Consumer price indices and the identification problem

Courtney, Mark

2013

Statistical Journal of the IAOS 29 (2013) 41–51
DOI 10.3233/SJI-130757
IOS Press

Consumer price indices and the identification problem

Mark M. Courtney
28 Ridgway Place, London SW19 4EP, UK
Tel.: +44 20 8946 7960; E-mail: mark.courtney@oxon.org

Abstract. Conventionally, consumer price indices are constructed on the assumption that we are observing a stable system of consumer demand, and that all price movements are, therefore, the result of supply-side changes. This often leads to an emphasis on consumer price substitution and to a recommendation that we should allow for it by using the geometric mean for first-stage aggregation.

This paper argues, on the basis of economic theory and from observations on the UK clothing sub-index, that demand-side changes are also important in generating price movements. For most items we are unable to solve the resulting identification problem of whether supply-side or demand-side influences predominate: in these circumstances, the appropriate formula to use for first-stage aggregation is one that makes no assumptions about the cause of price changes – i.e. one that uses an arithmetic rather than a geometric average. Allowing for both sources of price movements also affects the way in which elementary aggregates should be defined: this should be on the basis of both demand and supply characteristics, in order to minimise problems that arise when aggregating disparate products.

Keywords: Consumer price indices, first-stage aggregation, identification problem, demand-side changes, geometric mean, elasticity of substitution

1. Introduction

This paper is concerned with the choice of methods for first-stage aggregation in the construction of consumer price indices – the stage at which a number of price quotes for each item in the index are combined to get a price index for that item. This is a more difficult and more controversial choice than for upper-level aggregation, where the elementary aggregates are combined to get an overall consumer price index.

This is because, with few exceptions, there are no data – and certainly no timely data – on expenditure below the elementary aggregate level. Therefore, unlike for upper level aggregation, aggregation at the elementary level has to use unweighted prices. This means that there is always a degree of approximation, because there can only be stylized (and usually implicit) assumptions about what quantities are associated with the various price quotes and, therefore, about the importance that should be assigned to each price quote. This is where the identification problem becomes important, because it is usually impossible to know whether the observed price changes are being driven by changes in supply or changes in demand – and the two lead to very different results for the relationship between prices and quantities. Conventionally, the assumption in work on consumer price indices has been that price changes are driven only by changes in supply, but the usually neglected demand-side influences have important implications for the construction of consumer price indices.

The paper is organised as follows:

Section 2 summarises the implicit economic assumptions that are usually made in constructing consumer price indices.

Section 3 expands on this by adding the possibility of demand-side changes and provides a discussion and numerical example of the implications for the various formulae for first-stage aggregation. (An algebraic model is set out in the Annex to the paper).

Section 4 discusses the prevalence of demand-side changes and draws on recent experience in the UK, where demand-side changes have impacted markedly on the clothing component of the consumer price index.

Section 5 offers some suggestions for consumer price index construction in those cases where both supply and demand changes are at work and the identification problem becomes significant.

Finally, Section 6 makes some concluding remarks.

2. The economic assumptions in the standard approach

Various formulae have been developed for aggregating the price quotes into first-stage price indices and there are two broad approaches for deciding which formula to use: the axiomatic approach and the economic approach.

The axiomatic approach lists various statistical properties which it is desirable for an index to have. This approach is discussed in detail in Chapter 16 of the ILO Consumer Price Index Manual [9] and this paper will not deal further with the axiomatic approach. It is important both for choosing aggregation formulae and for indicating when practical adjustments to item definitions, price collection or index chaining might be desirable. But these choices can only come into play if the price index meets the basic economic criteria, and it is to these we now turn.

The economic approach starts by asking what is the purpose of the price index. There are various possible purposes, and the three most common are listed below:

(i) *A pure price index*. In this case, the consumer price index is trying to measure how the price of a fixed basket of goods has changed from one period to another. Although it is usual to construct a basket that is broadly representative of consumer purchases and to revise it from time to time, there is no necessary implication that consumers buy exactly this basket of goods in the initial period or that they buy the same basket in the subsequent period as in the initial period. In this case, there is no scope for the application of economic theory. What one can say is that the requirement for maintaining a fixed basket excludes price aggregation formulae, such as the geometric mean, that implicitly assume that the quantities purchased vary systematically with prices. Such a pure price index therefore requires one of the arithmetic averages for first-stage aggregation [18].

(ii) *An index used as a guide for monetary policy*. In recent decades there has been a growing use of inflation targeting, where monetary policy is directed towards keeping inflation within a target range. Inflation targeting may or may not be associated with the formal use of the rate of growth of the money supply as an indicator of inflationary pressure. But, in any case, a question of interest to the monetary authorities is the extent to which the measured price index signals an increase in monetary expenditure beyond the underlying rate of growth of the real economy. In this case, therefore, any systematic relationship between prices and quantities bought becomes important. For example, if price rises are systematically associated with decreases in quantities bought, then this means that a simple arithmetic average of price movements will overstate the extent to which higher prices are, in fact, leading to higher expenditure, which, in turn, would need to be financed by a higher money supply (or a higher velocity of circulation of money). If there were timely data on expenditure, an explicit *ex post* adjustment could be made but, in its absence, one must rely on stylised assumptions, based on economic theory and whatever empirical evidence is available.

(iii) *A cost of living index*. Historically, this was the first purpose for which consumer price indices were constructed, to see whether increases in wages were keeping up with the workers' cost of living. If the cost of living in question is some fixed or minimum standard, then the index in question should be a fixed basket index as described under the pure price index. However, it is more usual to calculate an index that is representative of all consumer expenditure, using plutocratic weights (i.e. weighting each individual's price experience by the amount she spends). In this case, it is legitimate to observe that, as a matter of fact, the basket of goods that is bought by consumers in the second period will not be exactly the same as the basket they bought in the first period. Hence, there is an ambiguity about whether, in calculating the increase in wages required to maintain the workers' purchasing power, one should work with the basket they (collectively) bought in the first period or in the second, or some average of the two.

From this extremely abbreviated discussion, it appears that it is useful, in many cases, to draw on the economic theory of how prices and quantities are related.

This is the justification for the economic approach to first-stage aggregation. It is usual to move on immediately to a discussion of consumer substitution, assuming that we are dealing with a representative consumer with stable preferences, who faces an exogenously given set of prices [10]. In other words, we assume that we are observing a set of prices and associated quantities that trace out or identify the consumer's stable demand curve.

As already mentioned, for products below the elementary aggregate level, we do not typically have data on quantities, so we turn to economic theory which tells us that, for normal goods, the consumer's demand curve is downward sloping – i.e. higher prices are associated with lower quantities bought. If there is a price rise, the consumer is assumed to substitute away from this product to another similar product (within the same elementary aggregate) whose price has risen relatively less.

What does this tell us about which formula to use for aggregation? The theoretical discussion is usually conducted for the case in which we know both prices and quantities, because that gives greater precision, so we will reproduce this first, before turning to the application to first-stage aggregation where we do not have data on quantities.

Consider a Laspeyres price index, which takes an arithmetic average of the price relatives (the ratio of the price in the second period relative to the price in the initial period) weighted according to expenditure in the *initial* period. If the quantities bought are, in fact, unchanged from the initial to the subsequent period, then the Laspeyres index will represent a true cost of living index, because there will then be no ambiguity about whether to use the initial or subsequent period baskets, since these are identical. (For the same reason, the Laspeyres index will also give a true indication of inflationary pressure – the extent to which higher prices affect monetary transactions).

However, if we assume a degree of consumer substitution, then in the subsequent period products with relatively high price rises will see a fall in the quantity bought, and products with price falls (or relatively low price rises) will see a rise in the quantity bought. Let us assume, for the sake of concreteness, that overall there are more price rises than price falls, so that price rises predominate and that therefore consumers on a fixed income would be worse off. Compensating them by increasing their income in line with the Laspeyres price index will allow them to buy the initial basket of goods at the new, second period set of prices. However, because of the changes in *relative* prices, consumers in the second period have, in fact, indulged in substitution and have bought relatively more of the goods that now have relatively cheaper prices. This means that with their income increased in line with the increase in the Laspeyres price index, the consumers will be able to buy exactly the same basket they bought in the second period and have some money left over. To that extent, the Laspeyres index has over-estimated the increase in the cost of living. (Alternatively, looked at from the point of view of monetary policy, an increase in the money supply in line with the Laspeyres price index will, at an unchanged velocity of circulation of money, be more than sufficient to finance the purchases actually made in the second period).

Let us now take a retrospective view and construct a price index that takes an arithmetic average of the price relatives, weighted according to the quantities bought in the *second* period. This is the Paasche price index. Assuming the existence of some consumer substitution, the Paasche index gives a greater weight to products whose price has fallen – or risen relatively slowly – from the initial period to period 2, so it will be below the Laspeyres index. If we start out in the initial period and increase incomes in line with the Paasche index, then, if the initial quantities *happened* to be the same as those bought in the second period, this compensation would enable the consumers to buy exactly what they did buy in the second period. But, in fact, the quantities purchased in the initial period will be determined by the initial period prices: specifically, there will be greater purchases (relative to second period purchases) of the products that were cheaper in the initial period and smaller purchases of the products that were more expensive in the initial period. This means that the income base that is uprated according to the Paasche index is smaller than would be necessary for the Paasche uprating formula to enable the purchase of the actual second period basket. The Paasche index thus offers less than complete compensation (or, in monetary terms, underestimates the increase in the quantity of money needed to finance expenditure at the new prices).

This gives us the result that a true cost of living index – one which, applied to expenditure in the initial period would allow the actual expenditure in the subsequent period to be financed – is bounded by the Laspeyres and Paasche indices. Intuitively, a symmetrical average of the two – such as the Drobisch-Sidgwick-Bowley index, which is their arithmetic average, or the Fisher's Ideal index, which is their geometric average – should give a good approximation of the true cost of living index.

Instead of using an arithmetic average of the individual price relatives, as in the Laspeyres or Paasche indices, it is possible to use a geometric average. If it is weighted by expenditure in the initial period it is the counterpart to the Laspeyres index, and if it is weighted by expenditure in the subsequent period it is the counterpart to the Paasche index. Because a geometric sum is always less than the equivalent arithmetic sum, the geometric index will always be below the corresponding arithmetic index.

Let us return now from the general case, where the quantities purchased are assumed to be known, to the problem of first-stage aggregation. Here, it is a question of aggregating the *unweighted* prices, while making the most realistic assumptions about what is happening to the unobserved quantities. Under certain, strict assumptions, these unweighted averages will have the same properties as the weighted averages, and will approximate to them if the assumptions are not too far wrong.

For unweighted averages, it makes a difference in calculating the arithmetic average, whether one expresses the index as the average of price relatives (the Carli index) or the ratio of averages (the Dutot index). If the *quantities* bought of every product are unchanged from one period to the next, and if the *expenditure* on each is equal in the initial period, then the Average of Relatives (Carli) will yield a true cost of living index. If the *quantities* bought of every product are unchanged from one period to the next, and if the *quantity* bought of each is equal in the initial period, then the Ratio of Averages (Dutot) will yield a true cost of living index.

In the case of the geometric average – it is one of its attractive properties – the result is the same whether it is specified as the geometric sum of price relatives or as the ratio of the geometric average of prices in the second period to the geometric average of prices in the first period. If *expenditure* on every product is the same in the initial period, and if *expenditure* on each product in the second period is the same as in the initial period (i.e. if the quantity bought moves in inverse proportion to the price), then the geometric index (the Jevons index) will yield a true cost of living index.

Because the Jevons index applies a geometric average to the price relatives, it will always show a lower rate of inflation than the Carli index, which applies an arithmetic average. (There is no such presumption when a comparison is made with the Dutot index, where the ratio of two average prices is taken).

Given these properties of the various indices, the choice of which to use for first-stage aggregation depends on the conceptually simple but in practice difficult judgement about how closely the real world situation corresponds to the stylised assumptions behind each of the formulae.

Of the two dimensions – quantity/expenditure in the initial period, and changes in quantity/expenditure from one period to the next – the compiler of the index has some control over the first. By changing the way in which items are defined or prices are sampled – taking more samples of popular lines – she can, for example, move the sample closer to the requirement for the Ratio of Averages (Dutot) method, that quantities purchased of each product are approximately equal.

However, the index compiler has no such influence over the consumers' reaction to prices from one period to the next. Hence, the economic approach tends to focus on the question of the extent to which the quantities that are bought of each of the products sampled react to changes in their prices. There is, as already mentioned, almost no timely data on this, so the decision tends to be made on intuition informed by a number of more or less relevant empirical investigations – of which, more in Section 4.

If it is believed that quantities react very little to relative price changes (or alternatively, if the reaction is random and not systematic), then one of the arithmetic averages will be appropriate. If, on the other hand, quantities move considerably and systematically in inverse proportion to the price of the product in question, then the geometric average (Jevons) approach is appropriate.

In Economics terms, an arithmetic average will be appropriate if the compensated own-price elasticity of substitution is close to zero, whereas a geometric average will be appropriate if the elasticity of substitution is close to or greater in absolute value than (minus) unity.

It is possible, in theory, to construct an intermediate measure, assuming a constant elasticity of substitution of between zero and one [5]. However, to estimate an elasticity of substitution credibly involves estimating either a complete set of demand equations at a very disaggregated level, or else being able to assume equal within-group substitution elasticities, which the data at the very disaggregated level do not support [1,17]. In practice, therefore, the choice facing national statisticians is either to use one of the arithmetic averages or else the geometric average for first-stage aggregation for each item or sub-index of the consumer price index. Although the choice between the use of arithmetic or geometric averages rests in large part on the elasticity of substitution of the goods concerned, there is very little empirical evidence on the size of these elasticities, with the choice relying mainly on casual empiricism or an appeal to theory.

Historically, only the arithmetic averages were used before the 1990's. This was attributable in part to the computational difficulty of calculating the geometric means in the pre-computer age, and in part to the difficulty of justifying anything other than a fixed-basket consumer price index to the public at large. Although all three formulae are still recognised as usable, twenty out of thirty-four OECD countries now use the geometric mean for some or all of their first-stage aggregation [12]. There are a number of reasons for this. There has been a strong trend towards setting inflation targets as the guideline for monetary policy and that has, naturally, led to more focus on the macroeconomic implications of the various compilation methods. Associated with this has been a growing economic and statistical sophistication and a willingness to move away from easier to understand fixed basket indices. And the lower calculated rate of inflation resulting from the use of the geometric mean rather than the average of relatives has sometimes been useful to governments that were under pressure to meet inflation targets. The 1996 US Senate Committee of Finance Advisory Commission report on the Consumer Price Index (the Boskin Commission report) was influential in this regard, concluding as it did that the CPI as then calculated considerably overstated US inflation and that one appropriate method of reducing it would be to switch to using the geometric mean for first-stage aggregation [19]. Another example is given by the change of the inflation target in the UK in December 2003 from 2.5% on the RPIX measure (the Retail Prices Index excluding mortgage interest payments) to 2.0% on the Harmonised Index of Consumer Prices measure. Given the differences in compilation methods, this actually represented a relaxation of the inflation target but there was speculation that the lower "headline" figure might influence wage bargainers and hence reduce inflationary pressure [11]. In the case of the EU countries, changes to the elementary aggregation method used in their Harmonised Index of Consumer Prices – which were often but not always carried over to their domestic price indices – resulted in reductions in the measured rate of inflation in all seven cases for which estimates are available [6].

3. An expanded economic approach

So far, we have merely summarised the standard approach. However, although it uses the so-called Economic approach, the economics behind it is curiously one-sided. It is a commonplace of economic theory that prices are determined by the interaction of supply and demand, but index-number compilation has concentrated almost exclusively on consumer demand. In looking at the effect of price changes on quantities bought, the implicit assumption has been that the aggregate demand curve for every good is stable over the relevant period and that any observed price change is caused by shifts on the supply side: in other words, the observed prices trace out or identify the demand curve for the good in question. However, in theory, a price change could equally well be brought about by a change in demand, with the supply curve being stable: in that case, the observed prices would trace out the supply curve. This is the identification problem: merely observing prices and quantities does not allow one to identify whether there have been changes in demand, supply or some combination of the two.

Econometricians were well aware of the identification problem and the need to identify the source of price changes when estimates of changes in demand were being made [3]. However, both the theoretical discussion and empirical work on consumer preferences and consumer price indices was entirely in terms of fixed consumer preferences and stable consumer demand curves [5]. This can be accounted for partly by the habit of working upwards from a single, representative consumer, and partly by the complexity of estimating simultaneous supply and demand systems. As discussed above, until the economic approach started being used to justify a move away from fixed weights in the 1990's, the assumption of unchanging demand did not affect the construction of consumer price indices in practice.

But once the economic approach is used, it becomes relevant whether a price change is caused by shifts in demand or in supply. Where there are changes in supply bringing about a movement along a stable demand curve, there will, for a normal good, be an inverse relationship between price and quantity: i.e. a decrease in the relative supply of a good will bring about an increase in its price and a decrease in the quantity purchased. This is the case assumed by the standard approach.

However, if changes in demand bring about a movement along a stable supply curve, there will be a positive relationship between price and quantity: i.e. an increase in the relative demand for a good will lead to an increase both in its price and the quantity purchased.

We have already seen that, in the standard approach, where prices and quantities are inversely re-

Table 1
A reduction in the supply of good x, with unchanged demand conditions

	P_x	Q_x	P_y	Q_y	Laspeyres price index	Paasche price index	Geometric price index
Original situation	100	100	100	100	100	100	100
New situation	122	82	100	100	111.2	110.1	110.7

These results can be generated from the model in the Annex by setting $\alpha = 1/2, \theta = 1, \phi = 1.5$.

Table 2
A shift in demand towards y and away from x, with unchanged supply conditions

	P_x	Q_x	P_y	Q_y	Laspeyres price index	Paasche price index	Geometric price index
Original situation	100	100	100	100	100	100	100
New situation	71	71	122	122	96.6	103.5	93.1

These results can be generated from the model in the Annex by setting $\alpha = 1/2, \beta = 1/4, \theta = 1$.

lated, the Laspeyres price index will always be above the Paasche index, with the Fisher's ideal index between the two. The Geometric price index will, as a matter of arithmetic, always be below the Laspeyres index (and will be equal to the Fisher's ideal index in the special case where there is a Cobb-Douglas utility function, implying a high level of price substitution, and equal expenditure on all goods in the index).

However, where price changes are caused by shifts in demand and where, as a consequence, prices and quantities move in the same direction, these results are altered. The Laspeyres price index will now always be below the Paasche index. (This is because the Laspyres index uses initial period weights, and these now give a smaller weight to the good whose price has risen; whereas the Paasche index, looking backwards, gives a greater weight to the good which has gained in popularity and has both an increased price and increased purchases in the second period). The Fisher's ideal index will still be a weighted index of the Laspeyres and Paasche indices. This means that the Laspeyres index, instead of tending to over-estimate inflation (as measure by the Fisher's ideal index) will tend to under-estimate it. Note, too, that the Geometric price index will still, from the arithmetic of its construction, always be below the Laspeyres' price index and thus will tend to under-estimate inflation even more.

These results are set out algebraically in the Annex for the two-good case. Numerical examples (in a format based on that used by the UK Office for National Statistics [14]) are given above, for two goods, x and y, with prices P and quantities Q. (Note that in these numerical examples, because the expenditure in the initial period on the two goods is the same, the weighted Laspeyres and Geometric indices will be the same as their unweighted variants).

4. How important is the identification problem?

We have seen that, in theory, observed changes in prices may be caused by changes in demand just as easily as they can be caused by changes in supply. And in this case, calculating a Laspeyres price index will under-estimate rather than over-estimate the "true" cost of living price index. A geometric price index will continue to show a lower price rise than the Laspeyres price index and thus will give an even greater under-estimate of the true cost of living index.

How likely is this to be a problem in practice? This identification problem is almost always assumed away in practical index number work. And, where one is dealing with highly aggregated data, this is not unreasonable. Consumer tastes are unlikely to shift significantly in the short run between broad categories of expenditure such as food, fuel, or clothing, and long-run trends can be accounted for by regularly updating the weights that are used, based on expenditure data. Conversely, it is likely that short-run supply-side effects will impact differently on the different broad categories, depending on how sensitive they are to changes in raw materials prices, exchange-rates, interest rates, tax rates etc. Thus, observed price changes will be caused predominantly by supply-side changes, and, consequently, will identify moves along a relatively stable system of demand functions [4].

However, when we look at sub-indices at the lowest level of aggregation, it is a different story. In this case, precisely because the goods involved are more likely to be good substitutes, there will be significant swings in consumer tastes, caused by changes in fashion, advertising, media coverage, brand awareness etc. This means that observed changes in relative prices are quite likely to be caused by demand-side changes, and, if that is what is happening, the observed price changes would, in fact, be identifying a relatively stable system of supply.

Moreover, the simple, two-good model in set out in the Annex and summarised in the preceding section rather underestimates the scope and impact of this sort of demand shift, because it assumes that suppliers have a well-defined supply function, based only on the relative market prices of the two goods. In fact, they have inventories of goods and variable price mark-ups, and so they can be expected to react more aggressively to changes in demand, reducing prices of items that have gone out of favour in order to clear inventories and increasing their profit margin on more fashionable lines. In addition, the whole point of advertising is to influence consumer taste, so that the supplier will be able to sell more of the good at a higher price.

The extent and importance of demand shifts will vary from sector to sector, depending on how quickly particular goods or services gain or lose popularity. Recreation and Culture, Entertainment and Clothing are all sectors where one might expect demand shifts to be important. Even though the theoretical importance of demand shifts was not recognised, difficulties in these sectors have been noted by those involved in the construction of index numbers, with the standard approach giving rise, for example, to a cumulative downward bias within the clothing and footwear group in Israel [15] and in the index for an apparel item in the United States [8].

The recent changes that the UK Office for National Statistics has made to the collection of prices in the clothing and footwear sector have constituted a particularly interesting experiment. It had become increasingly obvious over recent years that there was something wrong with the UK's clothing price statistics, since they showed an implausibly steep long-term decline in clothing prices. The ONS therefore introduced a number of changes to the collection of clothing and footwear prices from January 2010. Its concern was that it had not been picking up a representative sample of clothing prices during sale periods, particularly during January, which is the base month for the index: it had been tending to record the price decline when a previously available product had been discounted in a sale but had been less successful in picking up all the comparable price increase when a product that had previously been in a sale was increased in price thereafter. The ONS therefore increased the sample size and collected price quotes for products in January even if they were available for the first time or if there were an interruption in availability, relaxing the constraint on obtaining like-for-like comparisons. These changes also had the effect of increasing the within-period variance of price quotations [2,13].

The ONS' concern related specifically to the miss-measurement of demand during sale periods. But, as we have seen, a similar long-term under-estimation of clothing inflation would have arisen from a general miss-specification, if demand shifts were an important influence on clothing prices. And the UK's much broader specification of item definitions would make this effect show up more strongly than in other countries [7].

Table 3
UK clothing price indices, annual averages, 2005 = 100

	CPI index garments	CPI index other clothing	RPI index men's outerwear	RPI index women's outerwear	RPI index other clothing†
1996	172.9	122.5	122.6	148.4	99.7
2009	76.8	96.6	93.0	78.4	100.4
2010	75.7	97.2	99.9	83.6	104.3

†"Other Clothing" has a broader definition in the RPI than in the CPI.

As shown in Table 3, there was a marked fall in clothing prices from 1996 to 2009. The UK calculates two major consumer price indices, the Consumer Prices Index (the name given domestically to the EU's Harmonised Index of Consumer Prices) and the Retail Prices Index. They are based on the same price quotations but are aggregated differently. For clothing, the CPI uses the geometric mean for first-stage aggregation, whereas the RPI uses the Average of Relatives version of the arithmetic mean. Therefore, as would be expected, price declines on the CPI measure are, in every case, steeper than on the RPI measure. Moreover, it seems clear from the sub-categories that the price decline between 1996 and 2009 is being driven by the demand side rather than – as would be the standard assumption in index-number construction – by changes in supply. This is because supply side pressures, such as outsourcing production to the Far East, exchange-rate changes, changes in retail sales overheads etc. would be expected to affect garments and other clothing in roughly equal proportions. By contrast, it is plausible that demand-side changes, manifested in price discounting of unfashionable lines and the charging of premium prices for fashionable items, will show up in garments much more than in other clothing. And, if one looks at the RPI data, which differentiate between men's and women's outerwear, one finds that the measured fall in the price of women's outwear – which one would expect to be more subject to changes in demand – has been considerably greater than in men's outerwear.

The data from 1996 to 2009 indicate that demand-side changes were important. They do not, however, enable us to differentiate between the general downward bias caused by prices being driven by changes in demand and the specific problem, unique to the UK, of a miss-recording of sales price data. One might hope that the data available from 2010 onwards would help elucidate the relative strength of the two effects, since the new price-collection method ought to provide a better recording of data during sales: in particular, one would expect to see a narrowing in the gap between garment and "other clothing" inflation. However, the strong upward pressure on the supply side in 2010 from rising raw material prices and the effect of the depreciation of sterling on import prices makes it difficult to isolate the demand side effects.

What one can see clearly from 2010 onwards is the effect of the greater dispersion of price quotes under the new price collection regime. As one would expect, the RPI measure of clothing inflation, using the Average of Relatives for first-stage aggregation, has provided a consistently higher measure of inflation (less deflation) than the CPI measure. Over the period 1996 to 2009, for clothing as a whole, average annual inflation was 2.8% less under the CPI measure than for the RPI measure. However, between 2009 and 2010, the CPI measure showed clothing price inflation as 8.5 percentage points lower than under the RPI measure. Such a large divergence, attributable just to a different technique of first-stage aggregation, points to a continuing problem with the clothing component of the consumer price index.

5. Implications for the construction of consumer price indices

We have seen that, in theory, movements in consumer prices can be caused not just by the supply changes that are usually assumed in index-number construction, but also by changes in demand. Moreover, for some products, such as clothing, changes in demand are likely to be a very significant cause of price movements.

One implication is that one can no longer assume that the existence of consumer price substitution within an elementary aggregate – even if it is strong and extensive – is a valid reason to use the geometric mean for first-stage aggregation. One also needs to consider whether demand-side changes are a significant factor. If they are, then using the average of relatives variant of the arithmetic mean is likely to yield a more accurate measure of inflation, since its tendency to over-estimate inflation in the face of supply-side changes will tend to offset the under-estimate of inflation when demand-side changes are at work. By contrast, use of the geometric mean will lead to a consistent under-estimate of inflation if both supply-side and demand-side influences are operating. An alternative way of describing the difference would be to say that where both demand- and supply-side influences affect prices, the assumption behind the use of arithmetic averages, namely that quantities bought are not systematically related to the observed prices, is more likely to be correct than the assumption behind the use of the geometric average, that the quantity bought is inversely proportional to the observed price.

It is an accepted principle of index-number construction to define the items that constitute the index to be as homogeneous as possible, in order that they may experience similar price movements. In this way, the problems of aggregation are minimised and the use of unweighted averages for first-stage aggregation will be a good approximation for the underlying, unknown, weighted average (whichever formula for first-stage aggregation is used). If the standard assumption of a static demand curve is justified, this is achieved by a fairly tight definition of product characteristics, so that one can be sure that all the products sampled within the item definition will be meeting the same consumer need.

By analogy, if one were considering demand-led movements along a static supply curve, one would define items in a way that grouped together products with a similar supply response – e.g. one might distinguish imported from home-produced products.

In practice, changes in demand are going to exist simultaneously with changes in underlying supply conditions and therefore the definition of items will have to be a compromise. It is clear from the problems with the UK clothing indices that, where demand-side changes are significant, defining an item solely by function leads to an unacceptably large divergence of price movements, even if one narrows down the physical characteristics of the definition, for example by specifying material and finish. One possible way forward would be to partition the item according to supply characteristics, e.g. whether or not it is a "designer label" and whether sold in a full-service or discount shop. In this way the products between which there is significant competition are more likely to be grouped together.

A difficult problem arises where periodic sales and discounting are prevalent for the item in question, since both consumer and supplier behaviour are likely to differ between long-term trends and short-term sales. For example, there might be significant price substitution between two brands of designer skirt, such that – if they are equally highly rated in the fashion stakes –

a long-term price cut for one brand will win it an increase in market share. However, during sale periods suppliers are more likely to offer deep discounts on products that are currently unfashionable, and the observed relationship between a price cut and increased sales will be attenuated. Thus, having a large proportion of sales taking place during sales periods is another contra-indication for the use of the geometric mean for first-stage aggregation. One way of mitigating this problem might be to focus on another dimension of homogeneity, namely similarity in unit price. This would justify use of the Dutot (Ratio of Averages) index for first-stage aggregation, together with some relaxation of the requirement of like-for-like comparisons.

6. Conclusion

Recognising that prices are influenced by both supply and demand factors adds a degree of realism to the stylised assumptions that underlie the standard approach to consumer price index construction. This means that for most items, where there is an absence of empirical evidence one way or the other, the appropriate formula to use for first-stage aggregation is one that is agnostic about whether supply or demand changes predominate. It also means that the definition of elementary aggregates needs to be sensitive to both demand and supply characteristics, to minimise problems that arise when aggregating disparate products.

Acknowledgments

The author would like to thank Richmond, the American International University in London for the provision of research facilities. He would also like to thank Michael Baxter, Jill Leyland and Edward Tower for comments which have improved the paper – remaining faults are the author's.

References

[1] B.M. Balk, On curing the CPI's substitution and new goods bias, in: *proc. Ottawa Group fifth meeting*, 1999, available at http://www.statice.is/ottawa/balk.pdf.
[2] Bank of England, *Quarterly Inflation Report,* February 2011, p. 39.
[3] F.W. Bell, The Pope and the price of fish, *American Economic Review* **58** (1968), 1346–1350.
[4] S.D. Braithwait, The substitution bias of the Laspeyres Price Index: An analysis using estimated Cost-of-Living Indexes, *American Economic Review* **70** (1980), 64–77.
[5] W.E. Diewert, The early history of price index research, in: *Essays in Index Number Theory*, (Vol. 1), W.E. Diewert and A.O. Nakamura, eds, Amsterdam, 1993, Chapter 2, pp. 33–71.
[6] European Commission, *Compendium of HICP Reference Documents*, 2/2001/B/5, Luxembourg 2001, Table 9, pp. 60–61, available at http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-AO-01-005/EN/KS-AO-01-005-EN.PDF.
[7] D. Fenwick, The impact of choice of base month and other factors on the relative performance of different formulae used for aggregation of Consumer Price Index data at an elementary aggregate level, in: *proc. Ottawa Group fifth meeting*, 1999, available at www.statice.is/ottawa/fenwick.pdf.
[8] J.S. Greenlees and R. McClelland, Superlative and regression-based consumer price indexes for apparel using U.S. scanner data, *presented at the Conference of the International Association for Research in Income and Wealth*, St. Gallen, 2010, available at www.iariw.org/papers/2010/8aMcClelland.pdf.
[9] International Labour Office et al, *Consumer Price Index Manual, Theory and Practice*, International Labour Office, Geneva, 2004, Chapter 16.
[10] International Labour Office et al, *Consumer Price Index Manual, Theory and Practice*, International Labour Office, Geneva, 2004, Chapters 17,18.
[11] S. Nickell, Two Current Monetary Policy Issues, *Bank of England Quarterly Bulletin*, Winter 2003, pp. 504–517.
[12] OECD Statistics, *Main Economic Indicators*, Consumer Price Indices: Data and Methods.
[13] Office for National Statistics, *CPI and RPI: Increased impact of the formula effect in 2010*, Information Note, 2011, available at http://www.statistics.gov.uk/downloads/theme_economy/info-note-cpiandrpi-impact-formula-effect2010.pdf.
[14] Office for National Statistics, *The new inflation target: The statistical perspective*, London, 2003, available at http://www.statistics.gov.uk/downloads/theme_economy/New_inflation_target_031210.pdf.
[15] M. Sabag and Y. Finkel, The Israeli Consumer Price Index: Alternative computing methods for an 'Item Price Index', *Statistical Journal of the United Nations ECE* **11** (1994), 95–118.
[16] M.D. Shapiro and D.W. Wilcox, Alternative strategies for aggregating prices in the CPI, *Federal Reserve Bank of St Louis Review* **79** (1997), 113–125.
[17] I. Song and P.K. Chintagunta, Measuring cross-category price effects with aggregate store data, *Management Science* **52** (2006), 1594–1609.
[18] United Nations Economic Commission for Europe et al, *Practical Guide to Producing Consumer Price Indices*, United Nations, New York and Geneva, 2009, Chapters 10 and 13.
[19] U.S. Senate, Committee on Finance, Final Report of the Advisory Commission to Study the Consumer Price Index. Print 104-72, 104 Cong., 2 sess. Washington, D.C., Government Printing Office, 1996.

Annex: Price indices in an economy with two goods and a representative consumer and producer

Assume a Cobb-Douglas utility function: $U = x^{\alpha}y^{1-\alpha}$.

By taking a total differential and setting $dU = 0$ we get a family of indifference curves:

$$\frac{dy}{dx} = -\frac{\propto}{1-\propto}\frac{y}{x}$$

When the consumer is in equilibrium, the slope of the indifference curve = ratio of relative prices:

$$-\frac{P_x}{P_y} = \frac{dy}{dx} = -\frac{\propto}{1-\propto}\frac{y}{x}$$

which can be written: $\frac{P_x x}{P_y y} = \frac{\alpha}{1-\alpha}$.

Assume an elliptical production possibility frontier: $\theta x^2 + y^2 = k^2$.

Differentiating, the slope at any point on the production possibility frontier is given by: $\frac{dy}{dx} = -\theta\frac{x}{y}$ which, at the supplier's equilibrium, $= -\frac{P_x}{P_y}$.

In a market equilibrium between supply and demand, the price ratio faced by consumer and supplier is equal:

$$\therefore \theta\frac{x}{y} = \frac{\alpha}{1-\alpha}\frac{y}{x}$$

$$\therefore x = y\sqrt{\frac{\alpha}{1-\alpha}}$$

Substituting this result in the equation for the production possibility frontier gives, in equilibrium:

$$\theta y^2\frac{\alpha}{1-\alpha} + y^2 = k^2$$

$$\therefore \quad y = \sqrt{1-\propto}k$$

$$\text{and } x = \sqrt{\frac{\propto}{\theta}}k$$

Now introduce the consumer's income constraint: she allocates her income m between the two goods:

$$P_x x + P_y y = m$$

Using the unit elasticity expenditure property of the Cobb-Douglas, established above:

$$P_x x = \alpha m$$

$$\text{and } P_y y = (1-\alpha)m$$

Substituting in the values already calculated for x and y:

$$P_x = \sqrt{\alpha\theta}\frac{m}{k}$$

$$P_y = \sqrt{1-\alpha}\frac{m}{k}$$

Now, introduce a supply shock, altering the production possibility frontier to: $\varphi x^2 + y^2 = k^2$.

Table 4
(Annex): Price and Quantities before and after a Supply Shock

	P_x	x	P_y	y
period 0	$\sqrt{\alpha\theta}\frac{m}{k}$	$\sqrt{\frac{\alpha}{\theta}}k$	$\sqrt{1-\alpha}\frac{m}{k}$	$\sqrt{1-\alpha}k$
period 1	$\sqrt{\alpha\varphi}\frac{m}{k}$	$\sqrt{\frac{\alpha}{\varphi}}k$	$\sqrt{1-\alpha}\frac{m}{k}$	$\sqrt{1-\alpha}k$

The prices and quantities in the initial and subsequent periods are set out in Table 4.

A weighted Laspeyres price index is calculated as

$$\frac{P_{x,1}}{P_{x,0}}\left(\frac{P_{x,0}x_0}{m}\right) + \frac{P_{y,1}}{P_{y,0}}\left(\frac{P_{y,0}y_0}{m}\right)$$
$$= \frac{P_{x,1}}{P_{x,0}}\alpha + \frac{P_{y,1}}{P_{y,0}}(1-\alpha)$$
$$= \sqrt{\frac{\varphi}{\theta}}\alpha + (1-\alpha)$$

The weighted Paasche index is

$$\frac{P_{x,1}x_1 + P_{y,1}y_1}{P_{x,0}x_1 + P_{y,0}y_1}$$
$$= \frac{1}{\sqrt{\frac{\theta}{\varphi}}\alpha + (1-\alpha)}$$

The weighted Geometric index is calculated as

$$\left(\frac{P_{x,1}}{P_{x,0}}\right)^{\alpha}\left(\frac{P_{y,1}}{P_{y,0}}\right)^{1-\alpha} = \left(\frac{\varphi}{\theta}\right)^{\frac{\alpha}{2}}$$

Fisher's Ideal index is the geometric mean of the Laspeyres and Paasche indices:

$$\sqrt{\frac{\varphi\alpha + (1-\alpha)\sqrt{\theta\varphi}}{\theta\alpha + (1-\alpha)\sqrt{\theta\varphi}}}$$

A straightforward comparison shows that the Laspeyres index will always be greater than the Paasche index (equal when $\theta = \varphi$), with the Fisher's Ideal index between them.

The Geometric Mean index will equal the Laspeyres index in the trivial case where $\theta = \varphi$. Otherwise, it will always be below it. To see this, consider both indices as a function of $A = \sqrt{\frac{\varphi}{\theta}}$. The derivative of the Laspeyres index with respect to A is a constant, α. Likewise, the derivative of the Geometric Mean index is a diminishing function of A. The two derivatives are equal only where $\theta = \varphi$, when the values of the two indices also coincide. Therefore, for all values of A, the value of the Geometric Mean index is below that of the Laspeyres index.

The Geometric Mean index will be equal to the Fisher's Ideal index when $\alpha = 1/2$ (i.e. equal expen-

Table 5
(Annex): Prices and Quantities before and after a Demand Shock

	P_x	x	P_y	y
period 0	$\sqrt{\alpha\theta}\frac{m}{k}$	$\sqrt{\frac{\alpha}{\theta}}k$	$\sqrt{1-\alpha}\frac{m}{k}$	$\sqrt{1-\alpha}k$
period 1	$\sqrt{\beta\theta}\frac{m}{k}$	$\sqrt{\frac{\beta}{\theta}}k$	$\sqrt{1-\beta}\frac{m}{k}$	$\sqrt{1-\beta}k$

diture on the two goods), but will diverge from it as α goes towards 0 or 1.

Now, instead, introduce a demand shift, with the Utility function changing to $U = x^{\beta}y^{1-\beta}$, with unchanged income, m. The initial and subsequent period prices and quantities will be as in Table 5.

The Laspeyres price index will be:

$$\sqrt{\frac{\beta}{\alpha}}\,\alpha + \sqrt{\frac{1-\beta}{1-\alpha}}\,(1-\alpha)$$
$$= \sqrt{\alpha\beta} + \sqrt{(1-\alpha)\,(1-\beta)}$$

The Paasche index is:

$$\frac{m}{\sqrt{\alpha\theta}\frac{m}{k}\sqrt{\frac{\beta}{\theta}}k + \sqrt{1-\alpha}\frac{m}{k}\sqrt{1-\beta}k}$$
$$= \frac{1}{\sqrt{\alpha\beta} + \sqrt{(1-\alpha)\,(1-\beta)}}$$

Fisher's Ideal index $= 1$.

The Geometric Mean index $= \left(\sqrt{\frac{\beta}{\alpha}}\right)^{\alpha}\left(\sqrt{\frac{1-\beta}{1-\alpha}}\right)^{1-\alpha}$.

Given the properties of the Cobb-Douglas utility function with given expenditure and the unchanged elliptical production possibility frontier, the second period prices will settle such that the Fisher's Ideal price index will be unity for all α, β. The Laspeyres index will always be less than unity (given $0 < \alpha, \beta < 1$), except that it will equal unity in the trivial case when $\alpha = \beta$. Similarly, the Paasche index will always be more than unity.

The Geometric Mean index will always be below the Laspeyres index. To see this, set $\sqrt{\frac{\beta}{\alpha}} = A$ and $\sqrt{\frac{1-\beta}{1-\alpha}} = B$. Then both indices will, for a given α, be functions of A and B. The equation for the Lasperyres' index represents a plane in the space, L, A, B. That for the Geometric Mean index represents a surface concave from below (towards the A and B axes). It will have the same slope as the Laspeyres plane when $A = B = 1$ (implying $\alpha = \beta$), when the two indices will also have the same value. Therefore, for all other values of A, B (i.e. for all values of β given α), the Geometric index lies below the Laspeyres index.