INDEX NUMBERS:
A USERS' GUIDE

Nicholas Oulton

Centre for Economic Performance

London School of Economics

Version of September 2006

ABSTRACT

This is a guide to the principal types of index number and their main properties. The relationship between index numbers and concepts from economic theory like utility is covered. But the main focus is on the numerical behaviour of the indices: when will one type of index number grow more rapidly than another type? Quantity indices are given equal time with price indices. Differences between chain-linked and fixed base indices are analysed. Nine propositions about index number behaviour are presented and proved. Implications for macro modelling are discussed.

Contact address
Centre for Economic Performance
London School of Economics
Houghton Street
LONDON WC2A 2AE
UNITED KINGDOM
Email: n.oulton@lse.ac.uk

CONTENTS

1. Introduction ... 2

2. Approaches to index numbers ... 3

Axiomatic approach ... 3

Economic approach ... 5

3. Comparisons over two periods only ... 6

Formulas for Laspeyres and Paasche ... 8

Index numbers when some quantities can take zero values ... 10

Index numbers when some quantities can take negative values ... 11

4. Basic properties of fixed weight indices ... 11

5. Lower level indices ... 14

6. Chain indices: comparisons over more than two periods ... 15

The Divisia approach ... 15

Chain indices in practice ... 17

Properties of chain indices ... 18

Chain versus fixed base indices when expenditure shares are constant ... 21

Chained versus fixed base indices: the general relationship ... 22

7. Macro modelling implications of chain indices ... 22

8. Further reading ... 25

Theory ... 25

Practice ... 25

References ... 25

ANNEX A ... 28

ANNEX B ... 31

ANNEX C ... 41

ANNEX D ... 45

1. Introduction[1]

"For those who have made any attempt to penetrate their mysteries, index numbers seem to have a perennial fascination". Irving Fisher, *The Making of Index Numbers*, (3rd edition), 1927.

Not everyone will agree with Irving Fisher. But applied economists and analysts have to use index numbers almost every day of their working lives. They seem simple enough concepts: just weighted averages, which is hardly cutting edge mathematics. The fact that they seem so elementary perhaps explains why they are not commonly taught nowadays in graduate or even undergraduate economics degrees. But their simplicity is deceptive and their behaviour can sometimes be unexpected, even to the sophisticated. Hence this guide. Its purpose is to set out the main types of index numbers in common use and to give a concise account of their properties.

The emphasis is on explaining how index numbers behave in practice: for example, why might a chain index of output grow less rapidly than a fixed base index? Less time is spent on how index numbers are related to underlying economic concepts like utility, though this is covered in an Annex. Other omitted areas include important issues like possible bias in consumer prices and the general area of how index numbers are constructed in practice. Hence there is no discussion of quality adjustment, the merits of hedonic methods versus matched models, the use of scanner data, or sampling issues.

Subject to these limitations, the purpose is to set out the main results, together with simple proofs. Please note: this is a guide to results, not a survey of the literature, though some suggestions for further reading are given. Few if any of the results here are new. All are to be found scattered around in the literature. But no other concise guide seems to exist, so I hope that the present attempt will prove useful.

The guide is laid out as follows. The next section, section 2, discusses the two main approaches to index numbers, the axiomatic and the economic. The treatment here is brief, with the main results being more fully set out in Annexes C and D. Section 3 then introduces the four index numbers most often used in practice — Laspeyres, Paasche, Fisher and Törnqvist — in the simplest situation where a comparison is being made between two periods only. Section 4 discusses the basic properties of fixed weight indices and proves three propositions. Section 5

[1] The first version of this paper was written while I was employed by the Bank of England. I am grateful to colleagues at the Bank, in particular Ian Bond and Simon Price, for helpful comments and encouragement. I would particularly like to thank Erwin Diewert for taking the trouble to send me detailed comments on an earlier draft and correcting some of my errors. None of the above is responsible for any remaining errors.

looks at lower level indices, eg those used to average together the individual price quotes gathered by the price collectors for the U.K.'s Retail Prices Index (RPI) or Consumer Price Index (CPI).[2] Here somewhat different considerations apply since at this level the statistical agency usually has data only on prices and not on the quantities being purchased. Section 6 turns to the issues raised by the need to make comparisons over three or more periods. I consider first continuous time, Divisia indices, before setting out four propositions about discrete, chain indices. Section 7 considers the implications for macroeconomic modelling of the use of chain indices. Finally, section 8 provides a brief guide to further reading. Proofs of most of the main results will be found in Annexes A and B.

2. Approaches to index numbers

There are two main approaches to index numbers: (1) the so-called "test" or "axiomatic" approach and (2) the economic theory approach. Actually, there is also a third approach, the stochastic one, which lies behind notions like core inflation or trimmed means (Selvanathan and Prasada Rao 1994). This approach has been less influential than the other two. Also, as far as I know it has never been applied to output indices. So it will not be discussed further here.

Axiomatic approach

This approach is discussed in more detail in Annex C. Under the axiomatic approach one looks for an index which has "reasonable" properties. The basic idea of an output index is that its growth should be some sort of average of the growth rates of the constituent products. So, if we want an index of fruit output, where fruit is apples or bananas, we might require the following:

1. The growth rate of the index should lie between the growth rate of the quantity of apples and that of the quantity of bananas. If the quantities of apples and bananas are growing at the same rate, the index should growth at this common rate too.

2. The index should be more sensitive to products which are more "important".

---

[2] The RPI is used to adjust the value of inflation-proof government debt and social security payments. It was previously used too to measure the Bank of England's inflation target. The latter role is now played by the CPI, which is the U.K. version of Eurostat's Harmonised Index of Consumer Prices (HICP, commonly pronounced "hiccup").

This leads to the idea that the growth rate of the index should be a weighted average of the growth rates of its constituents, where the weights are the shares that each product has in the total value of the output of this group of products.

There are a few other requirements, such as that the index should be independent of the unit of quantity (lbs or kilos) and of the unit of currency (euros or centimes), but the best known index number formulas satisfy these tests. There are two further tests applicable to two period comparisons which are more demanding:

Product test. The product of the price index and the quantity index should be the expenditure index (the ratio of expenditure in period 2 to expenditure in period 1). For any single commodity, expenditure equals price times quantity. The product test simply generalises this notion to many commodities. A stronger form of the product test is the *factor reversal* test. The factor reversal test requires that the price and quantity indices should have the same functional form.

Time reversal test. Suppose the quantity and price vectors are (hypothetically) reversed, so that the quantities and prices observed in period 1 are now assumed to be observed in period 2, and those of period 2 are assumed to be observed in period 1. Then the new price and quantity indices should be the inverse of the old ones. Eg if with the original data the price index is 125 in period 2, then with the new data it should be [100 x (100/125) =] 80. This test is sometime misread to say that if prices and quantities first rise and then fall back to their original level, then the price or quantity index should return to its original value. But the latter is a different test (the circular test) which necessarily involves at least three time periods.

Since we never actually observe time running backwards, we will never observe a failure of the time reversal test. For an index number to fail this test indicates a conceptual flaw. Failure suggests that, in other situations that we might observe in practice, the index might behave in an undesirable or unacceptable way, though the test gives us no direct guidance on the nature of the misbehaviour.

Circular test This test can be interpreted as follows. Suppose that prices and quantities change after the first period, but then at some future date return to their first period values. Then the price or quantity index in this last period should equal its value in the first period. Note that this test assumes we are making comparisons over at least three periods. Unlike the time reversal test, we could in principle observe a failure of the circular test. And we may come quite close in practice to doing so, if we are concerned with monthly data. At monthly frequencies, due to seasonal factors, or special promotions, it is quite possible for prices to rise or fall but then return to their initial levels (this is known as "price bounce").

This test seems a very natural requirement, but it is in fact the most problematic. In general, fixed base indices satisfy the test but chain indices do not (see Proposition 7 below). As Annex C shows, there is a fundamental inconsistency between the product and the circular tests: one or other must be given up.

Economic approach

The economic approach is discussed in more detail in Annex D. As is well known, the consumer price index can be given an interpretation in terms of utility (see eg Deaton and Muellbauer 1980). The price index is the ratio of the minimum expenditure level at the new prices to the minimum expenditure level at the old prices, which would yield the same reference level of utility. In other words, if it costs a minimum of 100 euros to obtain 42 utils (the reference level) at the original set of prices and a minimum of 110 euros to obtain 42 utils at the new prices, then the price level has risen by 10%.

Utility is not the appropriate concept for output (input) quantity indices and the corresponding output (input) price indices. Here the central concept is the production frontier (function). We have two output (input) vectors, one for the base period, and the other for the comparison period, and we want to know: by how much has the production frontier (function) shifted between these two periods? (The shift may be due to changes in the quantities of inputs, to measure which we need an input quantity index, or it may be due to changes in technology or TFP, but this is taking us further afield).

In the economic approach, the issue becomes: which index number formula is likely to be the best approximation to the (usually unknown) shift in the production or utility frontiers? Diewert (1976) showed that there exist "flexible functional forms" which are good approximations (at least locally) to *any* production frontier (or productivity function or cost function or utility function) that is consistent with economic theory.[3] Also, some index numbers are "exact" in the following sense: if the frontier takes a particular form, then there is an index number which

[3] Flexible functional forms approximate to the second order, ie at a given point they have the same value and the same first and second derivatives as the function that they are approximating. But note that this does not necessarily mean that they approximate well the underlying function at some other point.

exactly measures shifts in such a frontier. An index number that is exact for a flexible functional form is called "superlative". For example, if the frontier is translog, then the Törnqvist quantity index (see below) is both exact and superlative.[4]

The point is that, if we believe that a particular flexible functional form is satisfactory, then we do not need to know its parameters to estimate the change in output or utility. We can simply employ the appropriate exact index number. This is important since in practice we will almost certainly have insufficient data to estimate the parameters of the production frontier or utility function econometrically. If there are n products, then a flexible functional form typically contains some $n(n+1)/2$ independent parameters. The U.K. Retail Prices Index (RPI) contains over six hundred items. To estimate all the parameters of a flexible functional form would require thousands of years of data. But there is no need to do this since we can calculate an exact index using just observed prices and quantities.[5]

3. Comparisons over two periods only

Here we want to compare output in some reference year with output in some later year. The following quantity indices are in common use and meet the most basic of the tests above, in the sense that they are all weighted averages of some sort:

Laspeyres quantity index (Q_t^L):

Paasche quantity index (Q_t^P)

Fisher quantity index (Q_t^F)

Törnqvist quantity index (Q_t^T)

In each case, the index number is for period t relative to period t-1. But the gap between the first and second periods could be of any length. Corresponding to each type of quantity index there is of course a price index. The formulas for both quantity and price indices are given in Annex A. The *base* of an index number is the period to which the prices (in the case of a quantity index) or the quantities (in the case of a price index) relate. The *reference* period is the period which is set

[4] Diewert (1976) is widely regarded as the most important single advance in index number theory of the last quarter century and must have been cited hundreds of times. It took over three years to publish and was rejected by the *Review of Economics and Statistics* and the *Quarterly Journal of Economics*, a fact that should give hope to us all (information from the Biographical Sketch on Diewert's home page).

equal to 1 or 100. The base is what determines the behaviour of the index number, while the reference period has no intrinsic importance: it can be chosen purely as a matter of presentation. Usually, but not invariable, the reference period is chosen to coincide with the base.

As is well known, the Laspeyres uses base (first) period weights, the Paasche uses current (second) period weights, while the Fisher and Törnqvist use the weights of both periods. This immediately suggests a limitation of Laspeyres and Paasche: why should one give a privileged position to the weights of either the first or the second period?

The Laspeyres and Paasche indices fail both the time reversal and product tests. The Törnqvist index passes the time reversal but fails the product test. Only the Fisher index passes both these tests.

To appreciate some of the issues here, it is helpful to consider the quantity index in conjunction with the corresponding price index. The Laspeyres *price* index answers the following question: what is the increase in expenditure which is needed in order to buy the *first* period's basket of products, when the buyer faces the *second* period's prices? The Paasche *price* index answers a different question: what is the reduction in expenditure which would occur if the buyer had to buy the *second* period's basket of goods but faced the *first* period's prices?

Let V be an index of expenditure, which can always be calculated without any conceptual difficulty as it is just the ratio of two nominal amounts:

$$V_t = \frac{\sum_i p_{it} q_{it}}{\sum_i p_{i,t-1} q_{i,t-1}} \tag{1}$$

Then we want the following to be true:

$$V_t = P_t \cdot Q_t \tag{2}$$

where P_t, Q_t are the price and quantity indices. Take the Laspeyres quantity index, which until recently was used in the U.K. to calculate constant price GDP for 1994 onwards with 1995 as the

---

[5] Of course in practice the RPI is calculated using a Laspeyres index. It is also chained, which raises other issues which are discussed below.

base. If we divide nominal GDP by GDP in constant prices, we get the GDP deflator. As shown in the Annex, the GDP deflator is a *Paasche* price index. So the GDP deflator for 1999 shows the growth in prices since 1995, using the weights of 1999. The GDP deflator for 2000 shows the growth of prices since 1995, *using the weights of 2000*. But then what does the growth of the GDP deflator between 1999 and 2000 actually mean, since different weights are involved in the comparison? This conceptual difficulty was alleviated, but not removed entirely, when the U.K.'s Office for National Statistics (ONS) moved to annual chain-linking in 2003 (see below).

A great attraction of Fisher price and quantity indices is that the product of a Fisher price index and a Fisher quantity index is the expenditure index, ie Fisher indices satisfy equation (2). The Fisher quantity (price) index is the geometric mean of the Laspeyres and Paasche quantity (price) indices, so like the Törnqvist it uses the weights of both periods. However Törnqvist price and quantity indices do not satisfy equation (2). But Törnqvist indices are attractive in other ways. First, by their nature they show the contribution of each component to the growth of the aggregate. Second, they correspond to translog cost and production functions which are relatively easy to estimate econometrically: the cost shares are linear functions of the logs of the prices or the quantities.

Formulas for Laspeyres and Paasche

The text book formulas for the Laspeyres and Paasche output indices are as follows:

$$Q_t^L = \frac{\sum_i p_{i,t-1} q_{it}}{\sum_i p_{i,t-1} q_{i,t-1}}$$

$$Q_t^P = \frac{\sum_i p_{it} q_{it}}{\sum_i p_{it} q_{i,t-1}} \qquad (3)$$

By interchanging prices and quantities, we get the Laspeyres and Paasche *price* indices (see Annex A). These formulas are important in theory but don't correspond to how the index numbers are calculated in practice. Normally, statistical agencies don't have access to individual prices and quantities at the level at which the indices are calculated. For example, the U.K. Retail Prices Index (and the EU's Harmonised Index of Consumer Prices) uses a Laspeyres formula but this is applied at a level of aggregation above that of individual price quotes. What the statistical agency does have is average prices for some type of product, say apples, and household budget shares, eg the proportion of the budget devoted to apples. It is therefore useful to express the formulas for Laspeyres and Paasche in terms of expenditure shares. In the case of the Laspeyres, a bit of manipulation produces

$$Q_t^L = \sum_i \frac{p_{i,t-1}q_{i,t-1}}{\sum_i p_{i,t-1}q_{i,t-1}} \frac{q_{it}}{q_{i,t-1}} = \sum_i w_{i,t-1} \frac{q_{it}}{q_{i,t-1}} \tag{4}$$

where $w_{i,t-1}$ is the expenditure share in period t-1:

$$w_{i,t-1} = \frac{p_{i,t-1}q_{i,t-1}}{\sum_i p_{i,t-1}q_{i,t-1}} \tag{5}$$

and $q_{it} / q_{i,t-1}$ is called the *quantity relative* (also known to macroeconomists as a gross growth rate). So the Laspeyres turns out to be a weighted average of quantity relatives, where the weights are expenditure shares in the *first* (base) period. Note however that this interpretation depends on all quantities in the base period being non-zero. If one or more base period quantities is zero, then equation (4) cannot be evaluated. But equation (3) can still be evaluated as long as at least one base period quantity is non-zero, which will obviously always be the case in practice. We return to this point below.

One might guess that the Paasche would also turn out to be a weighted average of quantity relatives, but with *last* (current) period shares as the weights. Actually, this is *not* the case. A bit of manipulation produces the following for the Paasche:

$$Q_t^P = \sum_i \left[\frac{p_{it}q_{i,t-1}}{\sum_i p_{it}q_{i,t-1}} \frac{q_{it}}{q_{i,t-1}} \right]$$

Here the weights are not the observed expenditure shares in any period but rather the first period shares revalued to second period prices.[6] Clearly, these are not observed. There *is* a formula for the Paasche in terms of observed, last period shares, but it looks rather different:

$$Q_t^P = 1 / \left[\sum_i w_{it} \left(\frac{q_{i,t-1}}{q_{it}} \right) \right]$$

[6] This means that the weights in this second formula for the Paasche index will not sum to 1.

(This can be established by straightforward algebra). Note that the quantity relatives are the other way up, $q_{i,t-1}/q_t$, *not* $q_{it}/q_{i,t-1}$, which is why we take the inverse of the weighted quantity relatives in this formula.[7]

By definition, the Fisher index is the geometric mean of Laspeyres and Paasche, and the Törnqvist index is related to them too. As shown in Annex A, the Törnqvist quantity index is approximately equal to

$$Q_t^T \approx \sum_i \left[\frac{w_{it} + w_{i,t-1}}{2}\right] \cdot \left(\frac{q_{it}}{q_{i,t-1}}\right)$$

So it resembles Laspeyres except than an average of the expenditure shares in the two periods are used, rather than the shares in the base period only.

Index numbers when some quantities can take zero values

It is reasonable to take prices as strictly positive. A price may appear to be zero, such as the price of treatment under the U.K.'s National Health Service, but closer examination shows that, properly defined to include the patient's cost of time, the price is positive. But it is quite common in empirical work to find that some quantities are zero in some time periods, at least at a sufficiently disaggregated level. For example, a consumer may buy a trip to Florida in 2006 but not in 2005, or a firm may purchase a component from a foreign source in one year, but from a domestic source in another, ie imports of this component are positive in one year, zero in another. This may create problems for quantity indices. The Törnqvist quantity index is a weighted average of growth rates. If a component is zero in one period, but positive in the other period, then its growth rate cannot be calculated (since one cannot take the log of zero) and the index does not exist.

This possibility shows that the Fisher index (or its constituents, the Laspeyres and the Paasche) possesses a further advantage: it can be calculated even when some base period quantities are zero. As far as I know this advantage has not been pointed out before.

---

[7] Similar formulas can be developed for Laspeyres and Paasche *price* indices. Here the expenditure shares are the same as in the corresponding quantity index, but price relatives ($p_{it}/p_{i,t-1}$) replace quantity relatives: see Annex A.

There is an economic interpretation of the impossibility of calculating the Törnqvist index when some quantities are zero: the Törnqvist assumes that all inputs are "essential", ie in a production function output is zero if any single input is zero. This is realistic at a sufficiently high level of aggregation, eg producing output is impossible without labour. But if labour is disaggregated into detailed types, then clearly this assumption need not hold, eg some output can certainly be produced without accountants or coal miners. By contrast, the Fisher index corresponds to a production function where no input is essential.[8]

Index numbers when some quantities can take negative values

Some economic magnitudes can be positive in some periods and negative in others. Inventory investment (the change in the stock of inventories) is one example. In this case even a Fisher index does not make sense, since we cannot calculate the growth rate of a variable which is sometimes positive, sometimes negative. The solution adopted in the U.S. National Income and Product Accounts is to calculate a Fisher index of the real *stock* of inventories, which is necessarily non-negative. Then the change in inventories is calculated as the first difference of the stock. Actually, the problem goes potentially wider than inventories. We tend to think of gross investment as a positive magnitude, but it is actually estimated as acquisitions less disposals. Conceptually therefore it can be negative and in practice this does occur at the industry level. So calculating an industry level index of investment may be problematic.

4. Basic properties of fixed weight indices

In order to derive important results about the behaviour of Laspeyres and Paasche indices, and also of chain indices, we need to derive first a basic result about any fixed base index. The question we want to answer initially is: what is the effect on a fixed base quantity index of a rise in the weight on one of its constituent goods? Let us express a fixed base quantity index in a general form:

$$Q_t = \frac{\sum_i p_i q_{it}}{\sum_i p_i q_{i,t-1}}$$

---

[8] Consider a translog function in unlogged form. If we take the limit as one of the constituent quantities goes to zero, then output (or utility) goes to zero too. The production function corresponding to a Fisher quantity index is $\left[\sum_i \sum_j a_{ij} q_i q_j\right]^{1/2}$, $a_{ij} = a_{ji}$, where the q_i are the input quantities (Diewert, 1976), from which it can be seen that output is non-zero as long as at least one of the quantities is non-zero.

where we don't specify the period to which the price weights belong. Now consider the effect on the size of Q if the weight on the ith good, p_i, were higher. If all other prices are held constant, this means we are asking: what is the effect on the index of an increase in the *relative* price of good i? By differentiating (see Result 3, Annex B), we derive the following

BASIC PROPERTY OF FIXED BASE INDICES Suppose the relative price of good i is raised. Then

(a) the fixed base quantity index Q_t *rises* if the quantity of good i is rising *faster* than average

($q_{it} / q_{i,t-1} > Q_t$)

(b) the quantity index Q_t *falls* if the quantity of good i is rising *slower* than average

($q_{it} / q_{i,t-1} > Q_t$).

(c) if the quantity of good i is growing at the same rate as the average, there is no effect on the fixed base quantity index.

From this basic property, some important propositions now follow immediately:

PROPOSITION 1 A Laspeyres quantity index will grow faster than a Paasche if the relative price of goods growing faster than average is falling. The reason is that the Paasche applies lower weights to goods growing more rapidly than average. Since the Fisher index is a geometric mean of the Laspeyres and Paasche, the Laspeyres will therefore also grow faster than the Fisher index.

It is quite likely that the relative price of fast growing goods will be falling, and this certainly applies in the case of computers. But there is no reason in theory why this should always be the case. For a counter-example, consider the GDP of an economy dominated by international trade, where domestic production is mainly for export and imports are of consumer goods. Here it is quite possible for quantity changes to be dominated by supply considerations, so that relative quantities are *positively* correlated with relative prices.[9]

[9] *National Accounts: Concepts, Sources and Methods* (ONS 1998), paragraph 2.74, suggests that oil price shocks provide a counter-example for the U.K.. In a recent article in the *Times*, Anatol Kaletsky claimed that chain-linking would increase the U.K. growth rate post 1995 (*The Times*, May 28, 2002, available at www.timesonline.co.uk). His argument was that service industries, whose prices have been rising, have been growing more rapidly than manufacturing. For an official assessment of the likely effect on GDP growth (a small cut), see Tuke and Reed (2001).

PROPOSITION 2 Suppose that the Laspeyres quantity index of (some component of) GDP is rising faster than the Paasche quantity index (say for the reason given in Proposition 1). Suppose that the national accounts use Laspeyres quantity indices. Then the implicit deflator will rise less rapidly than a Laspeyres price index.

Proof: Expenditure (V) is the product of the Laspeyres quantity and the Paasche price index, or of the Paasche quantity and the Laspeyres price index:

$$V_t = Q_t^L \cdot P_t^P = Q_t^P \cdot P_t^L$$

(See Annex A). So if $Q_t^L > Q_t^P$, then $P_t^L > P_t^P$, ie the Laspeyres rises more rapidly than the Paasche price index. But when the national accounts use Laspeyres quantity indices the implicit deflator is defined as V_t / Q_t^L, so it is a Paasche price index. Therefore the Laspeyres price index rises faster than the implicit deflator. QED.

PROPOSITION 3 Suppose that the national accounts are rebased throughout from an earlier to a later year. Then if goods with faster than average growth have falling relative prices, rebasing will lower the growth rate of GDP throughout.

The situation envisaged in Proposition 3 used to apply in the United States prior to the adoption by the Bureau of Economic Analysis (BEA) of annual chain-linking in 1996. For example, when the national accounts were rebased from 1982 to 1987, growth rates in all years (from 1929 to the most recent) were affected. In fact this was one of the reasons why annual chain-linking was adopted soon after the new, rapidly falling computer price index was introduced into the U.S. National Income and Product Accounts (NIPA). Without chain-linking, the computer price would have caused U.S. growth to be revised downwards every time rebasing occurred (Landefeld and Parker, 1997).[10]

Note that Proposition 3 does *not* apply generally to the U.K. (or to most other European countries), where periodic updating of the weights occurred even before the adoption of annual chainlinking. For example, in the U.K. when the 1995 base was adopted instead of the 1990 one, growth rates for years before 1994 were unaffected, since the weights for those years were the

---

10 This is an interesting example of endogenous methodological change. If the new computer price index had not been adopted, quite possibly the U.S. would still be using a fixed base index.

same as before. For these years, only the reference year was changed (1995=100 instead of 1990=100). The constant price levels were now in 1995 instead of 1990 prices, but growth rates both of totals and components were unaffected by rebasing.[11]

5. Lower level indices

To calculate any of the indices described above, knowledge of expenditure shares is required. But at the lowest level such knowledge does not exist. For example, in calculating the U.K. RPI the Office for National Statistics (ONS) knows from the National Food Survey what proportion of household budgets goes to apples. But the ONS does not collect the price of "apples" but rather the prices of different varieties of apple (Cox, Worcester, Golden Delicious, etc) in different shops in different regions, month by month. A similar situation applies to producer prices where the ONS has no information about the sales of the individual products whose prices it is collecting. There is no alternative here but to take an unweighted average of the individual price quotes. But there is more than one way to take even an unweighted average.

Three methods are in common use:

1. Ratio of arithmetic means (RA): $\dfrac{(1/N)\sum_i p_{it}}{(1/N)\sum_i p_{i0}}$

2. Arithmetic mean of price relatives (AR): $(1/N)\sum_i (p_{it}/p_{i0})$

3. Geometric mean of price relatives (GM): $\left[\prod_i (p_{it}/p_{i0})\right]^{1/N}$

Here p_{it} is now the price of (say) Golden Delicious in shop i in month t and there are N price quotes to average. The ONS uses the first of these formulas (RA) in the RPI and also the second (AR); it does not use the third (GM) in the RPI. But it does use GM in the Harmonised Index of Consumer Prices (HICP) since this is mandated by Eurostat. The U.S. Bureau of Labor Statistics (BLS) used to use AR prior to the Boskin Report. But the latter strongly criticised AR and recommended GM. As a result the BLS has largely gone over to GM.

---

[11] Growth in 1994 was affected since this was rebased to 1995 weights where previously 1990 weights were used.

The AR formula tends to produce higher results than RA. The reason is that it implicitly gives a relatively high weight to low prices and low prices tend to grow more rapidly than high ones: in other words, prices for a given item tend to regress towards some mean level (Carruthers *et al*. 1980). The third formula (GM) necessarily produces lower results than the second (AR) since the geometric mean is always lower than the arithmetic mean (unless all price relatives are equal).

The first formula (RA) corresponds to a "fixed proportions" type of preferences: the elasticity of substitution between apples in different shops is zero and the relative quantities of apples purchased in different shops is totally unaffected by relative prices. The geometric mean corresponds to Cobb-Douglas preferences where the elasticity of substitution is –1: the proportion of expenditure on apples in each shop is constant. In the case of apples, an elasticity of substitution of –1 seems much more realistic than one of zero. In fact, it would not be surprising if the elasticity were even higher. However, there might be cases where an elasticity nearer to zero was more realistic, for instance if the items being priced were rather disparate. So it is an empirical matter as to which is the most appropriate assumption in individual cases. But Boskin was in no doubt that the geometric assumption is more realistic in general.

A further disadvantage of RA is that it only makes sense when the price quotes are for homogeneous items. But often this is not the case. For example though an RPI price collector is pricing the same model of washing machine month by month in a given shop, another collector may be pricing a different model in another shop. Since the two models may well differ in quality and hence sell at widely different prices, it makes no sense to take an arithmetic average of the quotes as in RA. It is for this reason that the ONS sometimes uses the AR method. But because of the bias identified by Carruthers *et al*. (1980), the GM method (which also makes sense for non-homogeneous items) seems preferable.

Another reason for rejecting AR is that it fails the time reversal test while RA and GM pass it (Diewert 1995).

6. Chain indices: comparisons over more than two periods

The Divisia approach

The intuition behind a chain index is that we want the weights to be as up to date as possible. As we have seen in the two period case, it is better to use the weights of both periods, rather than just one period. In a chain index, we change the weights in every period. When we use a fixed base index over a long period of time, for most of the time we are calculating growth rates using weights which apply to none of the years being compared.

An attractive approach to chain index numbers derives from the work of François Divisia, as filtered through Hulten (1973). Let the total value of output of (or expenditure on) some set of products at time *t* be given by

$$V(t) = \sum_i p_i(t) q_i(t) \tag{6}$$

where the p_i are the prices and q_i the quantities. We want to be able to write this in the form

$$V(t) = P(t) Q(t)$$

where *P* is some sort of price index and *Q* a quantity index, just like one can for a single product. This equation can be rewritten in continuous growth rate terms as

$$\hat{V}(t) = \hat{P}(t) + \hat{Q}(t) \tag{7}$$

where a hat (^) denotes a continuous time growth rate. Divisia noted that if we totally differentiate equation (6) with respect to time, we obtain

$$\hat{V}(t) = \sum_i w_i(t) \hat{p}_i(t) + \sum_i w_i(t) \hat{q}_i(t) \tag{8}$$

where $w_i(t)$ is the value share: $w_i(t) = p_i(t) q_i(t) / V(t)$. Divisia suggested that we identify the first summation on the right hand side with the price index and the second with the quantity index:

$$\begin{aligned} \hat{P} &= \sum_i w_i(t) \hat{p}_i(t) \\ \hat{Q} &= \sum_i w_i(t) \hat{q}_i(t) \end{aligned} \tag{9}$$

Note that the weights are the value shares *at a point in time*, and so change continuously.

If we want the levels of the indices, we can set the price index equal to 1 in some reference year, say 1995, and the quantity index equal to nominal expenditure in the reference year. Then $P(1995) = 1,\ Q(1995) = V(1995)$ and we can generate the levels in any other period by applying the growth rates. The quantity index will now measure output in "chained 1995 euros". Alternatively, we can set both the price index and the quantity index equal to 1 in the base year. Then the product $P(t)Q(t)$ will give us an index of expenditure: the ratio of expenditure in one year to expenditure in the reference year. Clearly, growth rates are independent of the reference year, which affects only the levels of the indices. The reference year can be chosen purely as a matter of convenience or convention: but see the discussion below on non-additivity.

Another attractive feature of Divisia indices is that they display the property known as "consistency in aggregation". Suppose that we have a Divisia index of an aggregate like investment and that we also have Divisia indices of the sub-aggregates which make up the overall aggregate, eg investment in structures and investment in machinery. Then the aggregate Divisia index, directly calculated from the individual components (machine tools, computers, offices, shopping centres, etc), equals the indirect Divisia index, calculated in two stages as the Divisia index of the Divisia indices of the two sub-aggregates. This is obvious from equation (9).

Chain indices in practice

Though Divisia indices are a very useful theoretical tool, they cannot actually be calculated since they are defined in continuous time and real data comes in discrete time. Hence the issue becomes one of finding a good, discrete approximation to (9). From a theoretical point of view, the best choice would be, from the four options above, either the Fisher or the Törnqvist. The U.S. has chosen the Fisher for its national accounts, which has the great advantage that (as stressed above) the product of the price and quantity indices gives nominal expenditure. Eurostat has unfortunately chosen Laspeyres for the chained quantity index of GDP. This means that since annual chain-linking was introduced in 2003, the GDP deflator has been a chained Paasche index. This is better than the previous position, but still not ideal.

Neither the Fisher nor the Törnqvist inherit the property of exact consistency in aggregation from the Divisia indices that they approximate. However, Diewert (1978) has shown that they possess this property to a high degree of approximation.

It is well known that chain indices are non-additive: when expressed in the prices of some reference year, the components don't in general sum exactly to the total. The exceptions are that the components do sum to the total (a) in the reference year itself and (b) in the case of a chain-weighted Laspeyres quantity index, in the following year too. In contrast, in a fixed base index the components in constant prices sum exactly to the total in every year. When in 1996 the U.S. shifted the NIPA to an annual chain-linked index from a fixed base one, the loss of additivity caused a great fuss. But in Europe (including the U.K.) it has been customary to update the weights every five years or so. So except for the period where the most recent set of weights apply, the national accounts are non-additive. In other words, Europe has always used a form of chain-linking. The change, which in the U.K.'s case took place in 2003, is from what might be called quinquennial chain-linking to annual chain-linking.

Properties of chain indices

A chain index of the level of some quantity (eg GDP, investment, or consumption) at period t relative to period 1, Q_t^{Ch}, may be defined as follows:

$$Q_{1t}^{Ch} = Q_{12} \cdot Q_{23} \cdot \ldots \cdot Q_{t-2,t-1} \cdot Q_{t-1,t}$$

where $Q_{s-1,s}$ is the two period quantity index which (as before) shows the level of the aggregate in period s relative to its level in period s-1. The $Q_{s-1,s}$ are the individual "links" in the chain and in principle could be of any of the types considered earlier. Thus a chained Laspeyres index is:

$$Q_{1t}^{ChL} = Q_{12}^{L} \cdot Q_{23}^{L} \cdot \ldots \cdot Q_{t-2,t-1}^{L} \cdot Q_{t-1,t}^{L}$$

and a chained Fisher is defined analogously. We can pick a particular year as the reference year, say 1995. Then we can either set the index equal to 100 in that year or we can multiply the index in any year by the nominal value of the aggregate in 1995. If we choose the latter, then we are expressing the index in "chained 1995 euros".

PROPOSITION 4 Suppose that fast growing goods have falling relative prices. Then the replacement of a fixed base index by a chain-weighted one will *reduce* growth in the periods after the original base period, but *increase* it in the periods before the original base.

Proof This follows immediately from the Basic Property.

PROPOSITION 5 Non-additivity: the components of a chain quantity index, when expressed in the prices of a particular year, do not sum to the total except in years for which these base year prices are employed (unless the relative prices of the components are constant over time).

Proof In a fixed base index, the components in constant prices always sum to the total in constant prices. In a chain index, the base changes periodically. It is not possible for the components to sum to the total in two different sets of relative prices. Suppose we choose 2000 as our base year. Then in 2000 the chain index and the fixed base index of (say) GDP agree and both equal nominal GDP. Both before and after 2000 the chain index will in general grow at a different rate from the fixed base index. Therefore the level of the chain index in "chained 2000 euros" will diverge from the level of the fixed base index (measured in "2000 constant prices") before and after 2000. Hence the components cannot sum to the level of the chain index.

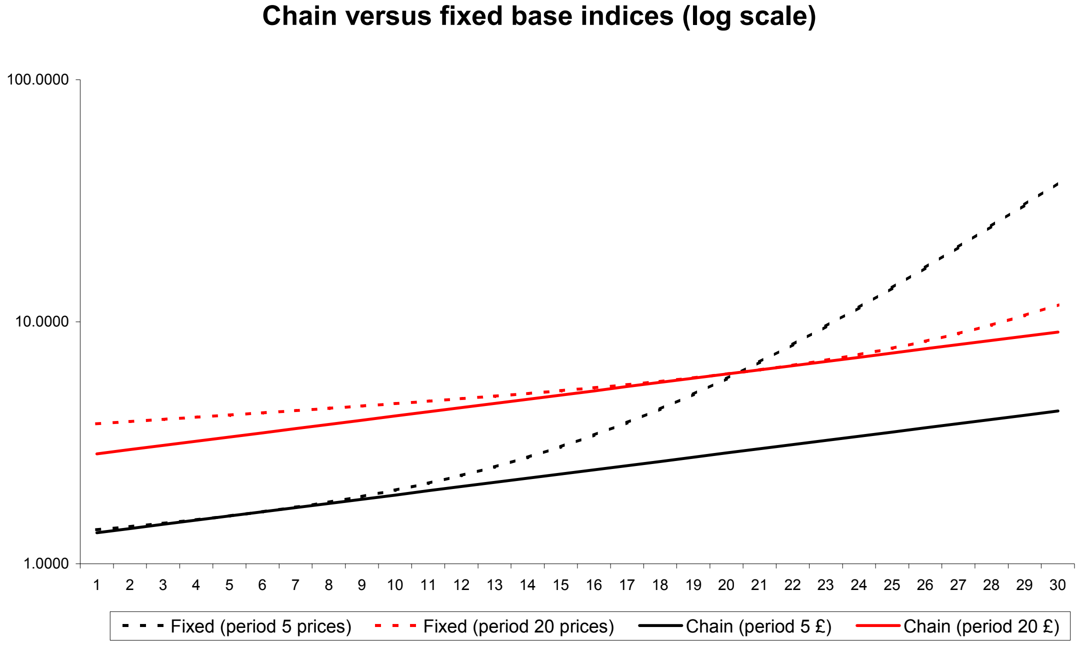
Suppose that fast growing goods have falling relative prices. Then by Proposition 4 the chain index will grow more rapidly than the fixed base index before the base year and less rapidly after it.[12] So its level will be below that of the fixed base index both before and after the base year. Hence apart from the base year itself (and for a chained Laspeyres, the subsequent period too), the components will sum to *more than* the chain index.

This point is illustrated in Chart 1, which compares a chained Törnqvist with a fixed base index on a log scale. There are assumed to be two components to the index; one component grows more rapidly than the other and its relative price is falling. For convenience a steady state is assumed in which expenditure shares are constant. Hence on a log scale the chain index appears as a straight line. Two separate reference years are illustrated, here period 5 (black lines) and 20 (red lines). Since the growth rate of the chain index is independent of the reference year, the two chain indices appear as parallel straight lines. For the two fixed base indices, the base year changes as well as the reference year. So their growth rates are affected by choice of base year. Each fixed base index (dashed line) is tangent to the chain index with the same reference year in that year. Prior to the reference year the fixed base index grows more slowly, afterwards more rapidly.

---

12 This suggests that the U.K.'s shift from updating the weights every five years to annual updating did not change long run average growth rates (say over 25 years) very much, since the errors tend to cancel out. Unfortunately,

Chart 1



Note It is assumed that the quantities of the goods are growing at 2% and 22% per period respectively and their prices at 7% and -13% respectively. The share of the fast growing good is 9.1%.

Charts 2 and 3 shows the ratio of the sum of the two components to the quantity index in "chained euros" for two reference or base periods, period 5 and period 20. The ratio is equal to 1 in the reference period but exceeds 1 in all other periods. With period 5 as the base, the ratio is nearly 9 times the chain index 25 periods later. Even with period 20 as the base the ratio exceeds the chain index by about 30% in both the first and the 30th periods. Obviously the results depend on the assumptions but these are not outlandish by reference to the case of computers.

Chart 2

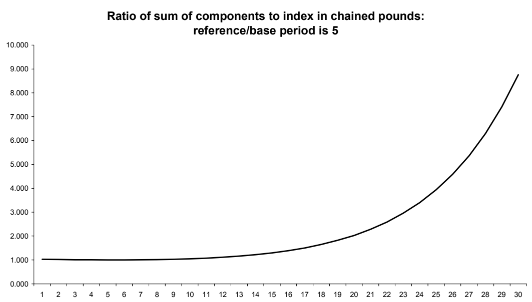
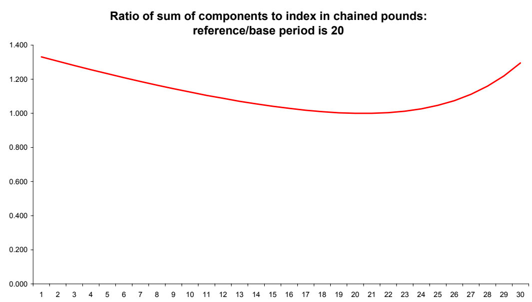


Chart 3



---

annual chainlinking was introduced alongside other changes including data revisions, so it is not possible to be sure about this.

PROPOSITION 6 In a chained quantity index, a part can be greater than the whole, when expressed in constant prices ("chained euros").

This is just a consequence of non-additivity. In a chained quantity index, a component can grow at a permanently faster rate than the aggregate. So the ratio of a part to the whole in constant prices (or, better, in "chained 1995 euros"), for example the investment-GDP ratio, can rise indefinitely. Eventually, it will exceed 1. See Annex B for proof.

PROPOSITION 7 Chain indices are not path-independent: they may fail the circular test.

This is discussed in and proved (for Törnqvist indices) as Result 5 of Annex B. This is a potentially serious defect of chain indices. But as discussed in Annex B, it is mitigated by two factors. First, if constant returns to scale prevail, then output price and quantity indices do satisfy circularity; the counterpart condition for consumer price indices is homotheticity of the utility function. Second, departures from circularity seem small in practice.

Chain versus fixed base indices when expenditure shares are constant

Consider a situation where the expenditure shares are constant over time. It is then tempting to conclude that a fixed base and a chain index will be numerically identical. Tempting, but wrong. The reason is that the Basic Property still applies. It is quite possible for prices and quantities to be negatively correlated while shares remain constant. So in this case a chain index will grow more slowly than a fixed base one. If prices and quantity are positively correlated then the opposite will be true. The point is that constancy of the shares tells us nothing by itself about the relationship between chained and fixed base indices.

To see why constancy of the shares does not make chain and fixed base indices identical, consider the example of a chained Laspeyres over three periods:

$$Q_{13}^{ChL} = Q_{23}^{L} \cdot Q_{12}^{L} = \sum_i w_i \left(\frac{q_{i3}}{q_{i2}} \right) \cdot \sum_i w_i \left(\frac{q_{i2}}{q_{i1}} \right)$$

(Here we write the shares without a time subscript since they are assumed constant over time). The fixed base Laspeyres of period 3 relative to period 1 is

$$\sum_i w_i \left(\frac{q_{i3}}{q_{i1}} \right)$$

so clearly the fixed base and the chained Laspeyres are *not* the same. The only time when they are the same is if all quantities are growing at the same constant rate over time. (Since shares are constant, this means that all prices are growing at the same rate too). If this common (discrete) growth rate is *g*, then both the chained and the fixed base index will yield the value $(1+g)^2$ in this case.

Chained versus fixed base indices: the general relationship

Chained and fixed base quantity (price) indices will yield identical results in the trivial case where all quantities (prices) are growing at the same rate. Apart from this case, the general proposition for quantity indices is:

PROPOSITION 8 If relative prices are constant over the relevant period, then chained quantity indices yield the same result as fixed base indices of the same type.

Proof Annex B shows that this proposition holds for Laspeyres and Paasche indices. Since a Fisher index is the square root of the product of Paasche and Laspeyres indices, it follows that chained and fixed base Fisher indices are equal when relative prices are constant.

Clearly an analogous proposition holds for price indices, where the condition is constancy of relative *quantities.*

7. Macro modelling implications of chain indices

Macro models inevitable simplify, which means that they must aggregate. Aggregation implies the use of index numbers. A common choice is to assume that the economy produces only one good. For purely theoretical purposes one could assume that this is literally true. Or, better, one could assume as many goods and services as one likes but that relative prices are constant. Then by Hick's aggregation theorem the economy behaves as if it produces only one good. Unfortunately, there may be important changes in relative prices that are significant at the macro level. A current example is the tendency for the prices of capital goods to fall relative to those of consumer goods and, within capital goods, for the relative price of computers to fall still more rapidly.

A macro model typically has a long run steady state whose properties are suggested by theory. For example, in a Solow, one-good growth model, output, consumption, investment and the capital stock all grow at the same rate in steady state. Now we do not believe that Hick's aggregation theorem applies literally, but we still may want to use a one-good model (as many central banks do). The question then is this: is it sensible to impose the long run properties of the one-good model on real life aggregates like GDP, investment or the capital stock?

The answer in general is no. This can be demonstrated by means of a simple model. Suppose there are two goods and the economy is closed. The first good can be used for either consumption or investment , the second for investment only. (One can think of the first good as a composite to which Hick's theorem applies). Following Bakhshi and Larsen (2001), let us label the two sorts of capital good "dull" and "exciting"; "dull" capital can also be consumed. The relative price of exciting capital is falling at a constant rate, because technical progress is faster in this sector. If such a model is to have a steady state, we must assume Cobb-Douglas technology, in which the current price shares of investment in output, and of each of the two assets in the aggregate capital stock, are constant, at least asymptotically.[13] So to keep the shares of exciting capital constant, investment in exciting capital must be growing faster than investment in dull capital. Denote the growth rates of the two kinds of investment by g_d, g_e, so $g_e > g_d$. These are also the steady state growth rates of the corresponding stocks. We also assume that the exciting capital depreciates faster: $\delta_e > \delta_d$. To complete the model we assume for simplicity that labour input is constant. Then we have

PROPOSITION 9 Let all aggregates be measured by Divisia indices. Then, in this model in steady state, aggregate investment (*I*) grows faster than the aggregate capital stock (*A*) and faster too than GDP (*Y*). The growth rate of capital *services* (*K*) will exceed that of the capital *stock*. In symbols,

$$\hat{I} > \hat{A};\ \hat{K} > \hat{A}; \hat{I} > \hat{Y} > \hat{C}$$

[13] Models of this type have been studied by Bakhshi and Larsen (2001). See also Whelan (2001).

If we assume an infinitely lived representative household maximising the present value of utility from consumption, where the instantaneous utility function is logarithmic, then a more complete ranking of growth rates can be shown to hold:

$$\hat{I} > \hat{K} > \hat{A}$$

Proof. See Annex B.

Proposition 9 is for Divisia indices. But examination of the proof in the Annex shows that it holds also for discrete growth rates. With this interpretation, Proposition 9 holds for Törnqvist indices too. Since Fisher indices approximate Divisia indices, and empirically yield very similar values to Törnqvist indices, one can expect the Proposition to hold for them too. Chained Laspeyres indices also approximate chained Törnqvist indices when shares are constant as here.

The fact that aggregate investment is growing faster than the capital stock makes the aggregate capital accumulation identity

$$A(t) = I(t) + (1 - \delta(t))A(t-1)$$

potentially misleading. Here the aggregate depreciation rate has been written as a function of time since it will not be constant, despite the fact that the individual depreciation rates are assumed constant along with the share of each asset in the value of the capital stock. Solving for δ, we find

$$\delta(t) = \frac{I(t)}{A(t-1)} - \frac{A(t) - A(t-1)}{A(t-1)}$$

Now I grows more quickly than A, so the first term on the right hand side rises without limit. So in this model in steady state, the aggregate depreciation rate rises without limit. Not only will it eventually exceed the highest individual rate, but it will also eventually exceed 1![14]

[14] See Whelan (2000) and Oulton and Srinivasan (2003).

8. *Further reading*

Theory

Fisher (1927) is mainly of historic interest. The basic reference is still Diewert (1987) who also considers a number of topics (such as cross-country comparisons) not discussed here. This paper gives numerous references to the theoretical literature. For the theory of Divisia index numbers see Hulten (1973). Diewert (1976) on exact and superlative index numbers is a basic paper. For the theory of the cost of living, see Deaton and Muellbauer (1980) and for the corresponding (though less well known) theory of output price and quantity indices see Fisher and Shell (1998). For hedonic index numbers, see Triplett (1987) and (1990). Many index number properties are listed in the SNA93 manual (Commission of the European Communities – Eurostat *et al.* (1993), chapter XVI), though without proof. For some pitfalls in interpreting chain indices, see Whelan (2000).

Practice

The construction of the U.K.'s RPI is described in ONS (1998); a general reference on consumer price indices is the ILO manual (Turvey *et al.* (1989)). The critique of the U.S. CPI by the Boskin Commission (Advisory Commission to Study the Consumer Price Index (1996)) has been very influential. For the relationship between deflators and price indices, see Triplett (1981). The U.S. approach to chain-linking is discussed in Landefeld and Parker (1997), the U.K. approach in Lynch (1996) and Tuke and Reed (2001). The OECD is producing a manual on producer price indices, including the use of hedonic indices. When this is available it is likely to be an important reference. The special problems of quarterly chain linking, and how to make quarterly and annually chain-liked data mutually consistent, are considered in the IMF's *Quarterly National Accounts Manual* (IMF 2001, chapter IX).

References

Advisory Commission to Study the Consumer Price Index [The Boskin Commission] (1996). *Towards a More Accurate Measure of the Cost of Living*. Final Report to the Committee on Finance of the U.S. Senate. Washington, DC.

Bakhshi, H. and Larsen, J. (2001). "Investment-specific technological progress in the United Kingdom". Bank of England Working Paper no. 129.

Balk, B.M. (1995). "Axiomatic price index theory: a survey". *International Statistical Review*, 63, 1, 69-93.

Carruthers, A.G., Sellwood,D.J. and Ward, P.W. (1980). "Recent developments in the Retail Prices Index". *The Statistician*, 29, 1-32.

Caves, D.W., Christensen, L.R. and Diewert, D.W. (1982). "Multilateral comparisons of output, input, and productivity using superlative index numbers". *Economic Journal*, 92 (March), 73-86.

Commission of the European Communities – Eurostat *et al.* (1993). *System of National Accounts 1993*. Brussels/Luxembourg, etc.

Deaton, A. and Muellbauer, J. (1980). *Economics and Consumer Behaviour.* Cambridge: Cambridge University Press.

Diewert, W.E. (1976). "Exact and superlative index numbers". *Journal of Econometrics*, 4, 115-46.

Diewert, W.E. (1978). "Superlative index numbers and consistency in aggregation". *Econometrica*, 46, 883-900.

Diewert, W.E. (1987). "Index numbers". In *The New Palgrave: A Dictionary of Economics,* edited by J. Eatwell, M. Milgate and P. Newman. London and Basingstoke: Macmillan Press.

Diewert, W.E. (1995). "Axiomatic and economic approaches to elementary price indexes". NBER Working Paper no. 5104. Cambridge, MA.

Fisher, F.M. and Shell, K. (1998). *Economic Analysis of Production Price Indexes*. Cambridge: Cambridge University Press.

Fisher, I. (1927). *The Making of Index Numbers: A Study of Their Varieties, Tests, and Reliability* (3rd edition). Reprinted in 1967 by Augustus M. Kelley, New York.

Fox, Kevin J. (2003). "An economic justification for the EKS multilateral index". *Review of Income and Wealth*, series 49, number 3 (September), pages 407-413.

Hulten, C.R. (1973). "Divisia index numbers". *Econometrica*, 41, 1017-1025.

International Monetary Fund (2001). *Quarterly National Accounts Manual — Concepts, Data Sources, and Compilation.* Available at www.imf.org/external/pubs/ft/qna/2000/textbook/index.htm.

Landefeld, J.S. and Parker, R.P. (1997). "BEA's chain indexes, time series, and measures of long-term economic growth". *Survey of Current Business*, May, 58-68.

Lynch, R. (1996). "Measuring real growth – index numbers and chain-linking", Economic Trends, no. 512 (June), 31-37.

Office for National Statistics (1998). *The Retail Prices Index: Technical Manual : 1998 Edition.* London: The Stationery Office. Downloadable from www.statistics.gov.uk.

Oulton, N. (1995). "The implications of the Boskin Report". *National Institute Economic Review*, no. 165 (July), 89-98.

Oulton, N. and Srinivasan, S. (2003). "Capital stocks, capital services and depreciation: an integrated framework". *Bank of England Working Paper no. 192.* Available from www.bankofengland.co.uk.

Samuelson, P.A. and Swamy, S. (1974). "Invariant economic index numbers and canonical duality: survey and synthesis". *American Economic Review*, vol. 64 (4, September), 566-593.

Selvanathan, E.A. and Prasada Rao, D.S. (1994). *Index Numbers: A Stochastic Approach.* Basingstoke: Macmillan Press.

Statistics Commission (2004). *Review of Revisions to Economic Statistics*. Report No. 17, volume 3, Annex 3: Submission by Office for National Statistics.

Triplett, J.E. (1981). *Monthly Labour Review*, September, pp 3-15.

Triplett, J.E. (1987). "Hedonic functions and hedonic indexes". In *The New Palgrave: A Dictionary of Economics*, edited by J. Eatwell, M. Millgate and P. Newman. London: Macmillan.

Triplett, J.E. (1990). "Hedonic methods in statistical agency environments: an intellectual biopsy". In *Fifty Years of Economic Measurement*, edited by E.R. Berndt and J.E. Triplett. Chicago and London: University of Chicago Press.

Tuke, A. and Reed, G. (2001). "The effects of annual chain-linking on the output measure of GDP". *Economic Trends*, no. 575 (October), 37-53.

Szulc, B.H. (1989). "Price indices below the basic aggregation level". Appendix 7 of Turvey *et al.* (1989).

Turvey, R. *et al.* (1989). *Consumer Price Indices: An ILO Manual.* Geneva: International Labour Office.

Whelan, K. (2000). "A guide to the use of chain aggregated NIPA data". Washington, D.C.: Board of Governors of the Federal Reserve.

Whelan, K. (2001). "A two-sector approach to modeling U.S. NIPA data". Washington, D.C.: Federal Reserve Board. Mimeo.

ANNEX A
INDEX NUMBER FORMULAS

The table below shows the index numbers in common use, together with their formulas. The two time periods are labelled t-1 and t, but the gap between them could be more than one period.

Table A.1
Common two period index numbers

	Price index		***Quantity index***	
Name	*Symbol*	*Formula*	*Symbol*	*Formula*
Laspeyres	P_t^L	$\sum_i w_{i,t-1}\left(\frac{p_{it}}{p_{i,t-1}}\right)$	Q_t^L	$\sum_i w_{i,t-1}\left(\frac{q_{it}}{q_{i,t-1}}\right),\ q_{i,t-1}>0$
Paasche	P_t^P	$\left[\sum_i w_{it}\left(\frac{p_{i,t-1}}{p_{it}}\right)\right]^{-1}$	Q_t^P	$\left[\sum_i w_{it}\left(\frac{q_{i,t-1}}{q_{it}}\right)\right]^{-1},\ q_{it}>0$
Fisher	P_t^F	$\left[P_t^L \cdot P_t^P\right]^{1/2}$	Q_t^F	$\left[Q_t^L \cdot Q_t^P\right]^{1/2}$
Törnqvist[a]	P_t^T	$\sum_i\left[\frac{w_{i,t-1}+w_{it}}{2}\right]\ln\left(\frac{p_{it}}{p_{i,t-1}}\right)$	Q_t^T	$\sum_i\left[\frac{w_{i,t-1}+w_{it}}{2}\right]\ln\left(\frac{q_{it}}{q_{i,t-1}}\right),$ $q_{i,t-1}>0$

a. The formula for the Törnqvist index defines its exponential growth rate (ie the log difference), not its level; equivalently, the formula is for the log of the index, since the value in the reference period t-1 is 1 (whose log is zero). To get the level in period t, raise e to the power of the expression in the table.

Note w_{it}: share of good i in the value of total expenditure at time t (budget shares in the case of a consumer price index, shares in GDP in the case of a GDP index).

Points to note:

1. The expressions $(p_{it}/p_{i,t-1})$, $(q_{it}/q_{i,t-1})$ are known as *price relatives* and *quantity relatives*, respectively.
2. The formulas for the Laspeyres and Paasche indices are not the ones usually given in textbooks. The usual formulas for these indices are:

$$Q_t^L = \frac{\sum_i p_{i,t-1} q_{it}}{\sum_i p_{i,t-1} q_{i,t-1}}$$
$$P_t^L = \frac{\sum_i p_{it} q_{i,t-1}}{\sum_i p_{i,t-1} q_{i,t-1}}$$
$$Q_t^P = \frac{\sum_i p_{it} q_{it}}{\sum_i p_{it} q_{i,t-1}}$$
$$P_t^P = \frac{\sum_i p_{it} q_{it}}{\sum_i p_{i,t-1} q_{it}}$$

The formulas in Table A.1 are algebraically equivalent to these (if quantities are non-zero), but correspond better to how such indices are actually calculated. In practice, statistical agencies start with price (or quantity) relatives and weight these together using budget shares (for the RPI) or output weights (eg base year shares in GDP for constant price GDP). The formulas in Table A.1 also make clearer the connection between Laspeyres and Paasche on the one hand and the Törnqvist index on the other. However, the text book formulas have the advantage of showing that Laspeyres, Paasche and Fisher quantity indices can still be calculated when some of the quantities are zero.

3. Note that in the Paasche formulas in the table, the price and quantity relatives are the other way up to the way they are in the Laspeyres formula. Hence we take the inverse “to get them the right way up”.
4. By definition, the Fisher quantity (price) index is the geometric mean of the Laspeyres and Paasche quantity (price) indices.
5. The product of the Fisher quantity index and the Fisher price index is the ratio of the values of expenditure in the two periods, ie it is the expenditure index:

$$V_t = P_t^F \cdot Q_t^F .$$

   This is *not* true of any other pair of price and quantity indices of the same type in Table A.1. However:

6. The product of the *Laspeyres* quantity index and the *Paasche* price index is the expenditure index. And the product of the *Paasche* quantity index and the *Laspeyres* price index is also the expenditure index. Points 5 and 6 are proved in Annex B (Result 1).
7. Using the result that $\ln(1+x) \approx x$ for small x, we obtain an approximation for the Törnqvist:

$$Q_t^T \approx \sum_i \left[\frac{w_{it} + w_{i,t-1}}{2} \right] \cdot \left(\frac{q_{it}}{q_{i,t-1}} \right)$$

which may be compared with the formula in the table for the Laspeyres. This shows that, approximately, the Törnqvist is just like a Laspeyres except that it uses an arithmetic average of the shares in the two periods, instead of just the base period shares.

ANNEX B
PROOFS OF PROPOSITIONS IN THE TEXT

Result 1. The product of the Laspeyres price (quantity) index and the Paasche quantity (price) index is the expenditure index.

Proof

$$P_t^L \cdot Q_t^P = \left[\frac{\sum_i p_{it} q_{i,t-1}}{\sum_i p_{i,t-1} q_{i,t-1}}\right] \cdot \left[\frac{\sum_i p_{it} q_{it}}{\sum_i p_{it} q_{i,t-1}}\right] = \frac{\sum_i p_{it} q_{it}}{\sum_i p_{i,t-1} q_{i,t-1}} = V_t$$

$$P_t^P \cdot Q_t^L = \left[\frac{\sum_i p_{it} q_{it}}{\sum_i p_{i,t-1} q_{it}}\right] \cdot \left[\frac{\sum_i p_{i,t-1} q_{it}}{\sum_i p_{i,t-1} q_{i,t-1}}\right] = \frac{\sum_i p_{it} q_{it}}{\sum_i p_{i,t-1} q_{i,t-1}} = V_t$$

Result 2. The Fisher price and quantity indices satisfy the time reversal and product tests.

Proof

(a) time reversal. Consider the Fisher *price* index. This is the square root of the product of the Laspeyres and Paasche price indices. With the original data:

$$P_t^F = \left[\frac{\sum_i p_{it} q_{i,t-1}}{\sum_i p_{i,t-1} q_{i,t-1}} \frac{\sum_i p_{it} q_{it}}{\sum_i p_{i,t-1} q_{it}}\right]^{1/2} = \left[\frac{A}{C}\frac{B}{D}\right]^{1/2}$$

in obvious notation. Suppose the price and quantity vectors are reversed, so that the price vector for time *t* becomes that for time *t*-1 and vice versa, and similarly for the quantity vectors. Then with the new data

$$P_t^F = \left[\frac{\sum_i p_{i,t-1} q_{it}}{\sum_i p_{it} q_{it}} \frac{\sum_i p_{i,t-1} q_{i,t-1}}{\sum_i p_{it} q_{i,t-1}}\right]^{1/2} = \left[\frac{D}{B}\frac{C}{A}\right]^{1/2}$$

(Here *D/B* is the Laspeyres price index with the new data and *C/A* is the new Paasche price index). Therefore the new index is the inverse of the old one. The same argument can obviously be used to prove that the Fisher *quantity* index satisfies the time reversal test.

(b) factor reversal. By definition,

$$P_t^F \cdot Q_t^F = \left[P_t^L \cdot P_t^P\right]^{1/2} \cdot \left[Q_t^L \cdot Q_t^P\right]^{1/2}$$

$$= \left[P_t^L \cdot Q_t^P\right]^{1/2} \cdot \left[P_t^P \cdot Q_t^L\right]^{1/2}$$

$$= V_t^{1/2} \cdot V_t^{1/2} = V_t$$

Here line 3 of this derivation uses Result 1.

Result 3. The basic property of fixed base indices

The fixed base quantity index is

$$Q_t = \frac{\sum_i p_i q_{it}}{\sum_i p_i q_{i,t-1}}$$

Obviously, doubling all prices leaves the index unchanged. The effect of a change in the *relative* price of good i can be computed by

$$\frac{\partial Q_t}{\partial p_i} = \frac{\left[\dfrac{q_{it}}{q_{i,t-1}} - Q_t\right] q_{i,t-1}}{\left[\sum_i p_i q_{i,t-1}\right]^3}$$

So $\partial Q_t / \partial p_i$ is positive, zero or negative according as $\left[(q_{it} / q_{i,t-1}) - Q_t\right]$ is positive, zero or negative.

There is an analogous result for price indices. The fixed base *price* index is

$$P_t = \frac{\sum_i q_i p_{it}}{\sum_i q_i p_{i,t-1}}$$

The effect of a change in the *relative quantity* of good i on the price index can be computed by

$$\frac{\partial P_t}{\partial q_i} = \frac{\left[\dfrac{p_{it}}{p_{i,t-1}} - P_t\right] p_{i,t-1}}{\left[\sum_i q_i p_{i,t-1}\right]^3}$$

So $\partial P_t / \partial q_i$ is positive, zero or negative according as $\left[(p_{it} / p_{i,t-1}) - P_t\right]$ is positive, zero or negative.

Consider a situation where the relative price of a good is falling and its relative price is rising. Suppose we compute price and quantity indices for aggregates including this good and then change the base to a later period when its relative price is lower. Then we have

$$\partial Q / \partial p_i > 0, \ \partial P / \partial q_i < 0$$

That is, under the circumstances envisaged (the relative price of good *i* is falling while its relative quantity is rising), the shift to the later base lowers both the price and the quantity index. So if the quantity aggregate is computed by a quantity index, then rebasing will lower the growth rate. But if it is computed by deflating the value by a price index, then rebasing the price index will *increase* the estimated growth rate.

Result 4. In a chained quantity index, a component can be larger than the total

This possibility is demonstrated by a simple example. Suppose that GDP is the sum of consumption and investment in current prices. Suppose too that the current price share of investment in GDP, w_I, is constant (as it would be in a steady state). Then a Törnqvist chain index of the growth of real GDP (*Y*) is

$$\Delta \ln Y_t = (1 - w_I) \Delta \ln C_t + w_I \Delta \ln I_t$$

Consider a situation where the growth rates of consumption and investment are constant, but investment is growing more rapidly than consumption. Then investment is also growing more rapidly than GDP, whose growth is also constant. Hence the volume ratio I_t / Y_t is rising without limit. It must therefore eventually exceed 1. Recall though that the current price ratio w_I is by assumption constant and less than one.

With fixed base indices, real GDP would be the sum of consumption and investment in constant prices:

$$Y_t = C_t + I_t$$

Then

$$\frac{Y_t}{Y_{t-1}} = \frac{C_t}{C_{t-1}} \frac{C_{t-1}}{Y_{t-1}} + \frac{I_t}{I_{t-1}} \frac{I_{t-1}}{Y_{t-1}}$$

Now the real shares C_{t-1} / Y_{t-1}, I_{t-1} / Y_{t-1} sum to 1. So if I_t is growing more rapidly than C_t, then the growth of Y_t rises asymptotically to equal that of investment. And the investment-GDP ratio in constant prices (I_t / Y_t) asymptotes to 1. So in contrast to the chained Törnqvist, with a fixed base we have

$$\lim_{t \to \infty} \Delta \ln Y_t = \Delta \ln I_t$$

That is, with a fixed base index the growth rate of GDP approaches the growth rate of the fastest growing component in GDP, here investment.

Result 5. Chain indices do not in general pass the circular test

The circular test requires that if prices and quantities change after the first period but later return to their period 1 values in the last period, then the index should be the same in the first and last periods. We can see that this cannot in general be true of chain indices by considering a Törnqvist index over four time periods:[15]

[15] Over just three periods a Törnqvist index always satisfies the circular test.

$$\ln Q_4^{ChT} = \ln Q_2^T + \ln Q_3^T + \ln Q_4^T$$

$$= \frac{1}{2}\left[\sum_i (w_{i2} + w_{i1}) \ln(q_{i2} / q_{i1}) + \sum_i (w_{i3} + w_{i2}) \ln(q_{i3} / q_{i2}) + \sum_i (w_{i4} + w_{i3}) \ln(q_{i4} / q_{i3})\right]$$

Expanding and cancelling terms,

$$\ln Q_4^{ChT} = \frac{1}{2}\left[\sum_i w_{i4} \ln q_{i4} - w_{i1} \ln q_{i1}\right] +$$

$$\frac{1}{2}\left[\sum_i w_{i1} \ln q_{i2} - w_{i2} \ln q_{i1}\right] + \frac{1}{2}\left[\sum_i w_{i2} \ln q_{i3} - w_{i3} \ln q_{i2}\right] + \frac{1}{2}\left[\sum_i w_{i3} \ln q_{i4} - w_{i4} \ln q_{i3}\right]$$

Suppose that price and quantities are the same in the fourth period as in the first. Then if the index is to satisfy the circular test, we must have $\ln Q_4^{ChT} = 0$. Now the first term in square brackets on the right hand side is indeed zero, under the assumption that prices and quantities are the same in the first and fourth periods. But if prices and quantities can vary freely, then there is no guarantee that the other three terms will net out to zero. Hence the index will fail the circular test. The operative words here are: "if prices and quantities can vary freely". It may be that economic theory can put constraints on behaviour which will make these other terms net out to zero, so that the chain index is circular after all.

Hulten (1973) proved that Divisia index numbers do not in general satisfy the circular test. Divisia index numbers are line integrals and mathematically circularity is equivalent to path independence, which is not generally true of line integrals. But he also showed that path independence is satisfied under certain restrictions. If the production frontier or utility function is homothetic, then circularity is satisfied.[16] In the case of utility functions, this implies that all income elasticities are equal to 1.

An equivalent result for discrete time index numbers was proved by Samuelson and Swamy (1974) and discussed by Diewert (1976). Consider a quantity index number which is exact for some aggregator function *f*(*q*), where *q* is a quantity vector. Denote the quantity index number relating output at time *b* to output at time *a* by

---

[16] In the case of production functions there is another reason why homotheticity may fail to hold, even in the case of constant returns to scale. This is when technical progress is biased rather than Hicks-neutral. Now the shape of the production function changes as it shifts out over time. And the estimated growth rate of TFP is a weighted average of

$$Q_{ab} = Q(p^a, p^b; q^a, q^b)$$

This index is a function of prices and quantities in the two periods. Suppose that this index number is exact. Then

$$Q_{ab} = f(q^b)/f(q^a)$$

Now consider the relationship between two consecutive one-period index numbers:

$$Q_{12} \cdot Q_{23} = \frac{f(q^2)}{f(q^1)} \cdot \frac{f(q^3)}{f(q^2)} = \frac{f(q^3)}{f(q^1)} = Q_{13}$$

That is, this index passes the circular test.

The assumption of homotheticity is perhaps tolerable for production frontiers; but homotheticity is not very attractive for utility functions. However Diewert (1987) argues that deviations from circularity are in practice small for superlative indices. He also states an approximation theorem, though unfortunately without giving the proof: a superlative index satisfies the circular test to first order. That is, if we compare a superlative index calculated directly between the first and last periods with the corresponding chain index (calculated over all the intervening periods as well), then the first derivatives of these two indices are equal; these derivatives are to be evaluated at a point where all prices and quantities are the same.

Proof of Proposition 8

The general proposition is that fixed base and chain indices will yield identical results if relative prices are constant over the relevant periods. We will prove this for Laspeyres and Paasche indices, and as a consequence, for Fisher indices as well. Compare first a chained with a fixed base Laspeyres quantity index over 3 periods (1, 2 and 3):

---

the growth rates of the efficiencies of the inputs, where the weights are the input shares. But these shares depend on the path of input prices, so the TFP index is path-dependent.

Fixed base Laspeyres for period 3 relative to period 1:

$$Q_{13}^{L} = \frac{\sum_i p_{i1} q_{i3}}{\sum_i p_{i1} q_{i1}}$$

Chained Laspeyres for period 3 relative to period 1:

$$Q_{13}^{ChL} = Q_{23}^{L} \cdot Q_{12}^{L} = \frac{\sum_i p_{i2} q_{i3}}{\sum_i p_{i2} q_{i2}} \cdot \frac{\sum_i p_{i1} q_{i2}}{\sum_i p_{i1} q_{i1}}$$

where $p_{i1}, p_{i2}, q_{i1}, q_{i2}, q_{i3}$ are the prices and quantities in periods 1, 2 and 3.

Suppose that relative prices are constant between periods 1 and 2, ie $p_{i2} = kp_{i1}$, where $k > 0$ is some constant. In this case, we can see that the chained and fixed base Laspeyres indices are equal, irrespective of what is happening to the volumes:

$$Q_{13}^{ChL} = \frac{k\sum_i p_{i1} q_{i3}}{k\sum_i p_{i1} q_{i2}} \cdot \frac{\sum_i p_{i1} q_{i2}}{\sum_i p_{i1} q_{i1}} = \frac{\sum_i p_{i1} q_{i3}}{\sum_i p_{i1} q_{i1}} = Q_{13}^{L}$$

By a parallel argument, one can show that a chained and fixed base Paasche quantity index will be identical if relative prices are constant between periods 2 and period 3. It then follows that a chained and fixed base Fisher index will be identical if relative prices are constant in all three periods.

Proof of Proposition 9

(a) Proof that $\hat{I} > \hat{A}$

Consider first the growth rates of aggregate gross investment and of the aggregate capital stock:

$$\hat{I} = u_d g_d + u_e g_e, \ \ u_d, u_e > 0, \ \ u_d + u_e = 1$$

$$\hat{A} = w_d g_d + w_e g_e, \ \ w_d, w_e > 0, \ \ w_d + w_e = 1$$

Here u_d, u_e are the shares of each type of investment in aggregate nominal investment and w_d, w_e are the shares of each asset in the aggregate nominal capital stock. Note that as we are assuming a steady state, the growth rate of dull investment is the same as that of the dull capital stock, and similarly for exciting capital.

We will show that in this model $u_e / u_d > w_e / w_d$. Let us take this for granted for a moment. Then it is easy to see that it implies $u_e > w_e$. The reason is that if $u_e / u_d > w_e / w_d$ then $u_e /(1-u_e) > w_e /(1-w_e)$ and if we assume to the contrary that $u_e \le w_e$, we find a contradiction. Hence $u_e > w_e$ as asserted. It then easily follows that $\hat{I} > \hat{A}$ given that by assumption $g_e > g_d$.

It remains to show that $u_e / u_d > w_e / w_d$. Let the prices of dull and exciting capital by p_d, p_e. Then

$$\frac{u_e}{u_d} = \frac{p_e I_e}{p_d I_d} = \frac{p_e I_e}{p_d I_d}\frac{A_e}{A_d}\frac{A_d}{A_e} = \frac{p_e A_e}{p_d A_d}\frac{I_e / A_e}{I_d / A_d}$$

$$= \frac{w_e}{w_d}\frac{g_e + \delta_e}{g_d + \delta_d} > \frac{w_e}{w_d}$$

The second line uses the fact that for any asset $I / A = g + \delta$ in steady state. This completes the proof that $\hat{I} > \hat{A}$.

(b) Proof that $\hat{I} > \hat{Y}$ and $\hat{A} \ne \hat{Y}$

In steady state the gross savings ratio *s* must be constant. Hence consumption (*C*) must grow at the same rate as investment in dull capital. So

$$\hat{Y} = (1-s)\hat{C} + s\hat{I} = (1-s)g_d + s[u_d g_d + u_e g_e]$$

$$= (1-su_e)g_d + su_e g_e < \hat{I}$$

Comparing the growth of GDP with that of the capital stock, we can see that there is no presumption that the one will necessarily grow faster than the other and there is no reason to expect the rates to be equal except by a fluke. If consumers are patient, they will prefer to have a high capital-output ratio in steady state, hence the (gross) saving rate will have to be high as depreciation will be a larger proportion of income. The higher the savings rate, the faster does GDP grow, so the more likely that it grows faster than the capital stock.

(c) Proof that $\hat{K} > \hat{A}$

The growth rate of the VICS (K) in this model is

$$\hat{K} = v_d \hat{A}_d + v_e \hat{A}_e = v_d g_d + v_e g_e, \quad v_d, v_e > 0, \; v_d + v_e = 1$$

where v_d, v_e are the shares in total profits generated by each asset. Applying the Hall-Jorgenson cost of capital formula, the ratio of these shares is

$$\frac{v_e}{v_d} = \frac{r + \delta_e - (\hat{p}_e - \hat{p}_d)}{r + \delta_d} \frac{p_e A_e}{p_d A_d}$$

$$= \frac{r + \delta_e - (\hat{p}_e - \hat{p}_d)}{r + \delta_d} \frac{w_e}{w_d} > \frac{w_e}{w_d}$$

Here r is the real rate of interest in terms of consumption, ie the nominal rate minus the growth of the price of consumption. The last line follows since we are assuming that the relative price of the exciting capital good is falling and that $\delta_e > \delta_d$. Hence by a similar argument to the one in part (a) $v_e > w_e$ and so $\hat{K} > \hat{A}$.

We can get a sharper result if we assume that the representative household maximises the present value of utility and utility is logarithmic, ie the household solves

$$\max U = \int_0^\infty \ln(C) e^{-\rho t} dt, \quad \rho > 0$$

The first order conditions for this problem imply that in steady state

$$r = g_d + \rho$$

Also, since shares are constant, eg $u_e / u_d = p_e I_e / p_d I_d$ is constant, it follows that

$$\hat{p}_e - \hat{p}_d = g_d - g_e$$

Plugging these last two results into the formula above for v_e / v_d, we get

$$\frac{v_e}{v_d} = \frac{\rho + g_e + \delta_e}{\rho + g_d + \delta_d} \frac{w_e}{w_d}$$

Now compare this with the formula for u_e / u_d derived in part (a). We see that $u_e / u_d > v_e / v_d$, hence $u_e > v_e$, so $\hat{I} > \hat{K}$. Therefore with logarithmic utility we have

$$\hat{I} > \hat{K} > \hat{A}$$

ANNEX C
THE AXIOMATIC APPROACH TO INDEX NUMBERS

This section is largely based on Balk (1995). In the axiomatic approach, we lay out a set of "reasonable" properties that we would like a price or quantity index number to possess. We then see whether these properties suffice to determine the index number uniquely. Traditionally, this approach has been applied to price indices, but as we shall see, it applies equally to quantity indices. We start with a definition of a price index, followed by a set of axioms.

Definition. A price index is a mathematical function of N prices and N quantities ($N > 0$) in two periods, the base period and the second or comparison period. It satisfies the following axioms:

Axiom 1. Monotonicity. Take base period prices and quantities and second period quantities as given. Consider two alternative second period price vectors, A and B, which are identical except that in vector A, one price is higher. Then the value of the price index using vector A is higher than its value using B.

Axiom 2. Linear homogeneity. Take base period prices and quantities and second period quantities as given. Consider two alternative second period price vectors, A and B. In A, each price is the same multiple h (> 0) of its value in the other vector B. Then the value of the price index using vector A is h times the value using vector B.

Axiom 3. Identity. If all prices remain constant, the value of the price index in the second period equals one, irrespective of any change in the quantities.

Axiom 4. Homogeneity of degree 0 in prices. If all prices are multiplied by a common (positive) factor in both periods, then the price index is unchanged. This covers the case of a change in the currency unit in which prices are measured, eg from dollars to cents or from euros to centimes.

Axiom 5. Dimensional invariance. If we change the quantity unit, and make a corresponding change to the unit prices, then the price index is unchanged. Eg suppose that the quantity unit is the kilo and prices are measured in centimes per kilo. This axiom says that if we change the quantity unit to grams, and measure prices in centimes per gram, then the price index is unchanged.

These axioms are self-evident, in the sense that if a function violated them, then we would not consider it to be a price index. We can set up exactly analogous axioms for quantity indices. The five axioms can be shown to be independent of each other: no one axiom can be derived from the other four. Also, the axioms can be shown to imply some further, highly intuitive properties:

1. *Proportionality*. If each price in the second period is the same multiple h (> 0) of the corresponding price in the base period, then the value of the price index in the second period is h. This follows from axioms 2 and 3.

2. *Homogeneity of degree zero in quantities*. If all quantities are multiplied by a common factor in both periods, then the price index is unchanged. This follows from axioms 4 and 5.

3. *The price index lies between the smallest and the largest of the price ratios (price relatives).* This follows from axioms 1, 2 and 3.

In addition to the axioms, a number of tests for price indices have been proposed. These tests are not included as axioms since they are not self-evident. In fact, as we shall see, some tests are inconsistent with the axioms. The most important of the tests are:

T1. Time reversal test This test states that the price index number for period 1 relative to period 0 should be the reciprocal of the index number for period 0 relative to period 1. Suppose that prices rose by 25% between periods 0 and 1. Then if time had run backwards (ie the price and quantity vectors for periods 0 and 1 were interchanged), the price index would have fallen by 20% (the index would have been 0.80 instead of 1.25).

Indices like the Fisher and the Törnqvist, which give equal weight to the pattern of expenditures in the two periods, pass the time reversal test. Indices like Paasche and Laspeyres, which privilege the expenditure pattern of one period, fail this test.

T.2 Circular (or transitivity) test Let the base period be labelled period 0 and two subsequent periods be labelled 1 and 2 respectively. The circular test says that the price index for period 2 relative to period 0 should be the product of the price index for period 1 relative to period 0 and the index for period 2 relative to period 1. This test tells us for example that if prices rose by 5%

in period 1 and by a further 2% in period 2, then the index calculated directly between periods 0 and 2 should rise by a factor of 1.05 x 1.02 = 1.071, ie by 7.1%.

These two tests apply to quantity indices as well. An important test involving both quantity and price indices is the product test:

T3. Product test The product of the price index and the quantity index should equal the expenditure index (the ratio of expenditure in period 1 to expenditure in period 0).

To understand this test, note that we can always define a quantity index as expenditure deflated by the price index. If the price index satisfies the five axioms, then so will the quantity index derived in this way. And the converse proposition holds too: if we start with a quantity index satisfying the five axioms and derive a price index as expenditure divided by the quantity index, then this price index will satisfy the five axioms too (this is easily checked by inspecting the axioms). In other words, if we find a form for the price (quantity) index with which we are happy, then we can always find a quantity (price) index satisfying the five axioms and the product test. The product test seems a very natural one, fully consistent with our common sense notion of the relationship between price and quantity indices. It simply generalises to indices what is true of a single commodity: price times quantity equals value.

A stronger form of the product test is the factor reversal test. This requires in addition that the quantity (price) index should have the same functional form as the price (quantity) index. The Fisher index passes the factor reversal test but other superlative indices such as the Törnqvist do not: the product of a Törnqvist price and a Törnqvist quantity index does not in general equal the expenditure index. However, in my view the factor reversal test is too strong and satisfies an aesthetic rather than an economic requirement. We know for example that a Törnqvist input quantity index corresponds to a translog production function and that under constant returns to scale a Törnqvist input price (unit cost) index corresponds to a translog cost function. But the translog production and cost functions are not dual to each other. So the Törnqvist price and quantity indices fail the factor reversal test. But so what? Life would be easier for economists if the Törnqvist passed the factor reversal test, but the fact that it fails does not reveal any flaw in economic theory. From an economic point of view, both the translog production and the translog cost functions meet all the requirements of economic theory. All that failing the test reveals is that, though each is both an exact and a superlative index in the sense of Diewert (1976), nevertheless they cannot both be precisely true of the same set of economic facts.

Let us stick to the less demanding product test. Unfortunately, it has been proved that there is a fundamental inconsistency between the product test, the circular test, and axiom 3 (identity): there are no price and quantity indices which can satisfy all three of these requirements.[17] Faced with the necessity of dropping one of the requirements, most students of index numbers have thought it best to abandon the circular test.

What does this inconsistency really mean? Under the axiomatic approach, no assumptions are made about economic behaviour. So prices and quantities are free to vary in an arbitrary manner. Inconsistency is telling us that it is possible to find combinations of prices and quantities which violate the axioms plus the circular and product tests. But if quantities are restricted to vary in response to prices, in a manner suggested by economic theory, then it is possible that the inconsistency will vanish. The text and Annex B discuss this possibility in the context of both Divisia index numbers and their discrete counterparts. It is shown that homotheticity of the relevant function (production, cost, utility, or expenditure function) can ensure that the circular test is not violated.

---

[17] Balk (1995) gives a proof assuming four time periods, based on the work of Abraham Wald.

ANNEX D
THE ECONOMIC APPROACH

D.1 Consumption[18]

In consumer theory, the utility function summarises all we know about a consumers' preferences. The consumer is assumed to maximise utility subject to a budget constraint. Duality theory teaches us that the expenditure function is dual to the utility function: all we know about preferences can equally well be expressed in the expenditure function which tells us the minimum cost of reaching any given level of utility, given the set of prices faced by the consumer. Given that the consumer maximises

$$u = u(q_1, q_2, \ldots, q_n)$$

subject to the budget constraint $E = \sum_{i=1}^{n} p_i q_i$, the solution to this problem defines the expenditure function $E = E(p,u)$. This function shows the minimum cost of achieving utility level *u* at the price vector *p*.

A "cost of living index" between period 0 and 1 can then be defined as the ratio of the minimum cost of achieving some given utility level $\bar{u}$ at the prices obtaining in period 1 (p^1), relative to the minimum cost of attaining the *same* level of utility at the prices prevailing in period 0 (p^0):

$$E(p^1, \bar{u}) / E(p^0, \bar{u}) \qquad \text{(D.1)}$$

In principle, given enough data, we could estimate the parameters of the expenditure function or equivalently of the utility function and then calculate the cost of living index. The economic approach suggests that we should call the result the consumer price index. The quantity of consumption can then be defined (in accordance with the product test) as expenditure deflated by the consumer price index.

In general, for given price vectors, the cost of living index depends on the particular level of utility chosen as the reference. But there is a special case where the cost of living is independent

[18] This section draws on Deaton and Muellbauer (1980).

of the reference level. This is when the expenditure function is separable in prices and utility: $E(p,u) = c(p)u$, where $c(p)$ is the cost of attaining a unit of utility. In this case

$$E(p^1,\bar{u})/E(p^0,\bar{u}) = c(p^1)\bar{u}/c(p^0)\bar{u} = c(p^1)/c(p^0)$$

This special case is where preferences are homothetic: all income elasticities are equal to one.

The dependence of the cost of living index on the reference level of utility may seem a bit mysterious but can be given a simple intuitive explanation. Consider an individual at a very low standard of living, who is spending say 60% of his income on food. Suppose that the price of food rises by 10%. Then to keep him at the same level of utility will require that his income be raised by close to (0.6 x 10 =) 6%, since he has limited opportunities for substituting clothing or shelter for food. Now consider the same individual at a much higher standard of living, spending say only 20% of his income on food. Faced with same 10% rise in the price of food, he will require at most an increase in income of (0.2 x 10 =) 2%. In fact, he will probably need a good bit less than this. By cutting his order from jumbo to regular fries, he will be able to afford to rent more DVDs in each month. So the rise in income necessary to keep him at the same utility level could work out to be considerably less than 2%.

How does a theoretical cost of living index such as the above compare with real life price indices like the Laspeyres or the Paasche? The standard result here is that

$$P_{01}^L > E(p^1,u^0)/E(p^0,u^0); \quad E(p^1,u^1)/E(p^0,u^1) > P_{01}^P$$

That is, the Laspeyres price index (P_{01}^L) overstates the cost of living change when the *initial* level of utility is the reference (this is called the substitution bias of the Laspeyres), while the Paasche price index (P_{01}^P) *understates* the change when the *second* period level of utility is the reference. In the homothetic case, we get the stronger result

$$P_{01}^L > c(p^1)/c(p^0) > P_{01}^P$$

ie the Laspeyres and the Paasche bound the true price index. Note however that this result is for a single consumer, whereas the theory is usually applied at the aggregate level. Even if all consumers behave in accordance with theory, it does not necessarily follow that the aggregate will behave like a representative consumer. With real life data, we cannot even be sure that

$$P_{01}^{L} > P_{01}^{P}$$

though this will be true if prices and quantities are negatively correlated as is generally found to be the case.

Since at best the Laspeyres and Paasche bound the true price index, are there any indices which can do better? This question is answered by the theory of flexible functional forms and exact index numbers. But before turning to this we consider index numbers in production.

D.2 Production

The counterpart in production theory to the utility function is the production function, to which the cost function is dual. For a given level of technology and a given set of input prices, the cost function measures the minimum cost of producing a given level of output. The cost function can be used to define an input price index, in the same way that the expenditure function is used to define a consumer price index. If there are constant returns to scale, then the input price index is independent of the level of output. If $c(p)$ is the minimum cost of producing a unit of output at input prices p (given the technology level), then the input price index at period 1 relative to period 0 can be defined as

$$c(p^{1})/c(p^{0})$$

The quantity of inputs can then be defined (in accordance with the product test) as total cost deflated by the input price index.

Suppose the economy consists of a number of industries producing different products. Standard theory shows that if all producers maximise profits under perfect competition, then the value of output will be maximised at the competitive prices, given the non-produced input levels (eg labour, natural resources and inherited capital). Alternatively, we can think of the economy as maximising the output of any one product, given the output of all the other products and the non-produced inputs: this defines a production possibility frontier. These relationships can be used to define an output price index. The quantity of output can then be derived by deflating the value of output by the output price index.

An alternative approach is to define a quantity index as a measure of the distance between the production possibility frontier in two different time periods: this leads to the so-called Malmquist quantity index. Of course, there is no unique number measuring this distance, since the frontier need not shift out in a radial fashion. It can be shown that, if there are constant returns to scale, the Malmquist measure and the deflated expenditure measure coincide (Diewert, 1987).

D.3 Flexible functional forms and exact index numbers

Diewert (1976) employed the expression "aggregator function" as a general term to cover production frontiers, production functions, utility functions and cost functions. He studied aggregator functions that took the form he called a "general quadratic mean of order *r*":

$$f_r(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} \left[a_{ij} x_i^{r/2} x_j^{r/2} \right]^{1/r}, \ a_{ij} = a_{ji}, \ r \neq 0$$

Here the x_i can be interpreted as prices or quantities depending on the context. Note that $f_r(x)$ is linearly homogeneous: doubling all the x_i doubles $f_r(x)$.

Suppose we believe that $f_r(x)$ is a good approximation of the true function. We wish to use the aggregator to calculate the change say in the cost of living. To evaluate an aggregator function it would seem necessary to know the values of the parameters. For a quadratic mean of order *r*, there are $(n^2 - n)/2 + n = n(n+1)/2$ independent parameters. Suppose there are 600 products, a not unreasonable number for a consumer price index. Then to estimate the parameters econometrically we would need monthly data extending over more than 15,000 years! However, Diewert (1976) showed that there exist index numbers which can calculate the change in $f_r(x)$ *exactly*. That is, there is no need for econometric estimation, since by definition an index number is a function solely of observable prices and quantities. More precisely, Diewert showed:

(1) $f_r(x)$ is a "flexible functional form", ie it is a second order approximation to any linearly homogeneous aggregator which is consistent with economic theory. [One function approximates

another to the second order if, at a given point, the values of the two functions, and the values of their first and second derivatives, are equal to each other].

(2) Suppose that economic agents maximise or minimise $f_r(x)$ subject to a budget constraint. Then there are index numbers that are "exact" for $f_r(x)$. A quantity index number is a function of prices and quantities in the two periods compared. In the quantity case exactness means that

$$Q_r(p^0, p^1; q^0, q^1) = \frac{f_r(q^1)}{f_r(q^0)}$$

where p^0, p^1, q^0, q^1 are the price and quantity vectors in the two periods and the quantity index numbers take the form

$$Q_r(p^0, p^1; q^0, q^1) = \frac{\left[\sum_{i=1}^{n} (q_i^1 / q_i^0)^{r/2} (p_i^0 q_i^0 / p^0 \cdot q^0)\right]^{1/r}}{\left[\sum_{j=1}^{n} (q_j^0 / q_j^1)^{r/2} (p_j^1 q_j^1 / p^1 \cdot q^1)\right]^{1/r}}$$

(3) When $r = 2$, the quantity index numbers are Fisher. The corresponding price index numbers are also Fisher when $r = 2$.

(4) If we take the limit as r goes to zero, the aggregator converges to the translog form. Then for this form Törnqvist indices are exact.

(5) If an index number is exact, then it passes the circular test. (This is proved as Result 5 of Annex B).