcox (1997)[14], the Lloyd Moulton index may be used to approximate a superlative index using the same information that is used in the construction of a fixed base Laspeyres index provided that we have an estimate for the parameter σ. We will test this methodology out using our artificial data set. The superlative index that we choose to approximate is the chain Fisher index[15] (which approximates the other chained superlative indexes listed in Table 11.11 very closely). The chained Fisher index P_F is listed in column 2 of Table 11.14 along with the fixed base Lloyd Moulton indexes P_{LM_σ} for σ equal to 0 (this reduces to the fixed base Laspeyres index P_L), 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 and 1 (which is the fixed base geometric index P_{GL}). Note that the Lloyd Moulton indexes steadily *decrease* as we *increase* the elasticity of substitution σ.[16]

Table 11.14 Fixed Base Fisher and Fixed Base Lloyd Moulton Indexes

Period	P_F	P_{LM0}	$P_{LM.2}$	$P_{LM.3}$	$P_{LM.4}$	$P_{LM.5}$	$P_{LM.6}$	$P_{LM.7}$	$P_{LM.8}$	P_{LM1}
1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	1.4011	1.4200	1.4005	1.3910	1.3818	1.3727	1.3638	1.3551	1.3466	1.3300
3	1.3185	1.3450	1.3287	1.3201	1.3113	1.3021	1.2927	1.2831	1.2731	1.2523
4	1.2689	1.3550	1.3172	1.2970	1.2759	1.2540	1.2312	1.2077	1.1835	1.1331

[14] Alterman, Diewert and Feenstra (1999) also used this methodology in the context of estimating superlative international trade price indices.

[15] Since there is still a considerable amount of dispersion among the fixed base superlative indices and practically no dispersion between the chained superlative indices, we take the Fisher chain index as our target rather than any of the fixed base superlative indices.

[16] This follows from Schlömilch's (1858) inequality again.

5	1.2226	1.4400	1.3940	1.3678	1.3389	1.3073	1.2726	1.2346	1.1932	1.0999

Viewing Table 11.14, it can be seen that *no single choice* of the elasticity of substitution σ will lead to a Lloyd Moulton price index $P_{LM\sigma}$ that will *closely* approximate the chained Fisher index P_F for periods 2,3,4 and 5. To approximate P_F in period 2, we should choose σ close to 0.1; to approximate P_F in period 3, we should choose σ close to 0.3; to approximate P_F in period 4, we should choose σ between 0.4 and 0.5 and to approximate P_F in period 5, we should choose σ between 0.7 and 0.8.[17]

We repeat the computations for the Lloyd Moulton indexes that are listed in Table 11.14 except that we now use the *chain principle* to construct the Lloyd Moulton indexes; see Table 11.15. Again, we are trying to approximate the *chained Fisher price index* P_F which is listed as column 2 in Table 11.15. In Table 11.15, P_{LM0} is the chained Laspeyres index and P_{LM1} is the chained geometric Laspeyres or geometric index using the expenditure shares of the previous period as weights.

Table 11.15 Chained Fisher and Chained Lloyd Moulton Indexes

Period	P_F	P_{LM0}	$P_{LM.2}$	$P_{LM.3}$	$P_{LM.4}$	$P_{LM.5}$	$P_{LM.6}$	$P_{LM.7}$	$P_{LM.8}$	P_{LM1}
1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	1.4011	1.4200	1.4005	1.3910	1.3818	1.3727	1.3638	1.3551	1.3466	1.3300
3	1.3185	1.3646	1.3242	1.3039	1.2834	1.2628	1.2421	1.2212	1.2002	1.1578
4	1.2689	1.3351	1.2882	1.2646	1.2409	1.2171	1.1932	1.1692	1.1452	1.0968
5	1.2226	1.3306	1.2702	1.2400	1.2097	1.1793	1.1488	1.1183	1.0878	1.0266

Viewing Table 11.15, it can be seen that again *no single choice* of the elasticity of substitution σ will lead to a Lloyd Moulton price index $P_{LM\sigma}$ that will *closely* approximate the chained Fisher index P_F for all periods. To approximate P_F in period 2, we should choose σ close to 0.1; to approximate P_F in period 3, we should choose σ close to 0.2; to approximate P_F in period 4, we should choose σ between 0.2 and 0.3 and to approximate P_F in period 5, we should choose σ between 0.3 and 0.4. However, it should be noted that *if we choose σ equal to 0.3 and use the chained Lloyd Moulton index $P_{LM.3}$ to approximate the chained Fisher index P_F, we will have a much better approximation than that provided by either the chained Laspeyres index* (see P_{LM0} in column 3 of Table 11.15) *or the fixed base Laspeyres index* (see P_{LM0} in column 3 of Table 11.14).[18] Hence our tentative conclusions on the use of the Lloyd Moulton index to approximate superlative indexes are:

- the elasticity of substitution parameter σ which appears in the Lloyd Moulton formula is *unlikely to remain constant over time* and hence it will be necessary for statistical agencies to update their estimates of σ at regular intervals and

[17] Unfortunately, for this data set, neither the fixed base Laspeyres index $P_L = P_{LM0}$ nor the fixed base weighted geometric index $P_{GL} = P_{LM1}$ are very close to the chain Fisher index for all periods. For less extreme data sets, the fixed base Laspeyres and fixed base geometric indexes will be closer to the chained Fisher index.

[18] For this particular data set, the fixed base or chained geometric indexes using either the expenditure weights of period 1 (see the last column of Table 11.14) or using the weights of the previous period (see the last column of Table 11.15) do not approximate the chained Fisher index very closely. However, for less extreme data sets, the use of chained Laspeyres or geometric indexes may approximate a chained superlative index adequately.

- the use of the Lloyd Moulton index as a *real time preliminary estimator* for a chained superlative index seems warranted, provided that the statistical agency can provide estimates for chained superlative indexes on a delayed basis. *The Lloyd Moulton index would provide a useful supplement to the traditional fixed base Laspeyres price index*.

6. Additive Decompositions for the Fisher Ideal Index

The final formulae we illustrate using our artificial data set are the *additive percentage change decompositions* for the Fisher ideal index that were discussed in section 9 of chapter 3. We will first decompose the *chain links* for the Fisher price index using the formulae (44) to (46) in chapter 3. The results of the decomposition are listed in Table 11.16. Thus $P_F - 1$ is *the percentage change in the Fisher ideal chain link* going from period t − 1 to t and *the decomposition factor* $v_{Fi}\Delta p_i = v_{Fi}\,(p_i^t - p_i^{t-1})$ is the contribution to the total percentage change of the change in the ith price from p_i^{t-1} to p_i^t for i = 1,2,…,6.

Table 11.16 An Additive Percentage Change Decomposition of the Fisher Index

Period t	$P_F - 1$	$v_{F1}\Delta p_1$	$v_{F2}\Delta p_2$	$v_{F3}\Delta p_3$	$v_{F4}\Delta p_4$	$v_{F5}\Delta p_5$	$v_{F6}\Delta p_6$
2	0.4011	0.0176	0.1877	0.0580	-0.0351	0.1840	-0.0111
3	-0.0589	-0.0118	-0.1315	0.0246	-0.0274	0.0963	-0.0092
4	-0.0376	-0.0131	-0.0345	0.0111	-0.0523	0.0635	-0.0123
5	-0.0365	0.0112	0.0316	0.0000	-0.0915	0.0316	-0.0194

Viewing Table 11.16, it can be seen that the price index going from period 1 to 2 grew about 40% and the major contributors to this change were the increases in the price of commodity 2, energy (18.77%) and in commodity 5, traditional services (18.4%). The increase in the price of traditional manufactured goods, commodity 3, contributed 5.8% to the overall increase of 40.11%. The decreases in the prices of high technology goods (commodity 4) and high technology services (commodity 6) offset the other increases by −3.51% and −1.11% going from period 1 to 2. Going from period 2 to 3, the overall change in prices was negative: −5.89%. The reader can read across the third row of Table 6.16 to see what was the contribution of the 6 component price changes to the overall price change. It can be seen that a big price change in a particular component i combined with a big expenditure share in the two periods under consideration will lead to a big decomposition factor, v_{Fi}.

Our final set of computations we illustrate using our artificial data set is the *additive percentage change decomposition* for the Fisher ideal index that is due to Van IJzeren (1987; 6) that was mentioned in problem 9 of chapter 3. The *price* counterpart to the *additive decomposition* for a *quantity* index, equation (41) in chapter 3, is:

(3) $P_F(p^0,p^1,q^0,q^1) = \sum_{i=1}^N q_{Fi}^* p_i^1 / \sum_{i=1}^N q_{Fi}^* p_i^0$

where the reference quantities need to be defined somehow. Van IJzeren (1987; 6) showed that the following reference weights provided an *exact additive representation for the Fisher ideal price index*:

(4) $q_{Fi}^* \equiv (1/2)q_i^0 + (1/2)q_i^1/Q_F(p^0,p^1,q^0,q^1)$; $i = 1,2,\ldots,6$

where Q_F is the overall Fisher quantity index. Thus using the Van IJzeren quantity weights q_{Fi}^*, we obtain the *following Van IJzeren additive percentage change decomposition for the Fisher price index*:

$$(5)\quad P_F(p^0,p^1,q^0,q^1) - 1 = \{\sum_{i=1}^{6} q_{Fi}^* p_i^1 / \sum_{m=1}^{6} q_{Fi}^* p_m^0\} - 1$$
$$= \{\sum_{i=1}^{6} q_{Fi}^* p_i^1 - \sum_{m=1}^{6} q_{Fi}^* p_i^0\} / \sum_{m=1}^{6} q_{Fi}^* p_m^0$$
$$= \sum_{i=1}^{6} v_{Fi}^* \{p_i^1 - p_i^0\}$$

where the *Van IJzeren weight* for commodity i, v_{Fi}^*, is defined as

(6) $v_{Fi}^* \equiv q_{Fi}^* / \sum_{m=1}^{n} q_{Fi}^* p_m^0$; $i = 1,\ldots,6$.

We will again decompose the *chain links* for the Fisher price index using the formulae (4) to (6) listed above. The results of the decomposition are listed in Table 11.17. Thus $P_F - 1$ is *the percentage change in the Fisher ideal chain link* going from period $t - 1$ to t and *the Van IJzeren decomposition factor* $v_{Fi}^*\Delta p_i$ is the contribution to the total percentage change of the change in the ith price from p_i^{t-1} to p_i^t for $i = 1,2,\ldots,6$.

Table 11.17 Van IJzeren's Decomposition of the Fisher Price Index

Period t	$P_F - 1$	$v_{F1}^*\Delta p_1$	$v_{F2}^*\Delta p_2$	$v_{F3}^*\Delta p_3$	$v_{F4}^*\Delta p_4$	$v_{F5}^*\Delta p_5$	$v_{F6}^*\Delta p_6$
2	0.4011	0.0178	0.1882	0.0579	-0.0341	0.1822	-0.0109
3	-0.0589	-0.0117	-0.1302	0.0243	-0.0274	0.0952	-0.0091
4	-0.0376	-0.0130	-0.0342	0.0110	-0.0521	0.0629	-0.0123
5	-0.0365	0.0110	0.0310	0.0000	-0.0904	0.0311	-0.0191

Comparing the entries in Tables 11.16 and 11.17, it can be seen that the differences between the Diewert and Van IJzeren decompositions of the Fisher price index are *very small*. The maximum absolute difference between the $v_{Fi}\Delta p_i$ and $v_{Fi}^*\Delta p_i$ is only 0.0018 (about 0.2 percentage points) and the average absolute difference is 0.0003. This is somewhat surprising given the very different nature of the two decompositions.[19] As was mentioned in section 9 of chapter 3, the Van IJzeren decomposition of the chain Fisher *quantity* index is used by the Bureau of Economic Analysis in the U.S.

References

Alterman, W.F., W.E. Diewert and R.C. Feenstra, (1999), *International Trade Price Indexes and Seasonal Commodities*, Bureau of Labor Statistics, Washington D.C.

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

[19] However, Reinsdorf, Diewert and Ehemann (2002) show that the terms in the two decompositions approximate each other to the second order around any point where the two price vectors are equal and where the two quantity vectors are equal.

Drobisch, M. W. (1871), “Ueber einige Einwürfe gegen die in diesen Jahrbüchern veröffentlichte neue Methode, die Veränderungen der Waarenpreise und des Geldwerths zu berechten”, *Jahrbücher für Nationalökonomie und Statistik* 16, 416-427.

Fisher, I. (1922), *The Making of Index Numbers*, Houghton-Mifflin, Boston.

Hardy, G.H., J.E. Littlewood and G. Polyá (1934), *Inequalities*, Cambridge: Cambridge University Press.

Hill, R.J. (2004), “Constructing Price Indexes Across Space and Time: The Case of the European Union”, *American Economic Review*, forthcoming.

Lloyd, P.J. (1975), “Substitution Effects and Biases in Nontrue Price Indices”, *American Economic Review* 65, 301-313.

Moulton, B.R., and E.P. Seskin (1999), “A Preview of the 1999 Comprehensive Revision of the National Income and Product Accounts”, *Survey of Current Business* 79 (October), 6-17.

Reinsdorf, M.B., W.E. Diewert and C. Ehemann (2002), “Additive Decompositions for the Fisher, Törnqvist and Geometric Mean Indexes”, *Journal of Economic and Social Measurement* 28, 51-61.

Szulc, B.J. (1983), “Linking Price Index Numbers,” pp. 537-566 *in Price Level Measurement*, W.E. Diewert and C. Montmarquette (eds.), Ottawa: Statistics Canada.

Shapiro, M.D. and D.W. Wilcox (1997), “Alternative Strategies for Aggregating Prices in the CPI”, *Federal Reserve Bank of St. Louis Review* 79:3, 113-125.

Sidgwick, H. (1883), *The Principles of Political Economy*, London: Macmillan.

Van IJzeren (1987), *Bias in International Index Numbers: A Mathematical Elucidation*, Dissertation for the Hungarian Academy of Sciences, Den Haag: Koninklijke Bibliotheek.

Vartia, Y.O. (1978), “Fisher’s Five-Tined Fork and Other Quantum Theories of Index Numbers”, pp. 271-295 in *Theory and Applications of Economic Indices*, W. Eichhorn, R. Henn, O. Opitz and R.W. Shephard (eds.), Würzburg: Physica-Verlag.