# 4

# What Is a Price Index?

Having seen how some of the early price indices were constructed in response to a specific need, in this chapter we discuss what a price index is and how it can be put together, using a relatively simple data set to highlight many of the issues. In order to help clarify the issues we are talking about in determining an appropriate measurement of price change, we first attempt to clarify the language that we will use to talk about measuring a change in the price level. Then we consider the potential inputs to such a process and discuss how such inputs might be used to produce meaningful estimates of the change in the price level, some of the methods for which we have already met in Chap. 3.

## 4.1 Defining a Price Index, Inflation and Index Numbers

In the course of this book and the practice of measuring changes in the general price level, we use a precise terminology. Before we define the mechanisms for producing numerical estimates of inflation, it is worth



R. O'Neill et al., *Inflation*, DOI 10.1007/978-3-319-64125-6_4

clarifying the way in which we talk about this process so that we start from a common platform of understanding.

We will often refer to **Index Numbers** as the subject, hence the capitalisation of the term. This covers the entire area of study of the design, properties, applications and potential uses of statistical tools which are designed to produce a single number to summarise the movement in one variable, constructed from many observations of other variables, between two or more different states of the world. These states could be spatially or temporally defined, as most price indices are. In this book we have focused explicitly on the use of Index Numbers in the pursuit of measuring changes in the general price level in the UK between different time periods, with each time period representing a distinct state of the world. We might otherwise have chosen to focus on using Index Numbers tools to measure the differences in contemporaneous price levels across a set of different countries (see Chap. 14), or we might have employed the techniques to indicate changes in industrial production across time periods in the UK. These represent just a few uses of the broad set of statistical tools which we have labelled as the domain of Index Numbers, and emphasises that this book focuses on a small subset of this broader subject area. We aim to help to make the study of Index Numbers in the context of UK inflation measurement more accessible.

Having defined the area of study in which our attention is focused on as **Index Numbers**, we will focus specifically on the different estimators which have been designed to measure the change in the price level through time. We collectively call this class of estimators **index number formulae** and they represent the way in which inputs can be combined in order to produce single, summary estimates of the price level, and percentage changes in these estimates across time then form our measure of inflation.

As we will see, the **index number formulae** which we employ will produce a different number to summarise our variable of interest (the price level) in each state of the world (time period). We will refer to the set of numbers which results from the computations from a single estimator as a set of **index numbers**. Note that this set of numbers is indicated by the use of lower case letters in "index numbers". This is in contrast to the use of capitalisation in the naming of the subject Index

Numbers. This is similar to the definition of the study of Statistics (the subject) and a set of statistics (e.g. average scores of students on a test). Although this may seem confusing at this point, hopefully our use of the two different terms will become clear as we make further headway in our consideration of inflation measurement in the UK context.

A set of index numbers is scale free, so it is usually scaled to be set equal to some value in a given state of the world, so for example in a set of index numbers to measure inflation one period is usually defined as having the value of 100. In this case, we will refer to this state of the world as the **reference period** for the index, as our states of the world will be exclusively time-based in this book. At the same time, we will label the arbitrary (and relatively unimportant) value to which the index is set in the **reference period** as the **reference value** of the index. The reference period is the one with which it will be most common for us to make direct comparisons and this can be facilitated by the setting of a sensible reference value, which is why a value of 100 is often used.

When constructing index numbers it is necessary to compare some states of the world with an initial or **base period**. This base period is the one which we compare our observations of price or quantity to. In many cases, this will be the earliest period for which an index is constructed, but need not be so. In more complex, long-running series, the **base period** is often updated on a regular basis. It is possible for an index to have different periods for base and reference periods, and the distinction between them will become more clear as we describe the uses of Index Numbers. In Chap. 2 we considered the nature and definition of **inflation**, which we concluded was an increase in the general level of prices. We will further abstract from this idea and in this chapter, the various index number formulae will produce a series of index numbers summarising the price levels, and the percentage change in these will define inflation as measured with that formula. Alternatively, we will consider **deflation** as the percentage decrease in our price level through time, as represented by a decrease in the value of an index number in the series when compared with some period in the past. It should be noted that elsewhere we will consider a further use of the term deflation, however it should be clear from the usage when we are using the term to describe a fall in the price level.

## 4.2 The Potential Inputs to Index Numbers Calculation

There are two main inputs to the set of index numbers formulae we will consider in this chapter: prices and quantities. While we might, in practice, make use of a further set of information relating to expenditure shares, much of the thinking around Index Numbers begins with a consideration of how prices and quantities might be used to measure a change in the general level of prices.

Consider an economy in which there are $n$ products[1] and that each of these $n$ products is available at a single price, $p_{it}$, in a given time period $t$, $p_{it}$ where $i = 1, 2, \ldots, n$. This in itself is not as simple an assumption as it might seem, as a number of goods have prices which differ according to the people buying them, for example many cinemas charge different prices to children, adults, students and pensioners. We will therefore use the simplifying assumption that each good is sold at a single price in a single time period. The $n \times 1$ column vector of prices is therefore represented by $\mathbf{P}_t = (p_{1t}, p_{2t}, \ldots, p_{nt})'$ in the rest of this discussion.

The other main input in our estimation of index numbers is the amount of a given product which is consumed at the price we defined above. We represent the quantity of a product consumed in a given time period, $t$, as $q_{it}$ where again $i = 1, 2, \ldots, n$. We can also represent this as a column vector $\mathbf{Q}_t = (q_{1t}, q_{2t}, \ldots, q_{nt})'$.

It is often difficult for those compiling inflation statistics to directly observe the quantities of products being purchased, however it is much more practical to be able to observe the amount of money which is spent on a given good in a period of time. For that reason, in our discussion of index numbers formulae, we will typically show the equation in two forms, one which specifies the estimator in terms of prices and quantities and another in terms of **expenditure shares**. The expenditure

[1]The terminology for individual items tends to vary across disciplines. Statisticians will be more likely to refer to products, while economists will more commonly refer to goods. We use the two terms interchangeably henceforth.

share of product $i$ in period $t$ represents the proportion of total expenditure which is related to purchases of the $i^{th}$ product. Hence, if we can denote total expenditure as the sum of the products of prices and quantities across the $n$ goods, $\mathbf{P}_t'\mathbf{Q}_t$ then the expenditure share of the $i^{th}$ good can be represented as:

$$w_{it} = \frac{p_{it}q_{it}}{\mathbf{P}_t'\mathbf{Q}_t}$$

and again we can create a column vector of these for the $n$ goods, $\mathbf{w}_t = (w_{1t}, w_{2t}, \ldots, w_{nt})'$.

As we will see, there are some specialised index numbers which require additional inputs, usually parameters governing economic behaviour, and we will discuss these for individual formulae as we come across them. However, we are now well equipped with the building blocks of our index numbers and can begin to consider how they might be combined in order to tell us something about the general level of prices and how it changes over time.

## 4.3 Some Popular Index Numbers Formulae

This section introduces a small selection of the index number formulae which have been suggested for the construction of an index number series to measure the change in the general level of prices. We begin by considering a slightly different question, the Index Number problem, which motivated many of the first attempts at identifying an appropriate estimator for the general level of prices.

### 4.3.1 The Index Number problem

The Index Number problem begins by looking at the change in the overall level of consumption between two time periods. We will label the first of these periods as time 0 and the later one as $t$. In this case, we can create a **value index** which measures the change in the amount spent (the sum of price multiplied by quantity for each individual good) on a set of $n$ goods in period $t$ compared to the base period. We will denote this value index as $V^{0,t}$ where

$$V^{0,t} = \frac{\mathbf{P}_t^{'}\mathbf{Q}_t}{\mathbf{P}_0^{'}\mathbf{Q}_0}$$

and note that in this case the base value of the index is 1, as $V^{0,0} = 1$ by definition. In this case, if $V^{0,t} > 1$ then the amount of total spending has increased in period $t$ compared to period 0, while if $V^{0,t} < 1$ then total spending is less in period $t$ than it was in period 0.

The Index Number problem was introduced by many early academics working in the area, summarised in Frisch (1936), and has driven much discussion in Index Numbers since. The crux of the problem is that there are only two things which can have changed between the two time periods. Either quantities can change or prices can change and it is thought that the change in total spending should therefore lend itself to being decomposed into a measure of changes in the level of prices and a measure of change in the level of consumption. That is, if we label the index of changing prices as $I_P^{0,t}$ and the index of changes in the quantity consumed as $I_Q^{0,t}$, then it should be possible to specify a price index which, alongside an appropriate quantity index, satisfies the property:

$$V^{0,t} = I_P^{0,t} \times I_Q^{0,t}$$

where the quantity index corresponding to a given price index can be inferred from the above identity.

We mention the Index Number problem here as it motivated much of the early development of weighted indices and we wish to highlight how such indices fit into the framework of this overarching problem below. Indeed, the Index Number problem remains relevant as price indices are used to deflate output measures so that changes in real economic activity can be measured in a currency with a consistent purchasing power. We will return to this application of index number series in the final section of this chapter.

### 4.3.2 Unweighted Index Numbers

Index number formulae can generally be split into two categories[2]: those that are weighted using quantity information and those that are not. In splitting the formulae in this way, we are stepping outside of the historical flow, which presents the main formulae in the order they were suggested.

We begin by considering the category of formulae for which we do not need quantities in order to be able to calculate the index numbers. It is rare for a measure of inflation to be constructed as a purely unweighted index, however it is common for such formulae to be used at the lowest level of a more complex index structure as we will see in Chap. 10 when we discuss elementary aggregates.

The first equation we will consider is the Carli index, which is the arithmetic mean of the **price relatives** for the $n$ goods under consideration. Denoting this index numbers formula as $I^{0,t}_{Carli(P)}$ where:

$$I^{0,t}_{Carli(P)} = \frac{1}{n}(\mathbf{R}_\mathbf{t}^{'}\mathbf{1}_n)$$

where $\mathbf{R}_t = \left(\frac{p_{1t}}{p_{10}}, \frac{p_{2t}}{p_{20}}, \ldots, \frac{p_{nt}}{p_{n0}}\right) = (R_{1t}, R_{2t}, \ldots, R_{nt})$ and $\mathbf{1}_n$ denotes an $n \times 1$ column vector in which every element is equal to 1.

An alternative to the Carli index is the Dutot index, which rather than taking the averages of the ratios of prices takes the ratio of the averages of prices. Hence,

$$I^{0,t}_{Dutot(P)} = \frac{\frac{1}{n}(\mathbf{P}_t^{'}\mathbf{1}_n)}{\frac{1}{n}(\mathbf{P}_0^{'}\mathbf{1}_n)} = \frac{\mathbf{P}_t^{'}\mathbf{1}_n}{\mathbf{P}_0^{'}\mathbf{1}_n}$$

the final statement is valid only where $n$, the number of goods, is constant between the two time periods.

[2]There are many other ways we might choose to classify index number formulae, this approach is chosen only for presentational purposes.

Hereafter, the unweighted indices become increasingly less obvious. The most important such index was proposed by W.S. Jevons in Jevons (1863), in his pamphlet on the changes in the value of gold. Jevons proposed that the change in the price level should be measured by the geometric mean of price relatives in the sample.

$$I^{0,t}_{Jevons(P)} = \left(\Pi^{n}_{i=1} R_{it}\right)^{(1/n)}$$

Our list of potential unweighted indices does not end here, and we might consider using the harmonic mean of price relatives as an alternative measure. In this case, the index formula is defined as:

$$I^{0,t}_{Harmonic(P)} = \frac{1}{\frac{1}{n}((1/\mathbf{R}_t)'\mathbf{1}_n)} = \frac{n}{(1/\mathbf{R}_t)'\mathbf{1}_n}$$

The Carruthers-Sellwood-Ward-Dalen (CSWD) index, is named after the combination of authors who have supported its use over the years. It produces index numbers which are the geometric mean of the corresponding Carli and Harmonic indices:

$$I^{0,t}_{CSWD(P)} = \sqrt{I^{0,t}_{Carli(P)} \times I^{0,t}_{Harmonic(P)}}$$

As far as we can discover the formula itself was first proposed in Coggeshall (1886). It is thought that the CSWD is an unweighted approximation to the Fisher index, which we will meet when we consider weighted index numbers formulae below. The properties of the Fisher index will be discussed in later chapters, however for the minute it is sufficient to note that this formula is one which many Index Numbers experts would be likely to include among their preferred formulae for inflation measurement, if sufficient data were available, hence the importance of the CSWD index, as a potentially unweighted version of the Fisher index. The CSWD formula seems like a much less natural estimator of an unweighted change in the price level, however as we will see several times in this book, as relates to Index Numbers, it is rarely the most obvious solution to a problem which turns out to be best in the eyes of the Index Numbers community.

Mehrhoff (2010) considers which unweighted index numbers formulae are equivalent to weighted indices, and doing this makes use of the concept of a generalised mean of price relatives. For a parameter value $r$, we can define the generalised mean of price relatives as:

$$I^{0,t}_{Gen.Mean(r)(P)} = \sqrt[r]{\frac{1}{n}(\mathbf{R}^r_t)'\mathbf{1}_n}$$

where $\mathbf{R}^r_t = (R^r_{1t}, R^r_{2t}, \ldots, R^r_{nt})'$. As Mehrhoff (2010) shows, this index number formula includes several of the unweighted formulae we have already seen as special cases. As the value of $r$ is changed, so does the nature of the average of price relatives we are taking. If we set $r = 1$, then the generalised mean reduces to the Carli index as described above. Similarly, as $r \to 0$ then $I^{0,t}_{Gen.Mean(r)(P)} \to I^{0,t}_{Jevons(P)}$ and if $r = -1$ then $I^{0,t}_{Gen.Mean(r)(P)} = I^{0,t}_{Harmonic(P)}$.[3] We then technically have an infinite number of ways to combine the price relatives to produce a series of index numbers as $r$ can take on any real value. In practice, however this will not have much practical appeal, particularly as it can be shown that the upper and lower limits of the generalised mean are the maximum and minimum values from the column vector of price relatives, hence $min(\mathbf{R}_t) \leq I^{0,t}_{Gen.Mean(r)(P)} \leq max(\mathbf{R}_t)$. Mehrhoff (2010) goes on to ask an interesting question: What value of $r$ allows an unweighted index to replicate the results of a given form of weighted index?

### 4.3.3 Weighted Index Numbers

We now turn our attention to the second collection of methods for producing estimates of the price level at a given point in time. All of these formulae use information other than prices to estimate the price level.

As we saw in Chap. 3, Etienne Laspeyres (1864) proposed that the quantities from the base period of the comparison could be used in

[3] Mehrhoff (2010) also considers $r = 2$ which produces a quadratic mean and $r = -2$ which produces a reciprocal quadratic, however we have never seen either of these formulae applied as price indices so do not include them in our discussion.

order to provide some useful weighting information. In essence, his famous index numbers formula measured the factor by which we would need to multiply income to ensure that a consumer could buy exactly the same goods at time $t$ in exactly the same numbers as was observed at time 0. His formula can be written as:

$$I^{0,t}_{Laspeyres(P)} = \frac{\mathbf{P}_t^{'}\mathbf{Q}_0}{\mathbf{P}_0^{'}\mathbf{Q}_0}$$

Laspeyres despaired that this might not be the most practical of index numbers formulae as it called for the dual collection of prices and quantities. It was soon shown by Irving Fisher, in his 1922 study of Index Numbers, that there is an alternative way of writing the Laspeyres formula which does not require the explicit use of quantities and uses expenditure weights instead:

$$\begin{aligned} I^{0,t}_{Laspeyres(P)} &= \frac{\mathbf{P}_t^{'}\mathbf{Q}_0}{\mathbf{P}_0^{'}\mathbf{Q}_0} = \frac{\sum_{i=1}^{n} p_{it}q_{i0}}{\sum_{i=1}^{n} p_{i0}q_{i0}} = \sum_{i=1}^{n} \frac{p_{it}q_{i0}}{\sum_{i=1}^{n} p_{i0}q_{i0}} \frac{p_{i0}}{p_{i0}} \\ &= \sum_{i=1}^{n} \frac{p_{i0}q_{i0}}{\sum_{i=1}^{n} p_{i0}q_{i0}} \frac{p_{it}}{p_{i0}} = \mathbf{W}_0^{'}\mathbf{R}_t \end{aligned}$$

where $\mathbf{W}_t = \left( \frac{p_{1t}q_{1t}}{\sum_{i=1}^{n} p_{it}q_{it}}, \frac{p_{2t}q_{2t}}{\sum_{i=1}^{n} p_{it}q_{it}}, \ldots, \frac{p_{nt}q_{nt}}{\sum_{i=1}^{n} p_{it}q_{it}} \right) = (w_{1t}, w_{2t}, \ldots, w_{nt})$ is the column vector of period $t$ expenditure shares. This allows for the estimation of a Laspeyres price index series if all we have access to price quotes from a given time period and the expenditure weights from the base period of the index.

In some cases, it is not possible to obtain expenditure weights without a significant delay, relative to the time scale demanded for release of inflation estimates. As a result, it may be necessary to use quantities from some time period $s < 0$ before the base period for prices. In this case, we are comparing the cost of the basket of goods from period $s$, obtained at period 0 prices with the cost of obtaining the same basket of goods and services at time period $t$.[4] This is the Lowe index, which we saw in Chap. 3, preceded the Laspeyres index, and can be written as:

$$I_{Lowe(P)}^{0,t} = \frac{\mathbf{P}_t^{'}\mathbf{Q}_s}{\mathbf{P}_0^{'}\mathbf{Q}_s}$$

We will deal with how this formula can be operationalised to produce index numbers at length, as it is the form which is most commonly adopted by NSIs in their production of official price statistics for a number of reasons. We could also write it as a weighted combination of price relatives:

$$I_{Lowe(P)}^{0,t} = \frac{\mathbf{P}_t^{'}\mathbf{Q}_s}{\mathbf{P}_0^{'}\mathbf{Q}_s} = \mathbf{W}_{s,0}^{'}\mathbf{R}_t$$

where $\mathbf{W}_{s,0} = \left(\frac{p_{10}q_{1s}}{\sum_{i=1}^{n} p_{i0}q_{is}}, \frac{p_{20}q_{2s}}{\sum_{i=1}^{n} p_{i0}q_{is}}, \ldots, \frac{p_{n0}q_{ns}}{\sum_{i=1}^{n} p_{i0}q_{is}}\right)$.

The Lowe index assumes that we are able to identify the quantities from period $s$, however we might also consider what happens if we are only able to obtain expenditures in period $s$. In this case, we could use the formula proposed in Young (1812) which is:

$$I_{Young(P)}^{0,t} = \mathbf{W}_s^{'}\mathbf{R}_t$$

which Arthur Young used in his consideration of the changing value of money in England for agricultural products.

Although it is often thought that the Laspeyres index is a fairly intuitive way of presenting a price index, there are some clear alternatives to this way of doing things which have been suggested and have stood the test of time in the Index Numbers literature. The most famous alternative to the Laspeyres formula is the Paasche index formula, as presented by Herman Paasche (1874). In this formula, we take the quantities not

[4]In many practical applications of this formula, it is normal for those producing price indices to attempt to minimise the distance in time between period $s$ and period 0 so that the basket of goods is as relevant as possible.

from period 0, but from period $t$, as there is no reason why the quantities purchased in this period should not be treated with as much emphasis as those from period 0. Hence the Paasche formula is:

$$I^{0,t}_{Paasche(P)} = \frac{\mathbf{P}^{'}_{t}\mathbf{Q}_t}{\mathbf{P}^{'}_{0}\mathbf{Q}_t}$$

which is very similar in structure to the Laspeyres formula. It is possible to write the Paasche index as a weighted combination of price relatives in a style similar to that of the Laspeyres index above:

$$\begin{aligned} I^{0,t}_{Paasche(P)} &= \frac{\mathbf{P}^{'}_{t}\mathbf{Q}_t}{\mathbf{P}^{'}_{0}\mathbf{Q}_t} = \frac{\sum_{i=1}^{n} p_{it}q_{it}}{\sum\limits_{i=1}^{n} p_{i0}q_{it}} = \sum_{i=1}^{n} \frac{\sum_{i=1}^{n} p_{it}q_{it}}{p_{i0}q_{it}} \frac{p_{it}}{p_{it}} \\ &= \sum_{i=1}^{n} \frac{\sum_{i=1}^{n} p_{it}q_{it}}{p_{it}q_{it}} \frac{p_{it}}{p_{i0}} = (\mathbf{W}^{'}_{t}(1/\mathbf{R}_t))^{-1} \end{aligned}$$

which means that the Paasche index is a current period weighted harmonic mean of the price relatives between the two time periods.

We have seen a few index numbers formulae which use weighting information and it is possible to alter these formulae to produce new formulae. For example, starting with the expenditure weighted version of the Lasperyes formula, there is no reason that the arithmetic weights need to come from the base period. We can replace these weights with those from period $t$, which leads us to the Palgrave price index:

$$I^{0,t}_{Palgrave(P)} = \mathbf{W}^{'}_{t}\mathbf{R}_t$$

which is a period $t$ expenditure share weighted arithmetic mean of price relatives between period 0 and period $t$.

In a similar fashion, we can ask why the weights in the harmonic version of the Paasche formula, must come from period $t$. It is a straightforward exercise to replace these weights with those from period 0 in order to obtain a further weighted price index, which we call the harmonic-Laspyeres index:

$$I^{0,t}_{Harmonic-Laspeyres(P)} = (\mathbf{W}_0^{'}(\mathbf{1}_n/\mathbf{R}_t))^{-1}$$

which will produce another alternative set of index numbers.

As we have seen already, the differences between the geometric and arithmetic mean have given rise to a large number of different index number formulae, and it is also possible to identify geometric versions of the Laspeyres and Paasche indices, in which the price relatives are first raised to the power of their expenditure share in a given period and then multiplied together to give either a geometric Laspeyres, when we use period 0 expenditure shares, or a geometric Paasche, when we use the period $t$ expenditure shares. This further expands the number of formulae available for combining prices and quantities to measure changes in the price level.

We could follow this path further along several other dimensions, for example we could take a generalised mean of the various combinations of weights and price relatives, which would then produce a huge number of new indices, many of which would be difficult to interpret in an economic sense. We therefore restrict the rest of our discussion to alternative formulae which have made an appearance in the existing Index Numbers literature in order to ease the potential burden we have in considering which formula to employ when considering the estimation of inflation.

### 4.3.4 Symmetrically Weighted Index Number Formulae

Having looked at index number formulae which try to isolate which set of quantities, or expenditure weights we should be using, we can now consider a class of indices which do not require such a choice but in some sense try to treat the weightings from the two periods as symmetric. Later on, we will say a lot more about the properties of the index number formulae which we present under this heading, however it is worth noting that they are of particular interest in the field of Index Numbers.

By far, the most famous index number formula using both sets of weights is the Fisher index, discussed at length in Fisher (1922) by the famous economist Irving Fisher. The formula for this index takes the geometric mean of the Laspeyres and Paasche indices, hence:

$$I^{0,t}_{Fisher(P)} = \sqrt{I^{0,t}_{Laspeyres(P)} \times I^{0,t}_{Paasche(P)}}$$

where we can see that the value of this index number series at time period must be somewhere between the values of the Lasperyes and Paasche indices. It is notable that this is the geometric mean of a weighted arithmetic mean and a weighted harmonic mean of price relatives, hence we can see the relationship between this formula and an unweighted version of it that we have seen above in the CSWD index, introduced in the unweighted collection of index numbers.

Unsurprisingly given the breadth of choice of index number formulae that we have already encountered in this chapter, there are further symmetrically weighted price index formulae options available. Törnqvist (1936) introduced an index number formula which takes a weighted geometric mean of the price indices, where the weights on individual price relatives are the arithmetic mean of the expenditure shares in the two periods.

$$I^{0,t}_{Tornqvist(P)} = \Pi^{n}_{i=1} R_{it}^{\frac{w_{i0}+w_{it}}{2}}$$

Alternatively, the Marshall-Edgeworth formula takes a weighted arithmetic mean of the price relatives; however, in this case, the weights chosen are the arithmetic means of the expenditure share for each of the goods across the two periods considered by the index:

$$I^{0,t}_{Marshall-Edgeworth(P)} = (\frac{1}{2}(\mathbf{W}_0 + \mathbf{W}_t))'\mathbf{R}_t.$$

Having considered the Marshall-Edgeworth formula, it was not clear to Walsh (1901, 1921) that the weights used should be estimated using an arithmetic mean. Instead, he suggested the weights be produced by the geometric mean of the expenditure shares in the two periods:

$$I^{0,t}_{Walsh(P)} = (\mathbf{W}_{0\times t})'\mathbf{R}_t$$

where $\mathbf{W}_{0\times t} = \left((w_{10}w_{1t})^{(1/2)}, (w_{20}w_{2t})^{(1/2)}, \ldots, (w_{n0}w_{nt})^{(1/2)})\right)$. Drobisch[5] (1871) had earlier suggested what now seems an obvious alternative to the Fisher formula; the arithmetic mean of the Laspeyres and Paasche indices:

$$I^{0,t}_{Drobisch(P)} = \frac{I^{0,t}_{Laspeyres(P)} + I^{0,t}_{Paasche(P)}}{2}$$

which is guaranteed to have higher numbers in its index numbers series than the Fisher index when $t \neq 0$ except in the special case in which $I^{0,t}_{Laspeyres(P)} = I^{0,t}_{Paasche(P)}$.

As we progress further into our study of Index Numbers, we will see that the subject is closely tied to the area of utility optimisation in microeconomics. It is no surprise then that this area of study should also have provided its own version of a price index for consideration alongside other measures. The index proposed in Lloyd (1975) uses as its basis a constant elasticity of substitution (CES) utility function, which contains the parameter $\sigma$, which represents the elasticity of substitution, which determines the rate at which consumers are willing to substitute goods with differential rates of marginal utility. The formula proposed in Lloyd (1975) was:

$$I^{(0,t)}_{Lloyd-Moulton(P)} = \left(\mathbf{W}_0^{'}(\mathbf{R}_t^{1-\sigma})\right)^{\frac{1}{1-\sigma}}$$

which was rediscovered by Moulton and Moses (1997) and has since become known as the Lloyd-Moulton formula. The new formula, which allows us to estimate an index which would ensure consumers have a fixed level of utility, requires the estimation of $\sigma$, which in itself is a complex task and therefore complicates the problem of operationalising

[5]As von Auer (2010) the contribution of Drobisch to the field of Index Numbers goes far beyond the suggestion of this formula, including first proposing the forms of the indices which carry the names of Laspeyres and Paasche and the suggestion of a unit value index.

such an index.[6] We might consider such a formula as a single-weighted index; however, as it attempts to fix utility across multiple periods, we have included it in this section of our introduction to Index Number formulae.

The differences between the various indices are subtle at first glance and the breadth of choice may appear overwhelming. The first question we might ask is whether it really makes a difference which formula we use to measure inflation? The short answer is yes, which we will try to demonstrate with the use of a small numerical example. However, it is possible to see that all of these formulae will produce slightly different sets of index numbers, and therefore slightly different estimates of inflation. In some cases, the difference will be relatively small, for example when the Laspeyres and Paasche formulae produce similar index number series then the Fisher and Drobisch indices will, by definition, produce similar sets of index numbers as well. In many cases, the weighted indices will produce similar sets of index numbers, which in some sense should be considered reassuring—if they were wildly different when using the same inputs, then the debate about which index number formula to use in measuring inflation would be even more lively.

### 4.3.5 Returning to the Index Number Problem

Having been introduced to a multitude of Index Number formulae, we can now return to the original question with which we began this section: do the above index number formulae, and the index numbers they produce, allow us to answer the Index Number problem?

The answer is perhaps less clear than we might have hoped. If we had a value index, then we could indeed divide this by any of the price indices we have considered, and this would give us a value which would theoretically represent the change in the quantity level implied by our price index. In some cases we are able to answer the question more forcefully; for example, if we divide our value index by a Laspeyres price index,

[6]For an example of attempts to estimate $\sigma$ for alcohol products in the UK, see Elliott and O'Neill (2012).

then the resulting quantity index will be a Paasche quantity index[7] as we can show:

$$\frac{V^{0,t}}{I^{0,t}_{Laspeyres(P)}} = \frac{\mathbf{P}'_t\mathbf{Q}_t}{\mathbf{P}'_0\mathbf{Q}_0}\frac{\mathbf{P}'_0\mathbf{Q}_0}{\mathbf{P}'_t\mathbf{Q}_0} = \frac{\mathbf{P}'_t\mathbf{Q}_t}{\mathbf{P}'_t\mathbf{Q}_0} = I^{0,t}_{Paasche(Q)}$$

Hence, by deflating a value index by a Laspeyres price index, we know exactly what we will get. When doing things the other way around, deflating a value index series by a Paasche price index series we obtain a Laspeyres quantity index series as seen below.

$$\frac{V^{0,t}}{I^{0,t}_{Paasche(P)}} = \frac{\mathbf{P}'_t\mathbf{Q}_t}{\mathbf{P}'_0\mathbf{Q}_0}\frac{\mathbf{P}'_0\mathbf{Q}_t}{\mathbf{P}'_t\mathbf{Q}_t} = \frac{\mathbf{P}'_0\mathbf{Q}_t}{\mathbf{P}'_0\mathbf{Q}_0} = I^{0,t}_{Laspeyres(Q)}$$

In a similar fashion, when dividing through our value index by a Fisher index then by definition the resulting quantity index must be a Fisher quantity index, that is

$$\frac{V^{0,t}}{I^{0,t}_{Fisher(P)}} = \sqrt{\frac{V^{0,t}}{I^{0,t}_{Laspeyres(P)}}}\sqrt{\frac{V^{0,t}}{I^{0,t}_{Paasche(P)}}} = \sqrt{I^{0,t}_{Laspeyres(Q)} \times I^{0,t}_{Paasche(Q)}}$$

As we progress to deflation of the value index with more complicated formulae, the resulting quantity indices are less easy to interpret. This may cause problems, as if we cannot clearly state the corresponding quantity index, and make it understandable, we will only be defining the quantity index via the form of the price index we have chosen. Although the form of the quantity index is not always a central consideration when choosing from the many price indices above, it is worth consideration as deflating series from current to constant values will be one of the key uses of the index numbers produced using the various estimators discussed in this chapter.

[7] Due to space restrictions, we do not spend more time discussing the quantity index versions of the above indices.

## 4.4 Differences in the Estimation of Inflation

Having seen so many estimators of index number series, it may be useful to see how these formulae might produce estimates of inflation using a small data series in which we can have most of the data we require. In this section we will therefore use an artificial data set to create series of index numbers using all of the formulae discussed above, with the exception of the Lloyd-Moulton index as this would require us to specify the utility function of consumers.

In Table 4.1 we provide the detail on the prices and quantities of 20 goods, as consumed by a group of people over a given time period. We observe quantities and prices over 10 time periods (labelled 0 to 9 so that when we use the first as the base period it is consistent with our notation). We therefore have the data required to estimate many of the index numbers series for each of the formulae we have considered in this chapter.

In Table 4.2 for each of the considered Index Number formulae we report the estimates of inflation (the percentage change in the index number series) compared to the base period. We choose the earliest period 0 as the base period, although we could easily re-base our estimations to another period, say period 5, which would change our estimated measures of price and quantity change. In order to see what this implies regarding quantity changes in the period, we also report the percentage change in the quantity index implied by the calculated price index. This means we have a number of estimates of inflation and of changes in the quantity index and below we will discuss some of the significant differences.

There are considerable differences in the results for different indices, although it is notable that all of the symmetrically weighted indices are similar. The differences between the final estimates of inflation are much larger for the unweighted indices, which therefore affects the corresponding quantity indices. Neither of these results is unexpected as we will see as we delve further into considerations of the nature of the indices we have considered. It is also clear that unweighted versions of indices do not do a very good job of approximating weighted indices, for example the CSWD is a poor approximation of the Fisher index, the harmonic mean is a poor approximation of the Paaasche and the Carli performs badly in replicating the results of the Laspeyres index.

**Table 4.1** Prices and quantities for the 20 items in our basket over ten periods

| Time | 0 | | 1 | | 2 | | 3 | | 4 | | 5 | | 6 | | 7 | | 8 | | 9 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Good | P | Q | P | Q | P | Q | P | Q | P | Q | P | Q | P | Q | P | Q | P | Q | P | Q |
| 1 | 2.5 | 27 | 2.81 | 28 | 2.9 | 28 | 3.04 | 30 | 3.09 | 28 | 3.46 | 28 | 3.66 | 29 | 3.48 | 31 | 3.61 | 34 | 3.59 | 31 |
| 2 | 6.2 | 6 | 6.38 | 6 | 6.25 | 6 | 6.81 | 6 | 6.52 | 6 | 6.62 | 7 | 6.85 | 7 | 6.63 | 7 | 7.39 | 7 | 7.94 | 7 |
| 3 | 7.4 | 8 | 8 | 8 | 7.77 | 7 | 8.59 | 7 | 8.55 | 6 | 8.74 | 5 | 9.84 | 5 | 10.85 | 5 | 10.45 | 5 | 10.81 | 5 |
| 4 | 2.5 | 29 | 2.56 | 27 | 2.79 | 25 | 3.18 | 23 | 3.51 | 25 | 3.69 | 27 | 4.14 | 28 | 4.46 | 26 | 4.77 | 27 | 5.39 | 29 |
| 5 | 11.3 | 28 | 12.76 | 30 | 13.37 | 28 | 13.83 | 28 | 14.81 | 29 | 14.85 | 29 | 15 | 28 | 17.12 | 27 | 16.91 | 26 | 17.75 | 27 |
| 6 | 11.5 | 12 | 11.3 | 11 | 11.89 | 11 | 12.35 | 12 | 13.68 | 11 | 13.11 | 10 | 13.84 | 9 | 14.13 | 9 | 13.43 | 9 | 15.38 | 9 |
| 7 | 9.7 | 12 | 9.45 | 12 | 9.51 | 12 | 10.69 | 12 | 10.84 | 12 | 12.34 | 12 | 11.9 | 12 | 12.06 | 13 | 11.62 | 13 | 12.52 | 13 |
| 8 | 10.9 | 12 | 11.04 | 12 | 11.15 | 13 | 11.25 | 12 | 12.61 | 13 | 12.73 | 13 | 13.05 | 12 | 14.6 | 13 | 13.97 | 12 | 15.74 | 12 |
| 9 | 7.1 | 11 | 7.2 | 10 | 6.93 | 10 | 6.84 | 10 | 7.47 | 10 | 7.73 | 9 | 8.38 | 10 | 8.99 | 11 | 10.06 | 12 | 10.35 | 13 |
| 10 | 9.1 | 12 | 8.89 | 11 | 8.68 | 10 | 8.75 | 10 | 9.04 | 9 | 8.95 | 10 | 10.16 | 9 | 11.14 | 8 | 11.79 | 8 | 13.47 | 8 |
| 11 | 11.4 | 28 | 12.62 | 30 | 12.54 | 32 | 13.52 | 33 | 12.87 | 32 | 13.42 | 29 | 13.86 | 28 | 13.4 | 30 | 14.41 | 33 | 14.46 | 32 |
| 12 | 5 | 9 | 5.1 | 8 | 5.53 | 7 | 5.34 | 7 | 5.81 | 7 | 6.61 | 7 | 6.44 | 8 | 6.9 | 8 | 7.9 | 8 | 7.62 | 8 |
| 13 | 3.2 | 30 | 3.37 | 32 | 3.33 | 35 | 3.41 | 35 | 3.68 | 35 | 4.04 | 38 | 3.93 | 41 | 4.43 | 43 | 4.55 | 47 | 4.84 | 51 |
| 14 | 7.5 | 26 | 7.8 | 28 | 8.12 | 27 | 8.48 | 28 | 9.17 | 29 | 9.99 | 28 | 11.09 | 26 | 12.7 | 28 | 14.41 | 27 | 15.93 | 29 |
| 15 | 7.8 | 14 | 7.51 | 13 | 8.15 | 13 | 7.93 | 12 | 8.57 | 12 | 8.39 | 12 | 8.69 | 11 | 8.69 | 11 | 8.8 | 11 | 8.71 | 12 |
| 16 | 8.8 | 28 | 10.03 | 30 | 9.93 | 27 | 10.46 | 26 | 11.04 | 28 | 12.01 | 27 | 13.58 | 28 | 14.1 | 29 | 14.42 | 28 | 14.35 | 28 |
| 17 | 8 | 29 | 8.58 | 28 | 8.78 | 28 | 9.65 | 27 | 9.8 | 25 | 10.75 | 25 | 10.7 | 23 | 11.99 | 23 | 12.08 | 21 | 11.97 | 21 |
| 18 | 3 | 7 | 3.3 | 7 | 3.75 | 8 | 4.19 | 7 | 4.77 | 6 | 5.14 | 6 | 5.42 | 6 | 6.11 | 6 | 5.87 | 6 | 6.72 | 7 |
| 19 | 2.2 | 19 | 2.53 | 20 | 2.87 | 20 | 3.2 | 22 | 3.18 | 24 | 3.51 | 25 | 3.95 | 24 | 4.35 | 22 | 4.24 | 21 | 4.21 | 22 |
| 20 | 2.3 | 11 | 2.31 | 11 | 2.4 | 12 | 2.59 | 12 | 2.74 | 11 | 2.76 | 11 | 2.79 | 11 | 3.12 | 12 | 3.59 | 11 | 3.98 | 11 |

**Table 4.2** Price and quantity indices between period 0 and period *t*

| | Time | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 |
|---|---|---|---|---|---|---|---|---|---|---|
| Carli | P | 4.99 | 8.41 | 14.72 | 20.68 | 27.31 | 34.13 | 43.74 | 48.61 | 57.35 |
| | Q | 3.01 | 0.46 | 0.04 | −1.02 | −3.03 | −5.68 | −3.65 | −3.86 | −2.65 |
| Dutot | P | 4.47 | 6.72 | 12.15 | 17.72 | 22.88 | 29.02 | 37.74 | 41.39 | 49.73 |
| | Q | 3.52 | 2.05 | 2.33 | 1.47 | 0.47 | −1.94 | 0.55 | 1.05 | 2.31 |
| Jevons | P | 4.83 | 8.07 | 14.06 | 19.93 | 26.14 | 32.62 | 41.53 | 46.41 | 54.67 |
| | Q | 3.17 | 0.78 | 0.61 | −0.4 | −2.13 | −4.6 | −2.14 | −2.41 | −0.96 |
| Harmonic | P | 4.67 | 7.74 | 13.44 | 19.22 | 25.01 | 31.24 | 39.48 | 44.32 | 52.22 |
| | Q | 3.32 | 1.09 | 1.16 | 0.2 | −1.24 | −3.6 | −0.7 | −1 | 0.63 |
| CSWD | P | 4.83 | 8.08 | 14.08 | 19.95 | 26.16 | 32.68 | 41.59 | 46.45 | 54.76 |
| | Q | 3.17 | 0.77 | 0.6 | −0.41 | −2.14 | −4.65 | −2.18 | −2.44 | −1.02 |
| Lasperyes | P | 6.33 | 8.71 | 14.59 | 19.79 | 25.85 | 32.17 | 41.5 | 45.83 | 53 |
| | Q | 1.71 | 0.18 | 0.15 | −0.28 | −1.9 | −4.28 | −2.12 | −2.02 | 0.12 |
| Paasche | P | 6.6 | 8.86 | 14.88 | 20.02 | 26.33 | 32.76 | 41.81 | 45.99 | 53.62 |
| | Q | 1.45 | 0.05 | −0.1 | −0.47 | −2.28 | −4.71 | −2.33 | −2.13 | −0.28 |
| Palgrave | P | 6.93 | 9.31 | 15.65 | 20.82 | 27.63 | 34.77 | 44.93 | 49.74 | 58.52 |
| | Q | 1.14 | −0.37 | −0.77 | −1.13 | −3.27 | −6.13 | −4.44 | −4.58 | −3.37 |
| Harmonic | P | 5.99 | 8.25 | 13.79 | 18.97 | 24.5 | 30.4 | 38.54 | 42.34 | 48.94 |
| Lasperyes | Q | 2.04 | 0.61 | 0.85 | 0.41 | −0.84 | −2.98 | −0.03 | 0.38 | 2.85 |
| Geometric | P | 6.16 | 8.48 | 14.19 | 19.38 | 25.18 | 31.26 | 40 | 44.04 | 50.88 |
| Lasperyes | Q | 1.87 | 0.4 | 0.5 | 0.06 | −1.38 | −3.62 | −1.07 | −0.81 | 1.53 |
| Geometric | P | 6.77 | 9.08 | 15.26 | 20.42 | 26.98 | 33.74 | 43.36 | 47.83 | 55.97 |
| Paasche | Q | 1.29 | −0.16 | −0.43 | −0.8 | −2.78 | −5.4 | −3.39 | −3.35 | −1.79 |
| Fisher | P | 6.46 | 8.78 | 14.73 | 19.9 | 26.09 | 32.47 | 41.66 | 45.91 | 53.31 |
| | Q | 1.59 | 0.12 | 0.03 | −0.37 | −2.09 | −4.5 | −2.23 | −2.08 | −0.08 |
| Tornqvist | P | 6.46 | 8.78 | 14.73 | 19.89 | 26.08 | 32.49 | 41.67 | 45.92 | 53.4 |
| | Q | 1.59 | 0.12 | 0.03 | −0.36 | −2.08 | −4.51 | −2.24 | −2.08 | −0.14 |
| Marshall-Edgeworth | P | 6.63 | 9.01 | 15.12 | 20.31 | 26.74 | 33.47 | 43.22 | 47.78 | 55.76 |
| | Q | 1.43 | −0.09 | −0.31 | −0.71 | −2.59 | −5.21 | −3.3 | −3.32 | −1.65 |
| Walsh | P | 6.47 | 8.8 | 14.78 | 19.92 | 26.19 | 32.68 | 42.21 | 46.54 | 54.09 |
| | Q | 1.58 | 0.1 | −0.02 | −0.39 | −2.17 | −4.65 | −2.61 | −2.5 | −0.59 |
| Drobsich | P | 6.47 | 8.78 | 14.73 | 19.9 | 26.09 | 32.47 | 41.66 | 45.91 | 53.31 |
| | Q | 1.58 | 0.12 | 0.03 | −0.37 | −2.09 | −4.5 | −2.23 | −2.08 | −0.08 |

## 4.5 Conclusions

We began this chapter by asking what a price index is and we have seen in the discussion that followed that an individual price index series represents estimates of a number which aggregates lots of information, usually regarding prices and quantities, the changes which then tell us about the rate of inflation. Identical data can be used to produce a wide range of index numbers, which means we potentially have a wide range of estimates of inflation. As we look more specifically at an individual inflation measure, the change in consumer price levels in the UK, we should be careful to remember that the index numbers produced are just one among many possibilities and that there are many ways to think about the measurement of price changes. As a result, there is no single answer to the question of what the value of the price index is in a given period and any price index we produce remains a single realisation using one among many potential methodologies.

## References


Coggeshall, F. (1886). The arithmetic, geometric, and harmonic means. *Quarterly Journal of Economics, 1,* 83–86.

Drobisch, M. W. (1871). Über die Berechnung der Veränderung der Waarenpreise und des Geldwerths. *Jahrbucher fur Nationalökonomie und Statistik, 16,* 416–427.

Elliott, D., & O'Neill, R. (2012). Estimating the elasticity of substitution for alcohol products. *Survey Methodology Bulletin 71,* 13–25.

Fisher, I. (1922). *The making of index numbers; a study of their varieties, tests, and reliability*. Boston: Houghton Mifflin.

Frisch, R. (1936). Annual survey of general economic theory: The problem of index numbers. *Econometrica, 4*(1), 1–38.

Jevons, W. S. (1863). *A serious fall in the value of gold ascertained, and its social effects set forth*. London: Edward Stanford.

Laspeyres, E. (1864). Hamburger Waarenpreise 1851–1863 und die californisch-australischen Goldentdeckungen seit 1848. Ein Beitrag zur Lehre von der Geldentwerthung. *Jahrbücher für Nationalökonomie und Statistik, 3*(81–118), 209–236.

Laspeyres, K. (1871). IX. die berechnung einer mittleren waarenpreissteigerung. *Jahrbücher FÃ¼r NationalÃ¶konomie Und Statistik, 16*(1), 296–318. doi:10.1515/jbnst-1871-0124

Lloyd, P. J. (1975). Substitution effects and biases in nontrue price indices. *The American Economic Review, 65*(3), 301–313.

Mehrhoff, J. (2010). Aggregate indices and their corresponding elementary indices. *Jahrbücher Für Nationalökonomie Und Statistik/Journal of Economics and Statistics, 230*(6), 709–725. doi:10.1515/jbnst-2010-0607.

Moulton, B. R., & Moses, K. E. (1997). Addressing the quality change issue in the consumer price index. *Brookings Papers on Economic Activity, 1997,* 305–366.

Paasche, H. (1874). Über die Preisentwicklung der letzen Jahre nach den Hamburger Börsennotirungen. *Jarbücher für Nationalökonomie und Statistik, 23,* 168–178.

Törnqvist, L. (1936). The bank of Finland's Consumption Price Index. *Bank of Finland Monthly Bulletin, 10,* 1–8.

Von Auer, L. (2010). Drobisch's legacy to price statistics. *Jahrbücher FÃ¼r Nationalökonomie Und Statistik/Journal of Economics and Statistics, 230*(6), 673–689.

Walsh, C. M. (1901). *The measurement of general exchange value*. New York: The Macmillan Company.

Walsh, C. M. (1921). The best form of index number: Discussion. *Quarterly Publications of the American Statistical Association, 17*(133), 537–544.

Young, A. (1812). *An enquiry into the progressive value of money in england, as marked by the price of agricultural products; with observations upon sir G. shuckburgh's table of appreciation*. London: Hatchard.